### D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

<table>
<thead>
<tr>
<th>Workpackage</th>
<th>WP4 - SKILLS: Skills, Education, and Centers of Excellence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Editor(s):</td>
<td>Niki Pavlopoulou (INSIGHT), Edward Curry (INSIGHT)</td>
</tr>
<tr>
<td>Responsible Partner:</td>
<td>INSIGHT</td>
</tr>
<tr>
<td>Contributors</td>
<td>Edo Osagie (INSIGHT), Dhaval Salwala (INSIGHT), Ernestina Menasalvas (UPM), Ana M. Moreno (UPM), Nik Swoboda (UPM), Aristide Rothweiler (UDE), Andreas Metzger (UDE), Angelo Giuliana (EIT DIGITAL), Susanne Kührer (EIT DIGITAL)</td>
</tr>
<tr>
<td>Internal Reviewer</td>
<td>ATOS, UDE</td>
</tr>
<tr>
<td>Status-Version:</td>
<td>Final – v3.0</td>
</tr>
<tr>
<td>Due to</td>
<td>M36</td>
</tr>
<tr>
<td>Submission Date</td>
<td>19/12/2019</td>
</tr>
<tr>
<td>EC Distribution:</td>
<td>Public</td>
</tr>
</tbody>
</table>

**Abstract:**
This document contains the Work package 4 (WP4) progress of activities from Month 19 to Month 36 (M19 to M36) of the Big Data Value ecosystem (BDVe) project according to the objectives set for this period. The main outcomes include an enhanced best practice guide, success stories with supporting new Centres of Excellence, an enhanced Education Hub, the pilot of Data Science Analytics badges, internship portal and mobility framework, as well as dissemination activities, like white papers and publications.
Contents

1 INRODUCTION ....................................................................................................................... 17
   1.1 WP4 OBJECTIVES ........................................................................................................... 17
   1.2 PURPOSE AND TARGET GROUP OF THE DELIVERABLE ........................................... 18
   1.3 DOCUMENT OUTLINE .................................................................................................... 18

2 KEY ACHIEVEMENTS ......................................................................................................... 20
   2.1 WP4 KEY COLLABORATIONS ......................................................................................... 20
   2.2 WP4 KEY IMPACTS ....................................................................................................... 21
       2.2.1 T4.1 Network of National BDV Centres of Excellence Key Impacts ............... 21
       2.2.2 T4.2 BDV EduHub Key Impacts ........................................................................... 23
       2.2.3 T4.3 Recognition of Skills for Big Data Professionals Key Impacts .............. 24
       2.2.4 T4.4 Data Scientist Mobility Programme Key Impacts ................................... 24

3 GENERAL ACTIVITIES ......................................................................................................... 26
   3.1 WP4 MONTHLY MEETINGS ........................................................................................... 26
   3.2 CONSORTIUM MEETINGS ............................................................................................ 27
   3.3 EVENTS AND EXTERNAL DISSEMINATION ................................................................. 28
       3.3.1 Presentation of the BDV badge programme and participation in the
            EDISON EDSF Release 4 Design Workshop 2018 ...................................................... 31
       3.3.2 Presentation of data science skills recognition framework at BDVA
            Activity Group Meeting (AG28) .............................................................................. 32
       3.3.3 Panel on “From Big Data to AI and Beyond” at EBDVF 2018 ......................... 32
       3.3.4 Workshop in “Shaping the European Big Data landscape: from i-Spaces
            and Centers of Excellence to Big Data Digital Innovation Hubs” at EBDVF 2018 33
       3.3.5 Participation in “Data Science Skills for Society: From Big Data to AI,
            Robotics and Beyond” at ICT 2018: Image Digital - Connect, Europe ................. 34
       3.3.6 Presentation in the SME Alliance Workshop: Boosting Security, Data and
            IoT Skills – Workshop #4 ..................................................................................... 35
       3.3.7 Presentation “Data 4 AI: For European Economic Competitiveness and
            Societal Progress” at European Industry Partnerships Collaborative Event 2019 36
       3.3.8 Participation in “Artificial Intelligence Session” and presentation on
            “Artificial Intelligence PPP Vision” at BDV PPP Summit 2019 .............................. 36
       3.3.9 Panel on “Towards an Ecosystem for Artificial Intelligence Value-
            Creation” at BDV PPP Summit 2019 .................................................................. 37
3.3.10 Panel on “Addressing the Growing Need for Skills in Data Science” at the BDV PPP Summit 2019 ................................................................. 38
3.3.11 Poster presentation of “A Proposal for Recognizing Skills in Data Science Using Open Badges” at the Conference on Innovation and Technology in Computer Science Education (ITiCSE) 2019 .................................................. 38
3.3.12 DataSci Awards ............................................................................. 40
3.3.13 Participation in the Round Table: “Reveal how more and better education and training offers can be tailored” at European Conference SME Alliance 41
3.3.14 Publication of “Addressing the Growing Need for Skills in Data Science” in the KDnuggets blogpost ................................................................. 42
3.3.15 Participation in the Steering Committee of the SME Alliance .......... 42
3.3.16 Presentation of the data science recognition programs from T4.3 and participation in the EDISON EDSF Release 4 Design Workshop ......................... 43
3.3.17 Presentation and discussion of work to recognise data science skills at a BDVA Activity Group Meeting (AG35) ...................................................... 43

4 NETWORK OF NATIONAL BDV CENTRES OF EXCELLENCE (T4.1-INSIGHT) ........................................................................................................ 45

4.1 WHITE PAPERS .................................................................................. 45
4.2 SHARING BEST PRACTICES ............................................................... 45
4.2.1 Best Practice Framework Overview ................................................. 46
4.2.2 Deep Dive on Capabilities ............................................................... 49
4.2.3 Critical Success Factors ................................................................. 49
4.2.3.1 Challenges ................................................................................. 50
4.2.3.2 Success Factors ......................................................................... 51
4.2.3.3 Mechanisms to Address Challenges ......................................... 52
4.2.3.4 Ideal Situation ........................................................................... 54
4.3 ENGAGEMENT LIFECYCLE ............................................................... 54
4.4 SUPPORTING THE ESTABLISHMENT OF NEW CENTRES OF EXCELLENCE .............. 57
4.4.1 GATE – Sofia Bulgaria - Success Story ........................................... 58
4.4.2 Reaching out Through IDEAL-IST ................................................ 59
4.5 NETWORK OF EXCELLENCE COLLABORATIONS .......................... 59

5 BIG DATA VALUE EDUCATION HUB (T4.2-UDE) .................................. 62
5.1 EDUCATION PROGRAMMES ............................................................. 62
5.1.1 Master and PhD Programmes ......................................................... 63
5.1.2 Professional Education Programmes .......................................................... 64
5.2 Dissemination Activities .............................................................................. 65

6 THE RECOGNITION OF SKILLS FOR BIG DATA PROFESSIONALS (T4.3-UPM) ........................................................................................................ 67

6.1 Task 4.3 Objectives .................................................................................... 67
6.2 A Brief Review of Activities in Period 1 .................................................. 67
6.3 Comments Received from the Reviewers and Their Impact on Work Proposed in Periods 2 and 3 ................................................................. 68
6.4 Summary of the Progress in Period 2 ....................................................... 70
6.5 A Badge-Based Framework for Recognising Data Science Skills in Formal Education .................................................................................. 71

6.5.1 The First Phase of the Pre-Pilot of the Data Science Analytics Badge ... 73
6.5.1.1 Feedback from Academia ................................................................. 74
6.5.1.2 Feedback from Industry ................................................................. 78
6.5.1.3 The September 2018 BDVA Activity Group Meeting (AG28) ....... 88
6.5.1.4 Version 0.2 of the Data Science Analytics Badge: Levels and Requirements ......................................................................................... 91

6.5.2 The Second Phase of the Pre-Pilot of the Data Science Analytics Badge 95
6.5.2.1 Feedback from Academia ................................................................. 96
6.5.2.2 Feedback from the BDVA Skills Task Force .................................. 98
6.5.2.3 Version 1.0 of the Data Science Analytics Badge ....................... 99

6.5.3 The Pilot of the Academic Level of the Data Science Analytics Badge 101
6.5.3.1 Applying to Issue the Badge ............................................................. 102
6.5.3.2 Reviewing Applications for Issuing the Badge .............................. 104
6.5.3.3 Web Site for the Pilot for Issuing the Data Science Analytics Badge 105
6.5.3.4 Running the Pilot for Issuing the Academic Level of the Data Science Analytics Badge ............................................................................. 108
6.5.3.5 Results of the Pilot for Issuing the Academic Level of the Data Science Analytics Badge ............................................................................. 110

6.5.4 The Platform for Issuing BDV Data Science Badges ............................ 110
6.5.4.1 A Quick Introduction to v2.0 of the Open Badge Standard ............ 110
6.5.4.2 Recommendations for the production of BDV Badges ................. 117
6.5.4.3 Description of how the badge issuing platform will be used ....... 118
6.5.4.4 The badge issuing platform’s functional requirements ..........119
6.5.4.5 Basic technical requirements for the badge issuing platform .....119
6.5.4.6 A review of all Open Badge v2.0 certified badge issuing platforms 120
6.5.4.7 The evaluation of the platforms........................................124
6.5.4.8 Instructions for Using the Badge Issuing Platform .................125
6.5.5 First Open Call for Issuing the Academic Level of the Data Science Analytics Badge................................................................125
6.5.6 Meetings with External Stakeholders ......................................126
6.5.6.1 Virtual Meeting with Warsaw University of Technology ........126
6.5.6.2 Virtual Meeting with the University of Adger, Norway (18/01/2019) 126
6.5.6.3 BDVA Tasks force on Skills.........................................................127
6.5.6.4 Workshop SME Alliance: Boosting Security, Data and IoT Skills – Workshop #3.................................................................127
6.5.6.5 Virtual Meeting with Politecnico di Milano.............................128
6.5.6.6 Virtual meeting with Young Universities Research Network (YERUN) 128
6.6 Endorsements for Non-Formal Training in Data Science .............128
6.6.1 Challenges when Considering Non-Formal Education ...............129
6.6.2 Timeline for the Implementation of an Endorsement System for Non-formal Education in Data Science............................................130
6.6.3 A Starting Point for the Endorsement System............................131
6.6.4 A Preliminary Proposal for an Endorsement System for Non-Formal Education in Data Science ..................................................133
6.6.5 The logistics of the Endorsement System....................................134
6.6.6 Work Planned in Period 3 .............................................................135
6.6.7 Meetings with External Stakeholders ........................................135
6.6.7.1 EBDVF 2018 – Meeting with Nokia...........................................135
6.6.7.2 Virtual Meeting with Big Data Lab – Corsi ................................136
6.6.7.3 Virtual Meeting with EIT Digital Professional School ..............136
6.6.7.4 Meeting with Big Data Lab – Corsi .............................................137
6.6.7.5 EBDVF 2019 - Meeting with Nokia .............................................137
A. HOW TO COMPLETE THE TEMPLATE FOR ISSUING APPLICATIONS .................................. 207
V. GENERAL INFORMATION .................................................................................................................. 207
VI. ACQUISITION OF SKILLS ............................................................................................................... 207

EVALUATION GUIDELINES FOR THE DATA SCIENCE ANALYTICS BADGE .............................. 209
A. CRITERIA FOR EVALUATION ........................................................................................................... 209
B. THE FINAL DECISION ................................................................................................................... 210

THE REVISION OF APPLICATIONS ........................................................................................................ 210

APPEALS ................................................................................................................................................. 210
A. ARTICLE 1 – PERFORMANCE OF THE REVIEW ............................................................................... 211
B. ARTICLE 2 - OBLIGATIONS OF IMPARTIALITY ............................................................................... 211
C. ARTICLE 3 - OBLIGATIONS OF CONFIDENTIALITY ......................................................................... 213
D. SIGNATURE ........................................................................................................................................ 213
A. SUMMARY OF THE EVALUATION ...................................................................................................... 216
B. DETAILED EVALUATION .................................................................................................................. 216

INTRODUCTION ....................................................................................................................................... 221

REQUIREMENTS FOR REVIEWERS ......................................................................................................... 221

EVALUATION GUIDELINES FOR THE ACADEMIC LEVEL OF THE DATA SCIENCE ANALYTICS BADGE ...................................................................................................................... 221
A. GENERAL GUIDELINES ..................................................................................................................... 222
B. CRITERIA FOR EVALUATION ............................................................................................................. 222
C. THE FINAL DECISION ...................................................................................................................... 224
A. INTRODUCTION ................................................................................................................................ 228
B. WHAT IS AN OPEN BADGE? .............................................................................................................. 228
C. REGISTERING WITH THE OPEN BADGE FACTORY ......................................................................... 228
D. SETTING UP THE BADGE FOR ISSUE ............................................................................................... 230
E. ISSUING BADGES ............................................................................................................................... 233
F. RECEIVING A BADGE ......................................................................................................................... 234
G. FINAL CONSIDERATIONS ................................................................................................................... 235
A. CREATOR ........................................................................................................................................... 236
B. GENERAL DESCRIPTION .................................................................................................................... 236
C. REQUIRED SKILLS ............................................................................................................................ 237
D. CREDITS ............................................................................................................................................ 237
A. MOBILITY FRAMEWORK SURVEY TEMPLATE .................................................................................. 238
B. MOBILITY FRAMEWORK PROJECTS CANDIDATE ........................................................................... 240
A. FUNCTIONAL SPECIFICATION FOR THE INTERNSHIP PORTAL ...................................................... 241
List of Figures

Figure 1: Panel on “From Big Data to AI and Beyond” at EBDVF 2018 ................. 33
Figure 2: Workshop in “Shaping the European Big Data Landscape: from i-Spaces and Centers of Excellence to Big Data Digital Innovation Hubs” at EBDVF 2018. 34
Figure 3: Skills Session in the Riga BDV PPP Summit ........................................ 38
Figure 4: Poster of “A Proposal for Recognizing Skills in Data Science Using Open Badges” ................................................................. 39
Figure 5: DatSci Awards 2019 ........................................................................... 41
Figure 6: SME Alliance Round Table Discussion ................................................. 42
Figure 7: BDVA Activity Group Meeting (AG35) .................................................. 44
Figure 8: BDAICoE Framework ........................................................................ 46
Figure 9: Engagement Lifecycle for BDAICoEs ................................................... 54
Figure 10: Geographic heat map with all European Big Data Centres of Excellence 57
Figure 11: GATE Director Prof. Silvia Ilieva with EU Commissioner Mariya Gabriel at the Launch of GATE ......................................................... 59
Figure 12: Education Hub Landing Page ............................................................... 63
Figure 13: Education Programme Profile Information ........................................ 64
Figure 14: Design of the Data Science Analytics Badge V0.1 ........................... 68
Figure 15: Summary of Task 4.3’s Outcomes in Periods 1-3 ............................ 70
Figure 16: Work on the Design of the Data Science Analytics Badge .......... 71
Figure 17: The relevance of basic level skills for Industry 1 ............................ 78
Figure 18: The relevance of intermediate-level skills for Industry 1 ........... 79
Figure 19: The relevance of expert level skills for Industry 1 ..................... 79
Figure 20: The relevance of Basic Level skills for Industry 2 .................... 81
Figure 21: The relevance of Basic Level skills for Industry 3 .................... 82
Figure 22: The relevance of Basic Level skills for Industry 4 .................... 84
Figure 23: The relevance of Intermediate Level skills for Industry 4 ........ 84
Figure 24: Relevance of Expert Level skills for Industry 4 ...................... 85
Figure 25: The relevance of Basic Level skills for Industry 5 .................... 86
Figure 26: The relevance of Intermediate Level skills for Industry 5 .......... 87
Figure 27: The relevance of Expert Level skills for Industry 5 .................. 87
Figure 28: Data Science Analytics Badges with academic and professional levels (v1.0) .................................................................................. 101
Figure 29: Badges Application and Issuing Process ........................................ 101
Figure 30: Summary of the process for applying and issuing academic badges.... 102
Figure 31: Screenshot of the skills recognition program web site with the home 105
Figure 32: Deadline and eligibility details in the Skills Recognition Program ... 106
Figure 33: Information about the documents for applicants and reviewers ...... 107
Figure 34: Screenshot of the first part of the FAQ ........................................... 107
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

List of Tables

TABLE 1: DEFINITIONS, ACRONYMS AND ABBREVIATIONS ................................................. 14
TABLE 2: WP4 MEETINGS SCHEDULE .................................................................................... 27
TABLE 3: BDVe Consortium MEETINGS SCHEDULE .............................................................. 28
TABLE 4: EVENTS AND EXTERNAL DISSEMINATION ACTIVITIES OF BDVe WP4 .............. 31
TABLE 5: CORE OPERATIONAL CAPABILITIES OF THE BDAICoE FRAMEWORK .......... 48
TABLE 6: SUMMARY OF CHALLENGES ................................................................................. 51
TABLE 7: SUMMARY OF SUCCESS FACTORS ....................................................................... 52
TABLE 8: SUMMARY OF MECHANISMS TO ADDRESS CHALLENGES .............................. 53
TABLE 9: SUMMARY OF IDEAL SITUATIONS ....................................................................... 54
TABLE 10: TIMELINE OF TASK 3.4 FOR PERIODS 2 AND 3 ................................................ 69
TABLE 11: DETAILED TIMELINE FOR ACTIVITIES RELATED TO FORMAL EDUCATION .... 72
TABLE 12: REQUIREMENTS OF THE DATA SCIENCE ANALYTICS BADGE (v 0.1) ............ 73
TABLE 13: REPLIES FROM PEOPLE WITH BASIC, INTERMEDIATE AND EXPERT LEVELS ...... 83
TABLE 14: CONSOLIDATION OF RECOMMENDATIONS PROVIDED BY ACADEMIA AND INDUSTRY 92
TABLE 15: THE FINAL LIST OF RECOMMENDATIONS TO BE CONSIDERED IN THE REVISIOn OF THE DATA SCIENCE ANALYTICS BADGE .......................................................................... 94
TABLE 16: NEW FORMULATION OF DATA SCIENCE ANALYTICS BADGE (v 0.2) ............... 95
TABLE 17: SCHEDULE FOR THE SECOND PHASE OF THE PRE-PILOT ................................. 96
TABLE 18: ALTERNATIVES FOR DSA.1 AND DSA.2 OF THE DATA SCIENCE ANALYTICS BADGE 99
TABLE 19: DATA SCIENCE ANALYTICS BADGE VERSION 1.0 ........................................... 100
TABLE 20: MATCHING BETWEEN RECOMMENDATIONS AND V1.0 OF THE BADGE REQUIREMENTS .............................................................. 100
TABLE 21: EXTRACT OF THE TEMPLATE TO COLLECT EVIDENCES RELATED TO THE DATA SCIENCE ANALYTICS BADGE, FIRST SKILL ....................................................... 103
TABLE 22: THE PILOT TIMELINE ........................................................................................ 108
TABLE 23: TASK 4.3. TIMELINE FOR ACTIVITIES RELATED TO NON-FORMAL EDUCATION ...... 131
TABLE 24: EXAMPLE OF QUALITY INDICATORS 4 .............................................................. 133
TABLE 25: SUMMARY OF STRATEGY PRACTICES ................................................................ 151
TABLE 26: SUMMARY OF GOVERNANCE .......................................................................... 152
TABLE 27: SUMMARY OF STRUCTURE ............................................................................... 153
TABLE 28: SUMMARY OF FUNDING .................................................................................. 154
TABLE 29: SUMMARY OF PEOPLE .................................................................................... 155
TABLE 30: SUMMARY OF CULTURAL PRACTICES .............................................................. 156
TABLE 31: SUMMARY OF BUSINESS DEVELOPMENT ...................................................... 158
TABLE 32: SUMMARY OF COLLABORATION ..................................................................... 159
TABLE 33: SUMMARY OF TECHNICAL INFRASTRUCTURE ................................................ 160
TABLE 34: GOOD PRACTICES IN RESEARCH SUPPORT SERVICES .................................... 162
TABLE 35: GOOD PRACTICES IN IP AND DP .................................................................... 165
**Table 36:** Good practices in Policy Outreach .................................................. 166
**Table 37:** Summary of Public Outreach ............................................................... 169
**Table 38:** Good practices in technology and knowledge transfer ...................... 171
**Table 39:** Good practices in performance and impact assessment ...................... 173
**Table 40:** BDV Data Science Analytics Badge Skills ......................................... 195
**Table 41:** BDV Data Science Analytics Badge Skills ......................................... 206
**Table 42:** BDV Data Science Analytics Badge Skills ......................................... 221
## Definitions, Acronyms and Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG</td>
<td>Activity Group</td>
</tr>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>BBDC</td>
<td>Berlin Big Data Centre</td>
</tr>
<tr>
<td>BD</td>
<td>Business Development</td>
</tr>
<tr>
<td>BDAICoE</td>
<td>Big Data and Artificial Intelligence Centre of Excellence</td>
</tr>
<tr>
<td>BDV</td>
<td>Big Data Value</td>
</tr>
<tr>
<td>BDVA</td>
<td>Big Data Value Association</td>
</tr>
<tr>
<td>BDVe</td>
<td>Big Data Value ecosystem</td>
</tr>
<tr>
<td>CoE</td>
<td>Centre of Excellence</td>
</tr>
<tr>
<td>CSC</td>
<td>Centre Steering Committee</td>
</tr>
<tr>
<td>D</td>
<td>Deliverable</td>
</tr>
<tr>
<td>DP</td>
<td>Data Protection</td>
</tr>
<tr>
<td>DSM</td>
<td>Digital Single Market</td>
</tr>
<tr>
<td>DTI</td>
<td>Department of Trade and Industry</td>
</tr>
<tr>
<td>EBDVF</td>
<td>European Big Data Value Forum</td>
</tr>
<tr>
<td>EC</td>
<td>European Commission</td>
</tr>
<tr>
<td>EDSA</td>
<td>European Data Science Academy</td>
</tr>
<tr>
<td>EDSF</td>
<td>EDISON Data Science Framework</td>
</tr>
<tr>
<td>EduHub</td>
<td>Education Hub</td>
</tr>
<tr>
<td>EPE</td>
<td>Education and Public Engagement</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FDI</td>
<td>Foreign Direct Investment</td>
</tr>
<tr>
<td>GA</td>
<td>General Assembly</td>
</tr>
<tr>
<td>GDPR</td>
<td>General Data Protection and Regulation</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technologies</td>
</tr>
<tr>
<td>Acronym</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>IIOT</td>
<td>Industrial Internet of Things</td>
</tr>
<tr>
<td>IoT</td>
<td>Internet of Things</td>
</tr>
<tr>
<td>IP</td>
<td>Intellectual Property</td>
</tr>
<tr>
<td>IPC</td>
<td>Intellectual Property Committee</td>
</tr>
<tr>
<td>IRI</td>
<td>Internationalised Resource Identifier</td>
</tr>
<tr>
<td>IT</td>
<td>Information Technology</td>
</tr>
<tr>
<td>ITiCSE</td>
<td>Innovation and Technology in Computer</td>
</tr>
<tr>
<td></td>
<td>Science Education</td>
</tr>
<tr>
<td>IUA</td>
<td>Irish Universities Association</td>
</tr>
<tr>
<td>JWS</td>
<td>JSON Web Signature</td>
</tr>
<tr>
<td>KPIs</td>
<td>Key performance indicators</td>
</tr>
<tr>
<td>LGPL</td>
<td>Lesser General Public License</td>
</tr>
<tr>
<td>M</td>
<td>Month</td>
</tr>
<tr>
<td>MIS</td>
<td>Management Information Systems</td>
</tr>
<tr>
<td>MOOCs</td>
<td>Massive Open Online Course</td>
</tr>
<tr>
<td>MSc</td>
<td>Master</td>
</tr>
<tr>
<td>NCoE</td>
<td>Network of Centre of Excellence</td>
</tr>
<tr>
<td>NDAs</td>
<td>Non-Disclosure Agreements</td>
</tr>
<tr>
<td>NUIIG</td>
<td>National University of Ireland Galway</td>
</tr>
<tr>
<td>O</td>
<td>Objective</td>
</tr>
<tr>
<td>PhD</td>
<td>Doctor of Philosophy</td>
</tr>
<tr>
<td>PMO</td>
<td>Project Management Office</td>
</tr>
<tr>
<td>PPP</td>
<td>Public-Private Partnership</td>
</tr>
<tr>
<td>R&amp;D&amp;I</td>
<td>Research, Development and Innovation</td>
</tr>
<tr>
<td>RIN</td>
<td>Research Information Network</td>
</tr>
<tr>
<td>RRI</td>
<td>Responsible Research and Innovation</td>
</tr>
<tr>
<td>RSS</td>
<td>Research Support Services</td>
</tr>
<tr>
<td>SB</td>
<td>Strategy Board</td>
</tr>
<tr>
<td>Acronym</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>SIC</td>
<td>Scalable Innovation Cycle</td>
</tr>
<tr>
<td>SMEs</td>
<td>Small and medium-sized enterprises</td>
</tr>
<tr>
<td>T</td>
<td>Task</td>
</tr>
<tr>
<td>TF</td>
<td>Task force</td>
</tr>
<tr>
<td>TTOs</td>
<td>Technology Transfer Offices</td>
</tr>
<tr>
<td>UCC</td>
<td>University College Cork</td>
</tr>
<tr>
<td>UCD</td>
<td>University College Dublin</td>
</tr>
<tr>
<td>UIC</td>
<td>Universities-Industry Collaboration</td>
</tr>
<tr>
<td>URI</td>
<td>Uniform Resource Identifier</td>
</tr>
<tr>
<td>UUID</td>
<td>Universally Unique Identifier</td>
</tr>
<tr>
<td>WIPO</td>
<td>World Intellectual Property Organisation</td>
</tr>
<tr>
<td>WP</td>
<td>Work package</td>
</tr>
<tr>
<td>YERUN</td>
<td>Virtual meeting with Young Universities Research Network</td>
</tr>
</tbody>
</table>

Table 1: Definitions, Acronyms and Abbreviations
Executive Summary

The goal of the Big Data Value ecosystem (BDVe) is to support the Big Data Value Public-Private Partnership (BDV PPP) in realising a vibrant data-driven EU economy. WP4 “SKILLS: Skills, Education, and Centers of Excellence” tackles the important issue of Big Data Skills and Education within Europe. BDVe cannot transform the Skills and Education landscape of Europe alone, thus this WP emphasises on collaboration with existing initiatives to build on their progress and work in partnership. WP4 links with National BDV Centres of Excellence (CoE) to foster collaboration and support the establishment of new centres. To provide targeted access to offerings concerning European universities’ data scientist and data engineering programmes, WP4 established the BDV Education Hub (EduHub) to facilitate knowledge exchange on educational programmes. In addition, industrial needs are also met. BDVe leverages existing efforts by EIT Digital on its Master (MSc) and professional courses and collaborates with the European Data Science Academy (EDSA) and EDISON projects on Data Science Curricula. Finally, a mobility framework and internship portal for data scientists and professionals are being established, leveraging existing European and national initiatives.

D4.3 covers the WP4 progress from M19 to M36 of the BDVe project.

▪ T4.1 “Network of National BDV Centres of Excellence” led by INSIGHT, in the first period, conducted a survey on CoEs and defined a best practice guide based on interviews from different CoEs. In this period, the best practice guide was refined to reflect the increasing importance of Artificial Intelligence as a topic for Big Data CoEs. The guide was extended with additional knowledge and has been successfully tested on GATE, the first Big Data CoE in Bulgaria. Further work focused on a cross-case analysis for the identification of common best practices and critical success factors which have been captured in a white paper. The Network of CoE have collaborated together in a number of areas including site visits for best practice sharing and in the development of a number of funding proposals involving multiple members of the network. Finally, we have developed an Engagement Lifecycle to support a systematic approach to supporting CoEs at different stages of development (pre-start-up, Early Stage, and Mature).

▪ T4.2 “Big Data Value Education Hub” led by UDE, in the first period, created an EduHub with 166 MSc courses in BDV. In this period, the EduHub was enhanced with 181 MSc courses, 12 Doctor of Philosophy (PhDs), and 62 on-site/online professional training programmes. The EduHub is hosted as part of the BDV PPP web portal. Details are reported in D4.5 [6].

▪ T4.3 “The Recognition of Skills for Big Data Professionals” led by the UPM, in the first period, conducted a review of existing recognition frameworks for skills in data science and proceeded with Open Badges. After considering
different options, the initial version of the types and requirements for the badges was taken from work from the EDISON project. In this period, it conducted a review of the plan for the badges gathering feedback from academia and industry, it designed its first pilot of the Academic level of the Data Science Analytics Badge, and it is currently offering that badge to academia. Also, it began work on a second skills recognition system focusing on non-formal training. Details are reported in D4.6 [12].

- T4.4 “Data Scientist Mobility Programme” led by EIT DIGITAL, in the first period, conducted a survey on existing mobility programmes and it underwent an initial design for portal for internship positions and mobility programmes at the industry. After considering different options for the internship portal, it was decided to focus only on the supply side of the open Big Data internship position. In this period, a survey was conducted that led to a high interest in a mobility framework and internship portal, and initial plans were made on the future pilot programme. Additional solutions for the internship portal are investigated using social channels. Details on the scope and objectives are reported in D4.7.
1 Introduction

The goal of BDVe is to support the BDV PPP in realising a vibrant data-driven European Union (EU) economy. The main priorities of the project are:

- Being informed on the most important facts in Big Data for corresponding decision-making in the BDV PPP
- Supporting the implementation of the BDV PPP from an operational point of view
- Developing a vibrant community around the BDV PPP
- Supporting the development of a European network of infrastructures and centres of excellence around Big Data
- Setting up a professional Communications strategy
- Setting up a framework that supports the acceleration of data-driven businesses
- Ensuring the sustainability of the investments and actions triggered by the BDV PPP

WP4 is about “SKILLS: Skills, Education, and Centers of Excellence”. This work package tackles the important issue of Big Data Skills and Education within Europe. BDVe recognises that it cannot transform the Skills and Education landscape of Europe alone; thus, a key principle of this WP is a strong emphasis on collaboration with existing initiatives to build on their progress and work in partnership. In particular, WP4 links with National BDV Centres of Excellence to foster collaboration and support the establishment of new centres. To provide targeted access to offerings concerning European universities' data scientist and data engineering programmes as well as commercial training offerings, WP4 established the BDV EduHub to facilitate knowledge exchange on educational programmes. In addition, BDVe works with industry to ensure Data Skills education is meeting industrial needs. BDVe leverages existing efforts by EIT Digital on its MSc and professional courses and collaborates with former members of the EDISON project on Data Science Curricula. Finally, a mobility framework and internship portal for data scientists and professionals are being established, leveraging existing European and national mobility programmes and initiatives, as well as internship positions.

1.1 WP4 Objectives

The objectives of WP4 are to develop skills, education, and Centres of Excellence around Big Data that facilitate European-level coordination, help to align curricula with industry needs, and accelerate skills development, thus increasing the number of European data scientists by 2020. To this end, the overall objectives are:
O4.1: To establish a Network of National BDV Centres of Excellence to foster collaboration and share best practices between existing centres and support the establishment of new centres (Task (T) 4.1).

O4.2: To exchange knowledge on data scientist educational programmes across Europe by delivering a BDV EduHub as a platform and living repository for such knowledge (T4.2).

O4.3: To establish a Big Data skills recognition framework for BDV professionals to ensure the programmes’ alignment with industry needs (T4.3).

O4.4: To stimulate and promote mobility of students, confirmed data professionals and domain experts, and mobility opportunities beyond the BDV PPP such as Industrial Internships (T4.4).

1.2 Purpose and Target Group of the Deliverable

The aim of this deliverable (D4.3) is to provide a detailed report on the activities and in particular the results and offerings of all WP4 tasks including reports on (T4.1) Network of National BDV Centres of Excellence, (T4.2) Big Data Value Education Hub, (T4.3) The Recognition of Skills for Big Data Professionals and (T4.4) Data Scientists Mobility Programme. The target groups of the deliverable include Managers of Centres of Excellence (T4.1), Educators and Industry with interest in skills recognition and promotion of their educational programmes. (T4.2, T4.3) and members of the BDV PPP and external actors with interest in Mobility opportunities in Big Data (T4.4).

1.3 Document Outline

The remainder of this document is organised as follows:

- Section 1 describes an introduction on BDVe and WP4, WP4 objectives, purpose, and target group of the deliverable, links of WP4 to other project activities, WP4 key collaboration and impacts, and document outline.
- Section 2 describes the key achievements of WP4, including its main outcomes, key collaborations and impacts.
- Section 3 describes the general activities of WP4, like internal meetings, events, and external dissemination, WP4 monthly meetings, and management activities.
- Sections 4-7 provide a more detailed description of the individual tasks:

---

1 Only a summary is provided here, as details are provided in D4.5 [6]
o Section 4 describes all the achievements of T4.1, like updates on the best practice guide, supporting new Centres of Excellence, collaborations, and meetings that took place with the network of Centres of Excellence and dissemination activities.

o Section 5 describes all the achievements of T4.2, i.e., it describes the current status of the EduHub, its courses, and training programmes, as well as the events during which the EduHub was presented.

o Section 6 describes all the achievements of T4.3. It begins with a brief review of the activities from the previous period. Then it addresses issues raised by the reviewers at the end of that period. Next, it describes work from this period, including the phases of the pre-pilot of the badges and the feedback received from industrial and academic stakeholders. It also details the pilot of the Academic Level of the Data Science Analytics Badge, how it was disseminated, and the selection process for the platform for issuing the badges. Finally, an initial proposal of the endorsements for non-formal training in data science is given along with details of future work.

o Section 7 describes all the achievements of T4.4, like the data scientist mobility programme, a survey on interest for the mobility programme and its implementation plan. It also includes the definition, design, and implementation of the internship portal and a survey on the interest that took place and its results.

- Section 8 provides a conclusion for the deliverable.
2 Key Achievements

This section outlines the key collaborations and impacts of WP4, analysed below.

2.1 WP4 Key Collaborations

BDVe recognises that it cannot transform the Skills and Education landscape of Europe alone; thus, a key principle of this WP is a strong emphasis on collaboration with the existing initiatives to build on their progress and work in partnership. In this period, WP4 has continued its working relationships with the BDV Association (BDVA), European Centre of Excellence (CoE) Network, GATE, EDSA, and EDISON, and it has established new working relationships with the European DatSci & AI awards community.

- BDVe worked to gain support for the BDVA to join the Digital Skills and Jobs Coalition. This was approved by the Board of Directors and the General Assembly of the BDVA in December 2018.
- Network of Centres of Excellence is linked to WP4 through INSIGHT and Task 4.1. The purpose of the network is to boost collaboration and define the data-driven future of Europe. INSIGHT (T4.1) is part of the core group of this network.
- GATE is linked to WP4 though INSIGHT. GATE is the first Big Data CoE in Bulgaria. INSIGHT (T4.1) found that there is only one CoE in Eastern Europe, and since they are creating a best practice guide, they supported GATE, and at the same time, they tested the guide. Numerous meetings and workshops have been held between INSIGHT and GATE, with INSIGHT serving on the Advisory Board for the Centre.
- The European DatSci & AI awards is linked to WP4 through INSIGHT. This event gives awards through competitions that demonstrate and celebrate the power of
Data Science and Artificial Intelligence (AI) in driving real business, educational, and social outcomes. It is a place where research and applications of Data Science & AI are showcased and recognised. This event used to focus on awards only in Ireland, but with the assistance of INSIGHT, a partnership was established with the BDVA to celebrate the Data Science Ecosystem across the whole of Europe. This is an annual event; therefore, more of these events will be organised in the future.

- **Ideal-ist** is linked to WP4. We are engaged with the Ideal-ist network of National Contact Points for Information and Communication Technologies (ICT) in H2020 to promote the activities of BDVe and WP4 to Member states. This includes making them aware of tools such as the EduHub and the Best Practice Guide for the Centres of Excellence, to the developments with the Skills Badges. Ideal-ist offers us a knowledgeable local contact that can help us to connect to all member states within the EU.

- **EDISON** is linked to WP4 through the UPM. The use of the EDISON EDSF in both of the skills recognition systems being developed are the main focus of this collaboration with members of the EDISON project.

- **BDVA Skills Task Force** is linked to WP4 through the UPM. Input from both industrial and academic members of the BDVA Skills Task Force has been requested to get feedback regarding the progress on both of the skills recognition systems being developed.

- **EIT Digital Professional School** is linked to WP4 through the UPM. Collaboration with the EIT Digital Professional School in their role as non-formal education providers (their online courses). In particular, we hope that they will actively participate in the development of the endorsement system for non-formal training in data science.

### 2.2 WP4 Key Impacts

All tasks in WP4 delivered key impacts as described below.

#### 2.2.1 T4.1 Network of National BDV Centres of Excellence Key Impacts

As aforementioned, T4.1 goal is to create a Big Data and AI Centre of Excellence (BDAI-CoE) framework based on surveys & interviews from CoEs, to sustain a Network of Excellence via meetings and to support the establishment of new CoEs in Europe. Therefore, the key impacts of T4.1 are:

- Enhanced the CoE Best Practice Framework as follows
  - Extended scope to include relevant AI topics for Big Data CoEs
  - Extensive cross-case analysis of best practices in centres to define a deep dive on specific capabilities
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

- Deep-dive Case Studies of 3 Successful CoEs
- Interviews held with 15+ CoE Managers/Leaders
- Development of an Engagement Lifecycle to support a systematic approach to supporting CoEs at different stages of development (pre-start-up, Early Stage, and Mature).

### Supporting Network of Centres of Excellence

- A series of formal and informal meetings of the network has taken place in period 2. These include the BDV Summit, European Big Data Value Forum (EBDVF), ICT Conference, ICT Proposers Day. Meetings hosted with the core group of the network.
- Collaboration in a number of areas including site visits for best practice sharing and in the development of a number of funding proposals involving multiple members of the network.

### Continued engagement with a new Big Data Centre in Bulgaria (GATE):

- The Gate centre was formally established in 2019, and support from BDVe was noted as an important support.
- Continued support of the centre as it moves from a pre-start-up phase to operation.
- Virtual and Face-to-Face Workshops to support the development of the centre from its research agenda to its business plan, governance, and organisational structure.
- Extensive feedback on our framework.
- Joint promotion of the collaboration.

### Engagement with Data Science & AI awards community:

- Idea originated from a meeting of the Network of CoEs.
- Partnership with BDVA to scale Irish national awards to a European level.
- Involvement of CoE members and network to promote the awards.
- Inclusion of judges from Network of Big Data CoEs, BDVA, and BDVe.
- Inclusion of Member state with Minister for Digital Single Market/Innovation.
- A venue to promote mature offerings of the project in 2020.
- A venue where data scientists and their teams from across Europe connect.

### Dissemination:

- White papers:
  - On Case Studies of 3 Successful CoEs
On interviews held with CoE Managers/Leaders
- On the BDAICoE framework
- White papers will be hosted at https://www.big-data-value.eu/skills/bigdata-centres-of-excellence/
  - Construction and constant updates on a separate Skills website https://www.big-data-value.eu/skills/bigdata-centres-of-excellence/
  - in the BDV PPP portal instead of being incorporated to the Ecosystem (with the support of WP5)
  - Collaboration with WP5 on the design of glossy templates for the white papers and feedback received on the dissemination of work

### 2.2.2 T4.2 BDV EduHub Key Impacts
The goal of T4.2 is to create the Education Hub, an online platform, and a living repository for professional education programmes on “Big Data Value”.

The Hub helps to identify the right programmes for interns and employees to improve their skills. Also, the Education Hub helps to identify, compare, and contact institutions to engage in further activities. On the one hand, the Hub enables students and professionals to find relevant and valuable educational programmes in the key areas of Big Data across Europe. On the other hand, the Education Hub promotes programmes in the area of Big Data by providing an edited list of suitable programmes for prospective students. In addition to that, the Education Hub improves the evaluation of education offers and the identification of gaps and territories for targeted support.

Therefore, the key impacts are the following:
- Information about 181 Master Programmes and 12 PhD Programmes
  - With filter search by master, PhD, location, language, duration, fees and BDVA membership
  - 20 featured programmes are from BDVA members
- Most represented in:
  - The UK with 56 courses,
  - France with 25 courses,
  - The Netherlands with 22 courses,
  - Germany with 19 courses,
  - Italy with 17 courses, and
  - Spain with 15 courses.
- Information about 62 Onsite and Online Training Programmes
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

- With filter search by online, onsite, provider, language, Development and Engineering, Data Analytics, Data Visualization, Data Management, Data Protection, and Data Processing
  - Information about over 80 more Onsite and Online Training Programmes is being collected
  - Advertised at over 11 major events (big data and domain-specific) as well as in social networks (youtube etc.).

2.2.3 T4.3 Recognition of Skills for Big Data Professionals Key Impacts

- Starting from the EDISON EDSF and taking into consideration feedback from academic and industrial stakeholders, we completed the definition of a set of six required skills for the Data Science Analytics Badge. This list of skills is publicly available on the BDVe web and will provide guidance to Data Science educational programmes when evaluating whether their training meets the needs of the industry.
- The three universities which participated in the pilot of the Academic Level of the Data Science Analytics Badge received detailed comments from reviewers of their applications. These comments will help them to improve their programmes by encouraging them to make changes to meet further the requirements of the badge (and thereby align their training with the needs of industry).
- The first open call for applicants to issue the Academic Level of the Data Science Analytics Badge will increase awareness of the Data Science Analytics Badge and its requirements.
- Publication of the paper *A Proposal for Recognizing Skills in Data Science Using Open Badges* and its presentation as a poster in the ACM sponsored Conference on Innovation and Technology in Computer Science Education (ITiCSE) in July 2019. This conference is an internationally recognised forum where computer science educators share their experiences and its inclusion in the conference, and the proceedings provides wide visibility to the work on BDV Badges.
- Online publication in KDnuggets of the article: *Addressing the Growing Need for Skills in Data Science*. KDnuggets is a blog publishing articles on AI, Analytics, Big Data, Data Mining, Data Science, and Machine Learning. KDnuggets currently has over 500K monthly visitors, and over 300K subscribers and followers. The publication reports on the discussion held during the round table organised by Task 4.3 at the BDV PPP in Riga during June 2019.

2.2.4 T4.4 Data Scientist Mobility Programme Key Impacts
The objective of T4.4 is to define and support the implementation of a mobility framework and provide an internship portal for mobility opportunities targeting data scientists, students, and professionals.

The Internship Portal will be launched beginning 2020 and has the following characteristics:

- It is specifically focused on the Big Data available internship positions, and it could be integrated with the Big Data badges
- It will be measured and compared with other generalist based on the number of the companies posting their Big Data open internship positions and the number of students applying through BDVe Internship Portal
- Will be linked to BDVe Collaborative Environment
- Should be promoted starting from a subset of the most significative universities that are part of the EduHub portal showing them the benefits of a dedicated Big Data Internship portal. When ready, the process is to engage the universities with which some collaborations already exist (e.g. UPM, POLIMI, etc.)
- It could be expected to help contribute to a “virtuous circle” between Big Data demand and skills supply

Regarding the Mobility Framework, the aim is to set up a co-location team from different BDV PPP projects, all having a common challenge to be addressed. An awareness process has started within the BDV PPP community about the need to share knowledge and skills through the mobility of resources to increase the number of resources skilled on the Big Data topics and improve within the companies the capability to address the business challenges that Big Data will offer. A survey on the willingness to participating in a pilot was done, and most of the BDV PPP projects have given their availability to identify common challenges.

When the mobility framework pilot/s through a co-located team/s, will start, it will be possible to measure how the single project has improved in the technical/problem-solving issues, thanks to the sharing of their knowledge with other, external the project, professionals/researchers and how facing problems with a different perspective and with integrated competences has facilitated the project itself.
3 General Activities

This section outlines all the events, dissemination activities, and internal meetings that facilitated and promoted the work of WP4.

3.1 WP4 Monthly Meetings

WP4 continued in M19 of the project with all tasks active from that date. We have a monthly call with the WP4 team to ensure continuity and to monitor progress. We established interworking between tasks in WP4 and the wider project. Specifically, the monthly WP4 meeting takes place on the last Thursday of the month, when possible and when deemed necessary by the work. In these meetings, updates since last meeting (including the status of action items) are discussed, as well as decisions from the consortium, contract amendments, any blockers, and next steps. A list of these meetings is available in Table 2.

<table>
<thead>
<tr>
<th>Month</th>
<th>Date</th>
<th>Location</th>
<th>Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>M19</td>
<td>26/07/2018</td>
<td>Virtual</td>
<td>INSIGHT, UDE, UPM, EIT DIGITAL, ITI</td>
</tr>
<tr>
<td>M20</td>
<td>30/08/2018</td>
<td>Virtual</td>
<td>INSIGHT, UPM, EIT DIGITAL</td>
</tr>
<tr>
<td>M21</td>
<td>27/09/2018</td>
<td>Virtual</td>
<td>INSIGHT, UPM, EIT DIGITAL, ATOS</td>
</tr>
<tr>
<td>M22</td>
<td>25/10/2018</td>
<td>Virtual</td>
<td>INSIGHT, UPM, EIT DIGITAL</td>
</tr>
<tr>
<td>M23</td>
<td>Cancelled due to overlap with BDVe monthly call</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M24</td>
<td>17/12/2018</td>
<td>Virtual</td>
<td>INSIGHT, UPM, EIT DIGITAL</td>
</tr>
<tr>
<td>M25</td>
<td>31/01/2019</td>
<td>Virtual</td>
<td>INSIGHT, UPM, EIT DIGITAL</td>
</tr>
<tr>
<td>WP4-WP5 sync call</td>
<td>07/02/2019</td>
<td>Virtual</td>
<td>INSIGHT, UDE, UPM, EIT DIGITAL, ITI, OGILVY</td>
</tr>
<tr>
<td>M26</td>
<td>28/02/2019</td>
<td>Virtual</td>
<td>INSIGHT, UPM, EIT DIGITAL</td>
</tr>
<tr>
<td>M27</td>
<td>28/03/2019</td>
<td>Virtual</td>
<td>INSIGHT, UDE, UPM, EIT DIGITAL</td>
</tr>
<tr>
<td>M28</td>
<td>25/04/2019</td>
<td>Virtual</td>
<td>INSIGHT, UDE/PALUNO, UPM</td>
</tr>
<tr>
<td>M29</td>
<td>30/05/2019</td>
<td>Virtual</td>
<td>INSIGHT, UPM, EIT DIGITAL</td>
</tr>
</tbody>
</table>
### 3.2 Consortium Meetings

WP4 is also active on BDVe consortium meetings, where the work, any issues, and future plans are presented, and feedback is received. More details on the events are given in Table 3.

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>WP4 purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>14/09/2018</td>
<td>Brussels</td>
<td>BDVe REVIEW 1, where WP4 presented work done and future work. Successful reactions and feedback were received by reviewers.</td>
</tr>
<tr>
<td>15/11/2018</td>
<td>Vienna</td>
<td>Presentation of WP4 work done, future work, and actions based on reviewers’ comments in front of advisors. Feedback was received.</td>
</tr>
<tr>
<td>09/01/2019</td>
<td>Barcelona</td>
<td>Participation of all partners of BDVe and presentation of the project status. WP4 presented on achievements, key milestones for 2019, dissemination plans, the impact for the second review, and blockers.</td>
</tr>
<tr>
<td>25/06/2019</td>
<td>Riga</td>
<td>Participation of all partners of BDVe and presentation of the project status. WP4 presented on actions based on reviewers’ comments, achievements, key milestones for 2019, dissemination plans, and blockers.</td>
</tr>
<tr>
<td>17/10/2019</td>
<td>Helsinki</td>
<td>Presentation of WP4’s purpose, highlights, and difficulties or suggestions needed in front of advisors. Feedback was received.</td>
</tr>
</tbody>
</table>
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>WP4 purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>18/11/2019</td>
<td>Madrid</td>
<td>Participation of all partners of BDVe and presentation of the project status. WP4 presented on D4.3 status, key achievements, actions based on reviewers’ comments, plan for 2020, and updates of work.</td>
</tr>
</tbody>
</table>

Table 3: BDVe Consortium Meetings Schedule

### 3.3 Events and External Dissemination

BDVe WP4 has participated in many events and external dissemination activities (including publications) during the reporting period. An outline of all the events is found in Table 4.

<table>
<thead>
<tr>
<th>Date</th>
<th>BDVe WP4 participation</th>
<th>Event</th>
<th>Partner</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>18/07/2018 - 19/07/2018</td>
<td>Presentation of the BDV badge programme and participation in the design of the EDISON EDSF</td>
<td>EDISON EDSF Release 3 Design Workshop</td>
<td>UPM</td>
<td>Amsterdam, the Netherlands</td>
</tr>
<tr>
<td>07/09/2018 - 09/09/2018</td>
<td>-</td>
<td>DatSci awards</td>
<td>INSIGHT</td>
<td>Dublin, Ireland</td>
</tr>
<tr>
<td>12/09/2018</td>
<td>Present data science skills recognition framework and conduct a working session to gather comments regarding the BDV badges programme</td>
<td>BDVA Activity Group Meeting (AG28)</td>
<td>UPM</td>
<td>Brussels, Belgium</td>
</tr>
<tr>
<td>19/09/2018 - 20/09/2018</td>
<td>Invitation from the Department of Business, Enterprise and Innovation to a roundtable discussion on draft components of an 'Industry 4.0 Strategy for Ireland'</td>
<td>-</td>
<td>INSIGHT</td>
<td>Dublin, Ireland</td>
</tr>
<tr>
<td>12/11/2018</td>
<td>Panel on “From Big Data to AI and Beyond”</td>
<td>EBDVF 2018</td>
<td>INSIGHT</td>
<td>Vienna, Austria</td>
</tr>
<tr>
<td>14/11/2018</td>
<td>Participation in workshop “Shaping the European Big Data landscape: from i-Spaces and Centers of Excellence to Big Data Digital Innovation Hubs”</td>
<td>EBDVF 2018</td>
<td>INSIGHT</td>
<td>Vienna, Austria</td>
</tr>
<tr>
<td>14/11/2018</td>
<td>Presentation of the BDV badge programme in the BDV booth and the distribution of flyers</td>
<td>EBDVF 2018</td>
<td>UPM</td>
<td>Vienna, Austria</td>
</tr>
<tr>
<td>04/12/2018</td>
<td>Towards a European certification framework of digital skills Room</td>
<td>ICT 2018: Image Digital</td>
<td>UPM</td>
<td>Vienna, Austria</td>
</tr>
<tr>
<td>Date</td>
<td>BDVe WP4 participation</td>
<td>Event</td>
<td>Partner</td>
<td>Location</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>05/12/2018</td>
<td>Participation in “Data Science Skills for Society: From Big Data to AI, Robotics and Beyond”</td>
<td>ICT 2018: Image Digital Connect, Europe</td>
<td>INSIGHT, UPM, SAP</td>
<td>Vienna, Austria</td>
</tr>
<tr>
<td>16/01/2019</td>
<td>Webinar presentation of “Big Data in the Transport Domain: Transforming Transport” (incl. EduHub advertisement)</td>
<td>BDV PPP Webinar</td>
<td>UDE</td>
<td>-</td>
</tr>
<tr>
<td>08/02/2019</td>
<td>Distribution of badge programme and EduHub flyers to all participants</td>
<td>Meeting SME Alliance</td>
<td>UPM</td>
<td>Brussels, Belgium</td>
</tr>
<tr>
<td>31/03/2019</td>
<td>Participation in a workshop related to Skills.</td>
<td>BDVA Strategy Workshops</td>
<td>INSIGHT</td>
<td>Brussels, Belgium</td>
</tr>
<tr>
<td>12/04/2019</td>
<td>Presentation of Task 4 with a focus on the badge programme and SME’s.</td>
<td>Meeting SME Alliance</td>
<td>UPM</td>
<td>Brussels, Belgium</td>
</tr>
<tr>
<td>17/04/2019</td>
<td>Presentation of “Data 4 AI: For European Economic Competitiveness and Societal Progress”</td>
<td>European Industry Partnerships Collaborative Event</td>
<td>INSIGHT</td>
<td>Amsterdam, the Netherlands</td>
</tr>
<tr>
<td>29/04/2019</td>
<td>Participation in a workshop related to Skills.</td>
<td>AI PPP Workshop</td>
<td>INSIGHT</td>
<td>Brussels, Belgium</td>
</tr>
<tr>
<td>03/06/2019</td>
<td>Keynote presentation “Data-driven AI for Self-adaptive Information Systems” at BIOC/FAISE Workshop (incl. EduHub advertisement)</td>
<td>CAiSE 2019 Conference</td>
<td>UDE</td>
<td>Rome, Italy</td>
</tr>
<tr>
<td>25/06/2019</td>
<td>Organisation of a Skills Session</td>
<td>BDV PPP Summit 2019</td>
<td>UPM</td>
<td>Riga, Latvia</td>
</tr>
<tr>
<td>26/06/2019</td>
<td>Participation in “Artificial Intelligence Session” and presentation on “Artificial Intelligence PPP Vision”</td>
<td>BDV PPP Summit 2019</td>
<td>INSIGHT</td>
<td>Riga, Latvia</td>
</tr>
<tr>
<td>Date</td>
<td>BDVe WP4 participation</td>
<td>Event</td>
<td>Partner</td>
<td>Location</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>-----------</td>
<td>------------------</td>
</tr>
<tr>
<td>26/06/2019</td>
<td>Participation in “Towards an Ecosystem for Artificial Intelligence Value-Creation” panel</td>
<td>BDV PPP Summit 2019</td>
<td>INSIGHT</td>
<td>Riga, Latvia</td>
</tr>
<tr>
<td>27/06/2019</td>
<td>Participation in “AI Strategic Agenda” workshop &amp; Inclusion of Data-skills in the development of a new SRIDA for AI PPP</td>
<td>BDV PPP Summit 2019</td>
<td>INSIGHT</td>
<td>Riga, Latvia</td>
</tr>
<tr>
<td>04/07/2019</td>
<td>Presentation of “Maschinelles Lernen für Selbst-Adaptive Software” (presentation to students)</td>
<td>paluno Tech-Talk: Wie KI die Softwaretechnik aufmischt</td>
<td>UDE</td>
<td>Essen, Germany</td>
</tr>
<tr>
<td>15/07/2019</td>
<td>Poster presentation: “A Proposal for Recognizing Skills in Data Science Using Open Badges”</td>
<td>Conference on Innovation and Technology in Computer Science Education (ITICSE), 2019</td>
<td>UPM</td>
<td>Aberdeen, Scotland</td>
</tr>
<tr>
<td>05/09/2019</td>
<td>Broadening the awards to a European level with support from BDVA</td>
<td>DatSci awards</td>
<td>INSIGHT</td>
<td>Dublin, Ireland</td>
</tr>
<tr>
<td>14/10/2019</td>
<td>Presentation of “AI Opportunities in Mobility &amp; Transport” (incl. EduHub advertisement)</td>
<td>EBDVF 2019</td>
<td>UDE</td>
<td>Helsinki, Finland</td>
</tr>
<tr>
<td>14/10/2019</td>
<td>Distribution of badge programme flyers at the BDV booth</td>
<td>EBDVF 2019</td>
<td>UPM</td>
<td>Helsinki, Finland</td>
</tr>
<tr>
<td>Date</td>
<td>BDVe WP4 participation</td>
<td>Event</td>
<td>Partner</td>
<td>Location</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------</td>
<td>---------</td>
<td>----------------------</td>
</tr>
<tr>
<td>15/10/2019</td>
<td>Participation in the Round Table: “Reveal how more and better education and training offers can be tailored”</td>
<td>European Conference SME Alliance</td>
<td>UPM</td>
<td>Brussels, Belgium</td>
</tr>
<tr>
<td>31/10/2019</td>
<td>Workshop with GATE on Governance and operation Structures</td>
<td>GATE Event</td>
<td>INSIGHT</td>
<td>Online</td>
</tr>
<tr>
<td>20/11/2019</td>
<td>Presentation of the BDV badge programme and participation in the design of the EDISON EDSF</td>
<td>EDISON EDSF Release 4 Design Workshop</td>
<td>UPM</td>
<td>Amsterdam, the Netherlands</td>
</tr>
<tr>
<td>11/12/2019 - 12/12/2019</td>
<td>Present data science skills recognition framework</td>
<td>BDVA Activity Group Meeting (AG35)</td>
<td>UPM</td>
<td>Brussels, Belgium</td>
</tr>
</tbody>
</table>

Table 4: Events and external dissemination activities of BDVe WP4

The most important events are described in more detail below, where general information about the event is given and how WP4 contributed to them.

### 3.3.1 Presentation of the BDV badge programme and participation in the EDISON EDSF Release 4 Design Workshop 2018

**Date:** 18/07/2018-19/07/2018, **Location:** Amsterdam, **Partner:** UPM

**General Description of the event:** The event was organised by members of the EDISON project to discuss revision #3 of the EDSF and to bring together users of the framework. [https://github.com/EDISONcommunity/EDSF/wiki/Agenda---EDSFr3-Design-Workshop](https://github.com/EDISONcommunity/EDSF/wiki/Agenda---EDSFr3-Design-Workshop)

**Specific input from BDVe:**

- The main objectives were to present the current state of the BDVe badge project to give some publicity to the proposal and get feedback. UPM also wanted to learn more about the EDISON Competence Framework (EDSF) and participate in the preparation of Release 3 of the EDSF.
- UPM presented the ongoing data science skills recognition work. Interesting issues about the logistics of the proposal raised. The audience seemed to
favour a more traditional ‘certification’ type approach but thought that the proposal was interesting.

- The meeting was relevant to the BDVe’s work on clarifying the terminology used, mainly between the terms ‘competence’ and ‘skill’.

3.3.2 Presentation of data science skills recognition framework at BDVA Activity Group Meeting (AG28)

Date: 12/09/2018, Location: Brussels, Partner: UPM

General Description of the event: During the BDVA Activity Group Meeting held in September of 2018, a gathering of members of the BDVA Skills Task Force was organized to discuss ongoing work by T4.3.

Specific input from BDVe:

- The UPM gave an overview of the proposed data science skills recognition framework and presented the feedback gathered from surveys regarding the initial version of the Data Science Analytics Badge
- This feedback was discussed in order to reach a consensus regarding modifications to the initial proposal.
- Other issues regarding the potential complexity of the process were also gathered.

3.3.3 Panel on “From Big Data to AI and Beyond” at EBDVF 2018

Date: 12/11/2018, Location: Vienna, Partner: INSIGHT

General Description of the event: The goal was to outline a vision for future digital skills policy in Europe to celebrate the joint achievements of the Coalition in the year 2018 and to highlight European Best Practices in the development of digital skills. The keynotes and presentations ranged from cutting-edge industrial applications of Big Data technologies, AI, innovative business cases of the data economy, inspiring future visions, and insights on EU policymaking and Research, Development and Innovation (R&D&I) funding in this area. The conference included plenary panels on Data Literacy and Data Science Skills for Society.

Specific input from BDVe: INSIGHT was a speaker and moderator to the panel "From Big Data to AI and Beyond" that focused on the unprecedented gap between the growing demand and supply of data science engineers. The importance of formal education in data science and AI was discussed, as well as how can the society prepare itself for the future, especially focusing on whether the European society is prepared for AI-driven innovation. These questions were explored by bringing together a rich set of panellists from multi-disciplines to offer a different perspective from the academic and university community, industry, and public organisations. At the end of the session, the moderator provided some concluding remarks to engage the community in the skills activity of the BDV PPP. The outcome of the panel included
information on challenges to creating an ‘AI-ready’ society and boosting future collaborations in solving these challenges.

Moderator:

- Dr Edward Curry, Research Leader at the Insight Centre for Data Analytics and at LERO

Panel:

- Dr Edward Curry, Research Leader at the Insight Centre for Data Analytics and at LERO
- Viktoria Pammer-Schindler, Assistant professor at Graz University of Technology, Research Area Manager at the Know Centre
- Gert-Breitfuss, Senior Researcher at Know Centre
- Sylvia Ilieva, Full Professor of software engineering at the Faculty of Mathematics and Informatics, Sofia University and part-time at Bulgarian Academy of Sciences
- Gerhard Schagerl, AVL List GmbH

Figure 1: Panel on “From Big Data to AI and Beyond” at EBDVF 2018

3.3.4 Workshop in “Shaping the European Big Data landscape: from i-Spaces and Centers of Excellence to Big Data Digital Innovation Hubs” at EBDVF 2018

Date: 14/11/2018, Location: Vienna, Partner: INSIGHT

General Description of the event: The general description of the event has been described above.
Specific input from BDVe: INSIGHT was a speaker at the workshop. In this workshop, various ways of collaboration were explored between BDVA Data Innovation Spaces, Big Data Centres of Excellence, and other existing initiatives, so they can offer companies a complete ecosystem to improve their business, production processes, products and services on Big Data and AI. BDVA i-Spaces are cross-organisational and cross-sectorial hubs aimed at ensuring that research on BDV technologies and novel BDV applications are quickly tested, piloted, and exploited. On the other hand, Digital Innovation Hubs are the mechanisms that the European Commission (EC) has devised to foster the digital transformation within the existing industry. The ultimate objective of the workshop was to discuss building a federation between them and the different initiatives at the European level. The outcome of the session included contributions for future calls, reinforcement of crucial pillars identified by BDVA as crucial in a post-H2020 scenario, and ideas on establishing a European network/federation of initiatives linking different capacities and technologies.

Speakers:

- Stefanie Lindstaedt (Know-Center)
- Edward Curry (Insight, BDVe project)
- Julia Palma (Eurecat, BDVA TF4-SG1)
- Daniel Alonso (ITI, BDVA TF4-SG1)

![Workshop Image](image)

Figure 2: Workshop in “Shaping the European Big Data landscape: from i-Spaces and Centers of Excellence to Big Data Digital Innovation Hubs” at EBDVF 2018

3.3.5 Participation in “Data Science Skills for Society: From Big Data to AI, Robotics and Beyond” at ICT 2018: Image Digital - Connect, Europe

Date: 05/12/2018, Location: Vienna, Partner: INSIGHT, UPM, SAP
General Description of the event: This was an open event, organised by the EC and the Austrian Presidency of the Council of the EU. 4800 visitors, including citizens, scientists, policymakers, and ICT-enthusiasts, gathered to discuss the future in a digital Europe and to focus on the EU’s priorities in the digital transformation of society and industry.

Specific input from BDVe: INSIGHT moderated a panel on “Data Science Skills for Society: From Big Data to AI, Robotics and Beyond”. This panel emphasised on the skills needed to deal with Big Data and AI technology and the challenges involved in industry and society. Part of this discussion included the presentation of the ongoing work by T4.3 to develop badges to recognise skills in data science. Future policies and recommendations were also discussed. In the panel, a diverse group of big data and AI stakeholders participated, like technology experts, academics, practitioners, and policymakers.

Moderator:
- Dr Edward Curry, Research Leader at the Insight Centre for Data Analytics and at LERO

Panel:
- Reignhard Lafrenz, Secretary-General, euRobotics
- Ernestina Menasalvas, Professor, UPM
- Jean-Christophe Pazzaglia, Digital Studio Engagement Lead, SAP

3.3.6 Presentation in the SME Alliance Workshop: Boosting Security, Data and IoT Skills – Workshop #4

Date: 12/04/2019, Location: Brussels, Partner: UPM


Specific input from BDVe:
- The UPM presented the current state of the BDVe badge program to give publicity to the project and to get some feedback.
- Interested contacts for a possible skills panel in Riga were made with Edison and Capgemini. Other contacts were made with interested participants from Fraunhofer and the Aster association (which includes the Emilia-Romagna Regional Government, Universities and National Research Centres located in the region), for future collaboration in Task 4.3.
3.3.7 Presentation “Data 4 AI: For European Economic Competitiveness and Societal Progress” at European Industry Partnerships Collaborative Event 2019

Date: 17/04/2019, Location: Amsterdam, Partner: INSIGHT

General Description of the event: The event was organised by the IoT Large-Scale Pilots Programme in collaboration with several European partnerships and DG Connect. Its aim was to present the current European research and innovation landscape related to Internet of Things/Industrial Internet of Things (IoT/IIoT), the status and outcomes of the projects related to the programme, suggest future strategic directions for EU Research and Innovation, discuss the future European partnerships under the “Horizon Europe” EU Framework Programme for Research and Innovation (2021-2027) and emphasise on European industry-driven priorities for turning technologies such as IoT/IIoT, 5G, AI, DLTs, edge computing to manufacturing, processing, construction, ICT and automotive.

Specific input from BDVe: INSIGHT presented “Data 4 AI: For European Economic Competitiveness and Societal Progress” at the event. The presentation focused on the importance of data in AI, as well as how AI has been used in BDV PPP in projects like Data Bio and Transforming Transport. It also discussed the role of data ecosystems and the opportunities they give to business, citizens, science, and government. The challenges involved and suggestions made by BDVA, as well as the need for a partnership, were also presented.

3.3.8 Participation in “Artificial Intelligence Session” and presentation on “Artificial Intelligence PPP Vision” at BDV PPP Summit 2019

Date: 26/06/2019, Location: Riga, Partner: INSIGHT

General Description of the event: The aim of the event was driving European innovation in Big Data and AI. Industry from Europe, academics, policymakers, and organisations involved in the BDV PPP exchanged ideas and fostered collaborations to shape strategies for European leadership in data-driven AI and to create a thriving European Big Data Ecosystem. That year’s BDV PPP Summit focused on the Impact empowered by Data-driven AI. The first day of the Summit included keynote speeches on data, AI, and privacy, and how these three things can be used jointly to build a future of the nation. A whole session was dedicated to data-driven AI and a deep discussion on BDVA Strategic Research and Innovation Agenda on AI. There were a few awards and social events like the best story award. In the next two days, parallel workshops were conducted by the BDV PPP community on different topics of interest. The event also had an exhibition area where several BDV PPP projects showcased their most relevant results.

Specific input from BDVe: INSIGHT moderated the “Artificial Intelligence” session and gave a presentation on “Artificial Intelligence PPP Vision”. This session included many
speakers and talks ranging from keynotes, AI application impacts in Healthcare, Robotics, and Languages, as well as a panel on “Towards an Ecosystem for Artificial Intelligence Value-Creation”.

Moderator:
- Dr Edward Curry, BDVA/Insight

Speakers:
- Ana García Robles, BDVA
- Kimmo Rossi, EC
- Dr Edward Curry, BDVA/Insight
- Bernd Liepert, euRobotics
- David Bisset, euRobotics
- Sonja Zillner, Siemens
- Milan Petkovic, Philips
- Andrejs Vasiljevs, Tilde

3.3.9 Panel on “Towards an Ecosystem for Artificial Intelligence Value-Creation” at BDV PPP Summit 2019

Date: 26/06/2019, Location: Riga, Partner: INSIGHT

General Description of the event: The general description of the event has been described above.

Specific input from BDVe: INSIGHT moderated the “Towards an Ecosystem for Artificial Intelligence Value-Creation” panel, where panellists discussed the future of AI, the importance of fostering collaborations across European industrial, academic and policy-making partners to building an AI Ecosystem to address future needs and challenges.

Moderator:
- Dr Edward Curry, BDVA/Insight

Speakers:
- Kimmo Rossi, EC
- Andrejs Vasiljevs, Tilde/META-NET
- Laure Le Bars, SAP
- Milan Petkovic, Philips
- Michela Milano, UniBo/EuroAI/AI4EU
- Alun Foster, ECSEL
- Nizar Touleimat, CEA/ECSO
3.3.10 Panel on “Addressing the Growing Need for Skills in Data Science” at the BDV PPP Summit 2019

**Date:** 27/06/2019, **Location:** Riga, **Partner:** UPM

**General Description of the event:** The general description of the event has been given above.

**Specific input from BDVe:** The UPM organised and moderated the “Addressing the growing need for skills in data science” panel, where experts discussed the data science skills gap. A detailed report of the discussion can be found in the KD Nuggets blog: [https://www.kdnuggets.com/2019/10/growing-need-skills-data-science.html](https://www.kdnuggets.com/2019/10/growing-need-skills-data-science.html)

**Moderator:**
- Ernestina Menasalvas, BDVA/UPM

**Speakers:**
- Rocco Defina, Oxys Consulting
- Yuri Demchenko, University of Amsterdam
- Thomas Hauser, University of Colorado at Boulder
- Liesbeth Ruoff-van Welzen, KNVI and IP3

![Figure 3: Skills Session in the Riga BDV PPP Summit](image)

3.3.11 Poster presentation of “A Proposal for Recognizing Skills in Data Science Using Open Badges” at the Conference on Innovation and Technology in Computer Science Education (ITiCSE) 2019
Date: 15/07/2019, Location: Aberdeen, Partner: UPM

General Description of the event: ITiCSE is an annual European conference on Computer Science Education organised with the support of the ACM Special Interest group on Computer Science Education. The main purpose of the conference is to gather experts interested in computer science education and to encourage the exchange of educational experiences. [https://iticse.acm.org/ITiCSE2019/](https://iticse.acm.org/ITiCSE2019/)

Specific input from BDVe:  

- Presentation of a poster “A Proposal for Recognizing Skills in Data Science Using Open Badges”, to disseminate the work done on the BDVe skills recognition program and to invite the attendees of the conference to participate either in the pilot or the official call to issue badges. The poster is shown below:

![Poster of “A Proposal for Recognizing Skills in Data Science Using Open Badges”](image_url)
3.3.12 DataSci Awards

**Date:** 05/09/2019, **Location:** Dublin, **Partner:** INSIGHT

**General Description of the event:** The DatSci awards has started four years ago to celebrate and connect the most talented data scientists and their teams across Europe. It has emerged as Europe’s leading celebration of talent in the Data Science & AI community, with more than 450 attendees and applications being accepted from over 48 countries. The award has 10 categories and 35 judges. This year, ten awards were presented. These awards were a mean to appreciate leading talents in Data Science and gave the opportunity to discuss emerging technologies, network with C-Level decision-makers, and build connections with like-minded people.

The winners of 2019 were:

- Data Scientist of the year: Nuria Oliver, Vodafone & Data-Pop Alliance
- Award for Best application of AI for the year: Jonhson Controls
- Award for best data use of data science in an SME/Startup: Geowox
- Award for best data use of data science/AI for customer experience: Beauty Machine Engine
- Best use of data science for Industry 4.0: Tecnalia & Gestamp
- Best use of data science to achieve social impact: IBM Ireland Lab
- Award for data science technology innovation of the year: Logical Clocks AB
- Best technical advance in the field of data science/AI from a research organisation: Accenture Labs
- Award for best data use of data science/AI for health and wellbeing: AXIAL 3D
- Data Science Student of the year: Rory Boyle, The When Lab-Trinity College Dublin

The Advisory Committee was:

- Sheamus Causer, Chief Technology and Information Officer, Ulster Bank
- Freda Cunningham, Director, Deloitte Ireland
- Dr Edward McDonald, Director, CeADAR Centre for Big Data Analytics
- Dr Mick Kerrigan, Chief Science Officer, RecommenderX DAC
- Stefan Decker, Professor for Databases and Information Systems, RWTH Aachen University, Director, Fraunhofer Institute for Applied Information Technology (FIT)

**Specific input from BDVe:** One of the main challenges is how to create a community or a venue where data-driven insights from different sectors (academia, industry, government, etc.) could be shared with other sectors and wider audiences. At the same time, certain regions of Europe might benefit more from frequent face-to-face collaboration or sharing of ideas, whereas some other regions may face some constraints.
INSIGHT facilitated the partnership of DatSci Irish national awards with BDVA and succeeded in broadening the scope of the awards to a European level. The event was supported by the Network of CoEs and its members, giving the opportunity for further exchange of academics across the network, successful contribution to European projects with industry use-cases, implementation and testing of innovative and data-driven solutions, identification of synergies and sharing of best practices and visibility internationally to define the data-driven future of Europe. Judges included members of BDVA and BDVe and support was received from the Irish government as the opening speech was given by the Irish Minister for the Digital Single Market (DSM). The aim of this partnership was to show how the BDVe with the support from BDVA can bring the community together and to provide a venue to promote mature offerings of projects, including BDVe in 2020 and a place that data scientists across Europe could connect and network.

![Figure 5: DatSci awards 2019](image)

3.3.13 Participation in the Round Table: “Reveal how more and better education and training offers can be tailored” at European Conference SME Alliance

**Date:** 15/10/2019, **Location:** Brussels, **Partner:** UPM

**General Description of the event:** In this event of around 150 people, there were 4 round tables focused on major themes addressing the importance of skills development and upskilling of SMEs to support their digital transformation.

**Specific input from BDVe:** The UPM participated in the Round Table “Reveal how more and better education and training offers can be tailored”, with the aim of presenting an overview of the skills recognition program developed in the BDVe project and to disseminate this work among institutions which might hire future professionals who have received BDV badges.
3.3.14 Publication of “Addressing the Growing Need for Skills in Data Science” in the KDnuggets blogpost

**Date:** 10/2019, **Location:** -, **Partner:** UPM

**General Description of the event:** "KDnuggets™ is a leading site on AI, Analytics, Big Data, Data Mining, Data Science, and Machine Learning. KDnuggets is currently reaching over 500,000 unique monthly visitors, and over 260,000 subscribers via email, Twitter, LinkedIn, Facebook, feedly/RSS, and Google+. KDnuggets is the most popular blog about data management.

**Specific input from BDVe:**
- The publication reports on the discussion held during the round table organised by the UPM at the BDV PPP in Riga during June 2019.
- The publication is available at https://www.kdnuggets.com/2019/10/growing-need-skills-data-science.html#.Xa84b0rPYbU

3.3.15 Participation in the Steering Committee of the SME Alliance

**Date:** 18/11/2019, **Location:** Brussels, **Partner:** UPM

**General Description of the event:** The SME Alliance is managed by a consortium of Capgemini Invent, Technopolis Group, and the European DIGITAL SME Alliance. Together with SMEs and other stakeholders, it developed a vision, a roadmap consisting of support measures, and a monitoring mechanism. The aim is to contribute to strengthening SMEs' workforce adaptability and capacity for the short to medium term.

**Specific input from BDVe:** The steering committee consists of five independent experts renowned in the field, including Ernestina Menasalvas. The objective of the meeting was to discuss the comments of the SC members on the final report, in order to validate it. As a result of the meeting in which the report was analysed, some
comments were giving, and the importance of the up-skilling and re-skilling for employees in SMEs was highlighted.

3.3.16 Presentation of the data science recognition programs from T4.3 and participation in the EDISON EDSF Release 4 Design Workshop

**Date:** 20/11/2019, **Location:** Amsterdam, **Partner:** UPM

**General Description of the event:** The event was organised by members of the EDISON project to discuss revision #4 of the EDSF and to bring together users of the framework. [https://github.com/EDISONcommunity/EDSF/wiki/(2)-EDSF-Release-4-Design-Workshop-20-November-2019](https://github.com/EDISONcommunity/EDSF/wiki/(2)-EDSF-Release-4-Design-Workshop-20-November-2019)

**Specific input from BDVe:**

- The current state of the formal and non-formal recognition frameworks being developed by T4.3 was presented in the workshop, and comments were received from participants.
- The next version of the EDISON EDSF was discussed, and feedback was given regarding the proposed changes.

3.3.17 Presentation and discussion of work to recognise data science skills at a BDVA Activity Group Meeting (AG35)

**Date:** 11/12/2019-12/12/2019, **Location:** Brussels, **Partner:** UPM

**General Description of the event:** During the BDVA Activity Group Meeting held in December of 2019, a workshop for members of the BDVA Skills Task Force was organised to present and discuss ongoing work by T4.3.

**Specific input from BDVe:**

- On the 11\textsuperscript{th} of December the UPM organised a three-hour skills workshop with the following agenda:
  - Give a review of previous work to establish the BDV Data Science Badge Programme for formal training
  - Gather feedback on the new proposal for recognising skills in non-formal training
  - Discuss/plan work for 2020 and beyond
- On the 12\textsuperscript{th} of December, a summary of the main points discussed during the previous day’s workshop was presented in a plenary session to all participants in the activity group meeting.
Figure 7: BDVA Activity Group Meeting (AG35)
4 Network of National BDV Centres of Excellence (T4.1-INSIGHT)

In this section, the main achievements of T4.1 “Network of national BDV centres of excellence” are outlined.

4.1 White papers

WP4, with the collaboration of WP5, has created a dissemination plan to share the outcomes of D4.1 [10]. The plan involved disseminating the work by using media like blogposts, tweets, white papers, and a dedicated website. Also, the work should continue to be promoted through national events, the BDVe community, and workshops.

Glossy templates have been designed by Ogilvy through our input and feedback. These templates have been designed for disseminating the work of our 3 case studies (Insight, CeADAR, SIRIUS), the BDAICoE framework, and the interviews that we have taken from the top leadership of the centres. This is a process that demands several reviews and approvals from the interviewees. The glossy templates can be found in Appendix B.

4.2 Sharing Best Practices

The current data explosion combined with recent advances in computing power and connectivity allows for an increasing amount of Big Data to be analysed anytime, anywhere. These technical advances enable addressing industrial relevant challenges and foster intelligent industrial application in a shorter time and with higher performance. We are now seeing how AI can increase value creation from Big Data and its use to rapidly emerging business-to-business, business-to-government, business-to-consumer, government-to-business and government-to-consumer scenarios in many AI application domains.

Since the creation of the first version of the framework, the relevance of the European ecosystem to AI has increased. In this time, many Big Data CoEs have extended their remit to also include activities in AI. At the same time, the EC has stimulated the creation of Networks of Excellence in AI via the ICT-48 call, of which many Big Data CoEs are actively participating. As a reaction to this change in the landscape, we have started to consider AI-specific topics within the development of our model. Based on our initial work in these areas, the majority of the work on Big Data CoE is applicable
to AI CoEs, especially those with a data-driven AI focus. As such, we have decided to rename the model to BDAICoE.

4.2.1 Best Practice Framework Overview

The task of the deliverable D4.1 [10] is to discover existing best practices in the domain of Big Data research. Additionally, the relevant information gathered from various sources about the practices has been grouped and designed into a framework under various themes. These themes, in the context of this deliverable, are named as elements of the BDAICoE best practice framework (Figure 8). The framework is also hosted at the BDV Skills website on https://www.big-data-value.eu/skills/bigdata-centres-of-excellence/. For the purpose of quick reference, a brief explanation is provided for each element of the BDAICoE framework.

**Environment**

“Environment means forces [that are] difficult to control from inside that demand a response” [14]. Another view of the external environment of an organisation is that it comprises of forces that initiate organisational change [2]. The environment is subdivided into three areas:

- **Industry:** Industry is defined as the ecosystem of companies surrounding a BDV CoE, that is associated with the creation of economic value at both national and European levels.
▪ **Policy:** Policy is defined as the set of public laws, regulations, and policies that govern research and innovation activities at national and European level, as well as dictate the access, manipulation, and distribution of data.

▪ **Societal:** The societal environment of a BDAICoE comprises of state of human development as measured by composite statistics and indexes, and the national priorities for human development in terms of the United Nations Sustainable Development Goals and H2020 Societal Challenges.

**Core Organisational Model**
This subsection provides detailed information on the main element or concepts within the BDAICoE core model.

▪ **Strategy:** Strategy represents the means by which a CoE intends to achieve its overall mission and goals.

▪ **Governance:** Governance in centres of excellence refers to the level of decision-making about strategy and operations.

▪ **Structure:** The structure is how a CoE is designed (i.e., levels, roles, units, decisions, rights, and accountability).

▪ **Funding:** Funding refers to the availability, diversity, and sustainability of the monetary support for carrying out research and educational activities in a CoE.

▪ **People:** People are the human capital required to carry out specific tasks towards the goals of the organisation.

▪ **Culture:** Culture represents the underlying values, beliefs, and norms that drive the teams and the CoE as a whole.

**Capabilities**

▪ **People** – People are the human capital required to carry out specific tasks towards the goals of the organisation.

▪ **Process** - Process is the knowledge of procedures and tasks for the achievement of the goals of the CoE.

▪ **Infrastructure** - Infrastructure is the systems, practices, and tools that facilitate and reinforce the work within the organisation.

▪ **Outreach:** Outreach is the collection of information dissemination activities with which a research centre informs the public about the science and technology developments in the centre. The aim is to enable the public to appreciate science and technology.

▪ **Collaboration:** Universities-industry collaboration (UIC) refers to the formal and informal engagement and interaction between a higher educational institution and an industry partner with the aim to facilitate knowledge and technology exchange as well as to provide an ad hoc advice and networking opportunity for the professionals. This can be done through the establishment
of activities such as collaborative and contract research and the provision of consulting services.

Capabilities are analysed more in Table 5.

<table>
<thead>
<tr>
<th>Operational Capability</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Development</td>
<td>How the centre develops new business opportunities and manages its partnerships</td>
</tr>
<tr>
<td>Collaboration</td>
<td>How the centre enhances Academic to Academic and Academic to Industrial Interactions</td>
</tr>
<tr>
<td>Research Support Services</td>
<td>The local research support services implemented by the centre</td>
</tr>
<tr>
<td>Technical Infrastructure</td>
<td>Computing resources used to support the research and innovation activities of the centre</td>
</tr>
<tr>
<td>Experimentation/Demonstration Platforms</td>
<td>The platforms that support the scientific and innovation activities of the centre</td>
</tr>
<tr>
<td>Intellectual Property (IP) and Data Protection (DP)</td>
<td>How the centre approaches IP management and DP</td>
</tr>
<tr>
<td>Education and Public Engagement (EPE)</td>
<td>How the centre’s dissemination activities inform the public of the science and technology developments</td>
</tr>
<tr>
<td>Policy Outreach</td>
<td>How the centre tried to Influence future policy</td>
</tr>
<tr>
<td>Technology and Knowledge Transfer</td>
<td>How the centre drives the transfer of know-how and adoption of its technology</td>
</tr>
<tr>
<td>Performance and Impact Assessment</td>
<td>How the centre identifies and tracks its performance and impact</td>
</tr>
</tbody>
</table>

**Table 5: Core operational capabilities of the BDAICoE framework**

**Impact**

The direct and indirect ‘influence’ of research or its ‘effect on’ an individual, a community, or society as a whole, including benefits to the economic, social, human, and natural capital.

- **Economic**: The economic impact is the effect on commerce, employment, or incomes generated from big data research in general and by the CoE in particular.

- **Scientific**: This relates to the influence a research centre has on the entire science and technology communities around the world. It includes the contributions it makes to the invention of novel ideas or concepts and the development of general science and technology principles.

- **Societal**: This relates to the useful impact of the result of a research centre on the entire human society, including the impact on awareness about science
4.2.2 Deep Dive on Capabilities

In period 2, we focused on the more detailed elaboration of the best practices within the model. This involved identifying specific practices within each part of our framework. This was achieved by performing a detailed analysis of the case studies and interviews with subject experts. The model has been extended with key practices for the core model capabilities (Strategy, Governance, Structure, Funding, People and Culture)

Key practices were also identified for each of the capabilities for the centres.

- Business Development
- Collaboration
- Research Support Services
- Technical Infrastructure
- Intellectual Property (IP) and Data Protection (DP)
- Education and Public Engagement (EPE)
- Policy Outreach
- Technology and Knowledge Transfer
- Performance and Impact Assessment

Full details on these practices are found in Appendix A.

4.2.3 Critical Success Factors

Critical Success factors are a range of key enablers that research centres, like corporate bodies, employ to achieve success in their operations. While some are very easily identifiable, e.g. funding availability and a mix of employees’ capabilities and cooperation, other success factors are not quite salient, e.g. the role of culture in the success of a research centre. Similarly, some success factors are common to a majority of research centres, e.g., the importance of enough funding to success, possession of world-class researchers, collaboration with important partners, output publicity, and so on. Other factors are very peculiar to individual research centres because certain factors apply to the research focus of a research centre. However, whatever the key success factor is, it is the responsibility of the management team to identify it early enough and to harness it to drive success in the required direction.

This section reports the findings of the BDAI CoE case study centres as success factor recommendations for existing Bid Data Value Centres and potential ones for their research operations. These factors are gathered from interviews with the centres’ senior management using a series of open-ended questions:

1. What are the common difficulties faced by the centre in achieving its objectives?
2. What factors contribute to/enable the success of the centre?
3. What are the typical mechanisms deployed to address success factors and challenges in the centre?
4. What would you need to do to be more successful?

Challenges are the drawbacks on the progress of any organisation, while the success factors facilitate progress. Therefore, the management team of an organisation, according to its mandate, has to device strategies and practices to eliminate or at least mitigate challenges and other risks to success. On the contrary, success factors are good for progress because they facilitate operations, so they have to be leveraged to drive success in the required direction of the organisation’s goal.

4.2.3.1 Challenges
The key challenges identified are detailed in Table 6.

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>To stay as a going concern – sustainability in the research industry.</td>
<td>Sustainability</td>
</tr>
<tr>
<td>Ensure essential funding to pursue basic research.</td>
<td>Research Funding</td>
</tr>
<tr>
<td>Meeting the high-performance target set for the centre.</td>
<td>Performance</td>
</tr>
<tr>
<td>Encourage more collaboration and partnership arrangements achieved.</td>
<td>Management</td>
</tr>
<tr>
<td>Lack of autonomy. The lack of separate legal entity status.</td>
<td>Autonomy</td>
</tr>
<tr>
<td>The need to ensure that governance adds value to the centre’s operations creates some concerns.</td>
<td>Governance</td>
</tr>
<tr>
<td>Competing Interest – Funders’ objectives versus Researchers’ objectives.</td>
<td>Balancing interests</td>
</tr>
<tr>
<td>Human resource availability and retention, e.g. recruitment of PhD level graduates with significant industry experience, can be a challenge.</td>
<td>Human resource</td>
</tr>
<tr>
<td>Working with SMEs is challenging due to their resource availability problems, lack of clearly defined objectives, and often have short-term plans.</td>
<td>Collaboration with SMEs</td>
</tr>
<tr>
<td>Physical separation from important partners limits interaction and knowledge of themselves.</td>
<td>Interaction with partners</td>
</tr>
<tr>
<td>Facilitation of a flowing, open discussion of technology and solutions between the centre and industrial partners.</td>
<td>Interaction with partners</td>
</tr>
<tr>
<td>Capability and capacity to assure partners that the centre will help them to solve their challenges.</td>
<td>Assurance for partners</td>
</tr>
<tr>
<td>Bridging the knowledge gaps between academic IT, commercial IT with the associated research and business problems.</td>
<td>Academic IT/business Gaps</td>
</tr>
</tbody>
</table>
### Table 6: Summary of Challenges

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>The need to bridge the gap between people with knowledge of the business problem and those with knowledge of theory.</td>
<td>Knowledge Transfer</td>
</tr>
<tr>
<td>Maintaining a flow of new project contracts and adequacy in project management expertise.</td>
<td>Gaining contracts</td>
</tr>
<tr>
<td>Industry Funding policy demands up to 25%-50% of its funding needs from industry. It creates a challenge of how to balance the interests of researchers with partners.</td>
<td>Funding policy</td>
</tr>
<tr>
<td>Aligning portfolio with the strategy to meet partners’ demands. This also creates project selection and investment challenges, which often lead to frustrating researchers and industry partners.</td>
<td>Portfolio and strategy alignment</td>
</tr>
<tr>
<td>Work overload arises from too many activities at the centre which is, perhaps, contributed by the funding policy.</td>
<td>Performance</td>
</tr>
<tr>
<td>The trade-off between expediency and consensus in making decisions and at the same time, gaining staff commitment to achieve the centre’s goals.</td>
<td>Expediency vs consensus in decision-making</td>
</tr>
<tr>
<td>Leading knowledge workers who are not driven by ordinary incentives like salaries because they have their own career agendas.</td>
<td>Stakeholder management</td>
</tr>
<tr>
<td>There is a need for the ‘cross-pollination’ of cultures between research and industry environments.</td>
<td>Knowledge Transfer</td>
</tr>
</tbody>
</table>

#### 4.2.3.2 Success Factors

The factors with which the centres’ leadership contribute to their success are detailed in Table 7.

<table>
<thead>
<tr>
<th>Success Factors</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to attract grant funding is based on reputation (both the centre and individuals) for excellent research outputs.</td>
<td>Funding</td>
</tr>
<tr>
<td>Local presence of big industry players in tech, medical, pharma, etc. offers opportunities for collaboration and industry funding.</td>
<td>Economy influence</td>
</tr>
<tr>
<td>The stock of a talented team of people: The capability to assemble world-class academic talents attracts and satisfies stakeholders.</td>
<td>Talent availability</td>
</tr>
<tr>
<td>Ensuring that the people of the centre can develop themselves and their careers.</td>
<td>People support</td>
</tr>
<tr>
<td>Effective public outreach that translates science into something easy to understand for non-scientists.</td>
<td>Public Outreach</td>
</tr>
</tbody>
</table>
Success Factors | Keywords
--- | ---
Organisational structure: Maximises outputs by providing (i) *space* (infrastructure and labs) that attract academics, (ii) *money*, and (iii) *reputation* of the individual members and the team. | Organisational structure

The research turnover: Turnover ensures that the centre is fresh and relevant to the industry. | Research turnover

Teamwork and collaboration: The more the centre achieves collaboration, the more successful it will be. Thus, there is a direct relationship. | Teamwork and collaboration

Focus on projects that are proposed by industry members. This ensures that what is produced will have an immediate and beneficial impact. | Immediate impact output

The produce-for-immediate-impact dynamic is highly motivating for the centre to get to work on a huge variety of projects across many industries every six months. | Motivation from impact

Deep collaboration with industry partners provides the centre with a huge opportunity for success, as it is involved in industry-focused research. | Collaboration

The support of the funding agencies is received in two ways — in the form of funding supply and help in the prioritisation of the research agenda. | Funding

The centre is structured to support balancing scientific excellence and supporting business partners. | Structure

The centre supports academic researchers in their career development and the goal of the centre through operationalisation of both agendas in daily activities. This decision enables a robust structure that allows people to be focused both on their personal needs and the needs of the centre. | Talent development

The committed and hardworking young scientists of international combinations make significant contributions. | Partners’ contribution

The industrial experience of the management team, which possesses a unique skill set in communication and industry-research collaboration and capability to speak/understand the languages of both the academics and industry. | Industrial experience

Table 7: Summary of Success Factors

### 4.2.3.3 Mechanisms to Address Challenges

The mechanisms deployed by the centre’s leadership to address their challenges are detailed in Table 8.

Practice | Keywords
--- | ---
Planning and measuring process: | Planning
Measuring
<table>
<thead>
<tr>
<th>Practice</th>
<th>Keywords</th>
</tr>
</thead>
</table>
| • Development of a strategic plan, an annual appraisal plan, and key performance indicators (KPIs) plan to align with the centre’s goals.  
• Measured and reviewed on a monthly basis.  
• Iterative planning process: over time, a plan may need to be reviewed and adjusted because initial factors affecting the plan have changed. |  |
| Communicate the progress of the centre on a regular basis to all members of the centre to promote unity and focus on the common goal. | Effective communication |
| Publish a strategic plan and allow people at all levels to engage. This allows people to engage with the vision. | Engagement |
| To help attain very high targets: Break the KPIs down into manageable pieces that people can handle. | KPI management |
| Aligns research agenda with the National Government’s science and technology agenda and the goals of industry partners and domain trends. | Alignment of goals |
| Using media publicity on current trends and using the media to create awareness about its research output. | Media publicity |
| The centre maintains a market-focused approach by engaging with industry and other research centre representatives at different events. | Events |
| Enables funding agencies to help to prioritise its research agenda. | Funding |
| Meet with industrial stakeholders twice yearly to deliberate and to set research agenda as well as help in decision-making processes. | Collaboration |
| Arrangement for obtaining IP is simple and quick. This attracts industry partners to sign up a contract for a collaborative project. | IP |
| A one-on-one mentorship programme with industry to enrich the centre’s experience in the development of researcher talent. | Mentorship |
| A monthly meeting with industry partners’ representatives to monitor and discuss the progress of the centre. Meetings ensure regular engagement of industry partners and increase awareness of industry role in making the centre a success. | Management meeting |
| Internal meetings (weekly and monthly) enable the management team to get constant visibility of the centre’s internal operations. | Internal meeting |

**Table 8: Summary of Mechanisms to Address Challenges**
4.2.3.4 Ideal Situation

According to the centres’ leadership, the ideal conditions for the operation of their centre are detailed in Table 9.

<table>
<thead>
<tr>
<th>Ideal Situation</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separate legal entity status may allow the centre to evolve into a larger entity to deal with SMEs, become self-sustaining, and to be able to deliver all its mandates.</td>
<td>Autonomy</td>
</tr>
<tr>
<td>Right balance of resources to deal with all challenges, meet increasing knowledge and demands for Data Analytics outputs.</td>
<td>Resource mix</td>
</tr>
<tr>
<td>Academic Service Level Agreement between the research centre and academics who are working for the research centre.</td>
<td>Service Level Agreement (SLA)</td>
</tr>
<tr>
<td>Meritocracy – a basis for decision-making on funding, performance, and rewards.</td>
<td>Meritocracy</td>
</tr>
<tr>
<td>A Strategic Investment Fund can provide flexibility, particularly in a situation where a merit-based funding policy is lacking.</td>
<td>Funding</td>
</tr>
<tr>
<td>Collaboration-seeking techniques to attract people to collaborate across non-traditional boundaries, both internally and externally.</td>
<td>Collaboration</td>
</tr>
<tr>
<td>Having a less divided funding framework, an increase funding level and an aligning funding interest with stakeholders’ interest and centre’s ambition.</td>
<td>Framework</td>
</tr>
<tr>
<td>Increase cash contribution from industry partners.</td>
<td>Funding</td>
</tr>
<tr>
<td>Division of labour in a more balanced way among the centre’s people.</td>
<td>Workload</td>
</tr>
<tr>
<td>Need to develop international networks and collaborations.</td>
<td>Capacity</td>
</tr>
</tbody>
</table>

**Table 9: Summary of Ideal Situations**

4.3 Engagement Lifecycle

![Figure 9: Engagement Lifecycle for BDAICoEs](image-url)
In order to support centres in a systematic manner, we have identified a lifecycle based on a three-phased approach that has been developed to provide a structured and deliberate approach to the support of BDAICoE within T4.1. The three-phase lifecycle comprises of a series of strategies and activities which are led stimulated by BDVe to maximise the impact of our resources and ensure we provide the maximise support possible to CoEs.

We believe that by implementing the lifecycle phases, it will maximise our ability to attract, provide advice, and facilitate opportunities for CoEs. The lifecycle supports us in targeting our messages and providing segmentation of CoE and target groups for communication strategies. The lifecycle includes three phases: Attract, Advice, and Opportunities.

**Phase 1 – Attract**

The objective of the first phase of the lifecycle is to *Attract* the attention of the target CoE and stakeholders, which we would like to work with. The *Attract* Phase has three activities – Target, Raise Awareness, and Activate Ambition.

- **Target**: We have defined our ideal targets by researching their existing BDAICoE landscape in Europe. We have then identified that we want to target regions without a CoE, regional academic clusters, regional development agencies, and early stage CoEs who may be interested in the advice on establishing and running a CoE.

- **Raise Awareness**: In order to connection with our target groups, we look to raise awareness of the advice and supports available from BDVe. We do this by contacting National Contact Points as the first point of call for each member state. We seek to make them aware of our activity and to ask for support in connecting with CoEs, groups that may have an interest in creating a CoE, and local funding and development agencies. We also raise awareness of our activities at events such as BDVA Activity Groups, EBDVF, and BDV Summit.

- **“Activate Ambition”**: The critical ingredient needed to establish a new CoE is the ambition of a university or an academic (or a group of universities or academics) to want to create a CoE. To make the CoE a reality, it must also be supported by the wider ecosystem of stakeholders from funding agencies and local/regional development agencies to industrial clusters and government. The objective of this activity is to “activate” the ambition of these stakeholders to the possibility of creating a CoE. The key mechanism we deploy is the use of success stories of existing CoEs and the journey undertaken about new CoEs. It is our hope that these stories can spark the desire and belief of our target stakeholders and encourage them to establish a CoE.

**Phase 2 – Advice**
Advice is the second phase in the Lifecycle, and it too has three activities that target centres at different stages of development. The advice we can offer a CoE is to focus on the issues which are most relevant for a centre at their stage of development. The activities target centres at Pre-Startup, Early-Stage and Established.

- **Pre-Startup:** In this activity, we focus on providing advice and support to groups that have an aspiration to establish a new BDAICoE. Typically, these groups are interested in developing a vision and strategy for the CoE and securing initial funding to establish the centre. This group is also interested in understanding an overview of the operations of a CoE to understand what is needed to successfully run a CoE.

- **Early-Stage:** For CoEs which have been recently established, we are able to provide advice on key factors that are important to running a CoE. These include the creation of an appropriate Strategic Plan and associated Key Performance Indicators (KPIs), Organisational Structure and Governance Model, and developing an appropriate cultural and staff support of the CoE. CoEs at this stage may also be interested in creating or improving a range of capabilities from Technology Transfer to Education and Public Engagement.

- **Established:** Mature CoEs may have an interest in understanding how well their current practices are in terms of performance. The intent for 2020 is to define support to make it easier for CoEs to benchmark their practices against other similar CoEs. This can support them in improving their operations.

**Phase 3 – Opportunities:** The third phase in the Lifecycle is *Opportunities*, which has three stages: Network, New Practices, and Collaboration.

- **Network:** CoEs are encouraged to join in the activities of the Network of CoEs to participate in the associated networking activities and to make connections to other CoEs in Europe. This can support the CoE in participating in proposals for national and European projects.

- **New Practices:** Subject Matter Experts are engaged to identify new and update existing practices that have been captured within the BDAICoE guide. This ensures the guide remains relevant to the CoE community and evolves to meet the emerging challenges for CoEs.

- **Collaboration:** Within the Network for CoEs, participations are able to establish collaborative links with other CoEs. This can lead to formal and informal collaborations at the national level (i.e. CeADAR and Insight in Ireland) and European levels. These collaborations can take the form of a proposal activity that can lead to joint projects. A number of CoEs have participated in site visits to share learning and best practices.
4.4 Supporting the Establishment of new Centres of Excellence

One of the key tasks within T4.1 is to support the establishment of new CoEs within Europe. Within Period 1 of the project, a survey of CoEs was performed to identify Member States and regions where there were no CoEs. This survey identified 32 CoEs within existing Member states and regions, illustrated in Figure 10. Among them, 21 belong to Western Europe, 1 to Eastern Europe, 5 to Southern Europe, and 5 to Northern Europe. This led us to the conclusion that there is an abundance of centres of excellence in the West part of Europe, some in the North and South of Europe, and almost none in the East of Europe. This highlighted the need to support Eastern Europe in member states and regions in creating new CoEs.

![Figure 10: Geographic heat map with all European Big Data Centres of Excellence](image)

To this end, we have been in contact with a number of Universities and region representatives to highlight the support available in creating a new CoE. Working with WP5 on how best to contact potential new CoEs, it became clear that the target audience for these supports is a small group of stakeholders (often only 1-2 per member state), who has an interest in establishing a CoE. To this end, we have reached
out using the BDVA as a vehicle to make the community aware of these supports. We have also enlisted the support of the Ideal-ist Network of National Contact Points for ICT which will make us aware of any new CoE that is being considered.

Current engagement within this area ranges from informal conversations and online meetings to full-day workshops with relevant stakeholders. Our activities have focused on the following member states:

- **Latvia**: Meeting with Ralfs Nemiro, Ministry of Economics of the Republic of Latvia on the opportunities to develop the Latvian Data Economy. The establishment of a Big Data CoE was part of the discussion on actions that support the local data ecosystem.
- **Estonia**: Meeting with Dr Kristjan Vassil, Vice-Rector for Research, University of Tartu. The University is now a member of BDVA and actively engaging in BDV PPP Proposals. A CoE is under consideration.
- **Bulgaria**: Working with the University of Sofia on establishing the GATE Big Data CoE.

### 4.4.1 GATE – Sofia Bulgaria - Success Story

GATE was a Horizon 2020 WIDESPREAD-2016-2017 TEAMING Phase 1 programme that aspired to create a sustainable business plan for the creation of the first CoE in Big Data in Bulgaria. The purpose of this Big Data Centre is to produce excellent science by seamlessly integrating connected fields and associating complementary skills. GATE aspired to add value to knowledge, to strengthen the capacity of researchers, to educate and train early-stage researchers, to disseminate and promote projects, and to achieve international visibility and scientific, as well as industrial connectivity. With innovation pillars, like Data-Driven Government (Public Services based on Open Data), Data-Driven Industry (Manufacturing and Production), Data-Driven Society (Smart and Sustainable Cities) and Data-Driven Science (Big Data technology stack in the scientific community), GATE had set its aim high to fulfil its goal.

Our advisory involvement in the GATE project by sharing our best practice guide proved very successful in this aspect. A number of meetings and workshops took part in the GATE project in the first period, as well as joint dissemination activities in BDV Meet up at Sofia in Bulgaria. By the end of the project, our involvement was reviewing the research strategy and the business plan. In this period, we received very positive feedback since GATE received the green light from the EC to build the first Big Data CoE in Bulgaria, paving the way for more CoEs to start spreading in Eastern Europe in the future.
4.4.2 Reaching out Through IDEAL-IST

We are now in active discussion with Ideal-ist to reach stakeholders in the following member states without a CoE:

- **Southern Europe Region**: Cyprus, Malta, and Portugal.
- **Northern Europe Region**: Latvia and Lithuania.
- **Eastern Europe Region**: Croatia, Czech Republic, Poland, Romania, Slovakia, and Slovenia.

Within the last year of the project, we will focus our activities on supporting any new centres which emerge from these member states. We will work with Ideal-ist to identify these new centres as early as possible in the formation process to ensure we can maximise our support.

4.5 Network of Excellence Collaborations

The Network of Centres of Excellence in Europe aims to boost collaboration and define the data-driven future of Europe. 55 centres of excellence, all different in terms of scientific background and focus. The network’s main focus is to support big data research and innovation within Europe by exchange of academics across the network, the successful contribution to European projects with industry use-cases, the implementation and testing of innovative and data-driven solutions, the identification of synergies, the sharing of best practices and the visibility internationally, can all strengthen the European economy.

Network Organisers
The Network of Big Data Centres of Excellence in Europe is run by Know-Centre and the Berlin Big Data Centre (BBDC). The extended core group is constituted by INSIGHT, IMEC, RISE, and Big Data CoE Barcelona. The network collaborates with EDISON, EDSA, DATA SCIENCE, BDVA, and BDVe. This strong academic-led network provides a good balance to the Industry-led BDVA within the European Ecosystem.

Proposal Activity
During period 2, the members of the CoE Network have been involved in a number of collaborative actions to create new proposals within Horizon 2020. A proposal is a competitive activity, but the network has provided useful connections between CoEs to form and participate in consortia. Given the competitive nature of the calls, the proposal activity takes place outside of the network. Two or more members of the network have been involved in consortia that have targeted the following calls.

- **ICT-26-2018-2020 (AI)**: The goal is a European AI-on-demand platform mobilising the European AI community to support businesses and sectors in accessing expertise, knowledge, algorithms, and tools to successfully apply AI thereby generating market impact.

- **ICT-48-2020 (Towards a vibrant European network of AI excellence centres)**: To ensure European strategic autonomy in such critical technology as AI, underpinning most of our future professional and private activities, with huge potential socio-economic impact, it is essential to reinforce and build on Europe’s assets in AI, including its world-class researcher community, in order to stay at the forefront of AI developments.

- **DT-ICT-05-2020 (Big Data Innovation Hubs)**: The challenge is to break "data silos" and stimulate sharing, re-using, and trading of data assets by launching a second-generation data-driven innovation hub, federating data sources and fostering collaborative initiatives with relevant digital innovation hubs. Also, to promote new business opportunities notably for SMEs as part of the Common European Data Space.

Indeed, many other proposals could have benefited from the networking opportunities provided by the network.

**Formal/Informal Meetings**
Meetings of the network members take place within both formal and informal settings. Workshops have been held to support proposal activity of the network at the major events of the BDV PPP, including EBDVF and the BDV Summit. More
informal meetings have taken place at other events, including the ICT Proposers Day and ICT Conference and Infoday for ICT Event.

More specifically:

- Meeting at EBDVF 2018 Vienna (Focus on participation with Big Data Digital Innovation Hubs Call)
- Meeting at BDV Summit 2019 Riga (Focus on participation with Big Data Digital Innovation Hubs Call, Working Session for GATE centre for best practice exchange)
- EBDVF 2019 Helsinki (Multiple meetings focusing on ICT-48 calls for AI Network of Excellence)
- ICT Proposers’ Day Helsinki 2019 (Multiple meetings focusing on ICT-48 calls for AI Network of Excellence)

**Best Practice Exchanges and Visits**

The last two members of the network have participated in a number of site visits between centres to support the exchange of best practices. These visits have been organised in a bottom (organic) manner between centres that share common interests and want to promote collaboration. Centres from Ireland, Spain, Bulgaria, Germany, Austria, and Sweden have participated in these activities. In addition to these visits, bi-lateral meetings between network members have taken place at events, including EBDVF, ICT, BDV Summit, and the DatSci Awards. We will continue to support and encourage these meetings within the final year of the project.
5 Big Data Value Education Hub (T4.2-UDE)

In this part, the activities and achievements of Task 4.2 are outlined.

5.1 Education Programmes

The Education Hub was first released in 2017 and is offered as part of the BDV PPP portal (https://www.big-data-value.eu/education-hub/) and is also accessible via the more “catchy” URL http://bigdataprofessional.eu/. The landing page (see Figure 12) provides access to the different offerings based on the students’ interest and education level, i.e., Master, PhD, or professional Training.
5.1.1 Master and PhD Programmes

By 2018, the Education Hub featured profiles of 181 European Master Programmes and 12 PhD Programmes. These profiles contain information about (see the example in Figure 13):

- the name of the programme
- a description
- the suggested number of semesters
- a list of modules
- the language of teaching
- whether or not there are fees
- the name of the university
the location of the university
- an URL to the original programme
- whether or not the university is a BDV associate

EIT DIGITAL MASTER’S PROGRAMME IN DATA SCIENCE

Figure 13: Education Programme Profile Information

The newly established Data Science Master’s offers a unique academic programme, whereby students can study data science, innovation, and entrepreneurship at leading European universities. In this programme, students will learn about scalable data collection techniques, data analysis methods, and a suite of tools and technologies that address data capture, processing, storage, transfer, analysis, and visualization, and related concepts (e.g., data access, pricing data, and data privacy).

5.1.2 Professional Education Programmes
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

By 2019, the Education Hub featured profiles of 62 Onsite and Online Training Programmes. These profiles contain information about:

▪ the name of the programme
▪ a description
▪ the language of teaching
▪ the training provider name
▪ an URL to the original programme

Furthermore, the training programmes were classified by the following technical priorities and thus can be filtered accordingly:

▪ Development and Engineering
▪ Data Analytics
▪ Data Visualisation
▪ Data Management
▪ Data Protection
▪ Data Processing

To improve growth and reduce the cost of maintenance for the Education Hub, a new feature was added, which lets the user submit all necessary information on any Master, PhD, or Professional Training Programme. This new feature can be used to register new big data programmes to the Education Hub or to update wrong or missing information about existing programmes.

5.2 Dissemination Activities

Various activities were done with respect to the dissemination and marketing of the Education Hub.
The EduHub was presented during Period 2 (July 18 - Oct 19) in particular during the following events:


▪ **UDE presentation**: European Big Data Value Forum 2019, Helsinki, Finland. October, 2019. “AI Opportunities in Mobility & Transport”.

▪ **UDE presentation**: paluno Tech-Talk: Wie KI die Softwaretechnik aufmischt, July 2019, „Maschinelles Lernen für Selbst-Adaptive Software“ (part of scientific summer „AI“ of the city of Essen)
- **UDE and answare webinar presentation:** BDV PPP Webinar “Big Data in the Transport Domain: Transforming Transport”, January 2019

Complementing the presentation of the Education Hub at events, it was advertised as part of the BDVe outreach activities as carried out by WP5.

In addition to that, the Education Hub was promoted by posting more than 100 YouTube comments on European, data-science-related videos. Also, in order to increase visibility and searchability, 50 suitable keywords and meta-tags were added to the Education Hub website.
6 The Recognition of Skills for Big Data Professionals (T4.3-UPM)

6.1 Task 4.3 Objectives

Task 4.3 involves the design and implementation of a data science skills recognition framework that addresses the principal needs of the main stakeholders in the data science ecosystem: data scientists, employers, and educators. From informal discussions with members of those groups, the following needs were identified:

1. Data Scientists need to:
   - Get their training and experience recognised
   - Build credentials which are easily verifiable online
   - Display their skills online
2. Employers need to:
   - Find candidates who meet their needs
   - Easily evaluate and verify the skills of their job candidates at different levels of granularity
3. Educators need to:
   - Gain publicity for and the branded recognition of their programmes
   - Recognise the partial completion of their training
4. Employers and Educators need to:
   - Build a community to ensure that training meets demands

6.2 A Brief Review of Activities in Period 1

Tasks 4.3 began work in Period 1 on a data science skills recognition framework with a focus on competencies gained in formal education, i.e., university courses in the form of Bachelors or Masters programmes. First, different tools for recognising skills were considered and compared. This study concluded with the selection of Open Badges as the basis for the new data science skills recognition framework (see BDVe D4.6: A Framework for the Recognition of Data Science Skills in Europe [12] for further details). D4.6 also includes a draft of the workflows for the program. Then the categories and levels for the recognitions were considered, and a business plan was presented (see D4.2 [11] for more information).

The main outcome of Period 1 was the initial proposal of a data science skills recognition program consisting of five badges, each with three levels (basic, intermediate and expert) based upon work from the EDISON project:

- Data Science Analytics
6.3 Comments Received from the Reviewers and their Impact on Work Proposed in Periods 2 and 3

In response to recommendations received from the reviewers of work finished in Period 1, Task 4.3 reformulated its work proposal to include a greater emphasis on a second skills recognition system focusing on non-formal learning while at the same time scaling back on work related to the initial skills recognition system focused on formal education. In particular, the following changes were made:

1. Regarding non-formal education, we began work on a second skills recognition framework. This second system is based upon the idea of endorsements. It is specifically tailored to include training offered through non-formal education; that is, online courses, Massive Open Online Course (MOOCs), in-house company training, etc. Though the focus of this framework is on non-formal education, it will be applicable to non-formal educational programmes offered by traditional academic institutions (universities) as well.

2. Regarding formal education, from the five badges originally defined, we will limit our scope to just the Academic Level of the Data Science Analytics Badge. This decision is consistent with feedback received from industry and academia, stating that this badge contains the skills most commonly required of data scientists. Also, the professional level of this badge (which
encompasses both non-formal and informal training by data science professionals) would have overlapped partially with the second framework.

Accordingly, Table 10 shows the new timeline of work for Period 2 and 3 of Task 4.3.

<table>
<thead>
<tr>
<th>Activity Description</th>
<th>Start Date</th>
<th>End Date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>4.3.1. A Badge-Based Framework for Recognising Data Science Skills in Formal Education</strong></td>
<td>07/2018</td>
<td>12/2020</td>
</tr>
<tr>
<td>4.3.1.1 Analyse the results of the First Pre-Pilot of the Data Science Analytics Badge</td>
<td>07/2018</td>
<td>11/2018</td>
</tr>
<tr>
<td>4.3.1.2 Conduct a Second Pre-Pilot of the Data Science Analytics Badge</td>
<td>12/2018</td>
<td>03/2019</td>
</tr>
<tr>
<td>4.3.1.3 Launch and run the Pilot of the Data Science Analytics Badge</td>
<td>04/2019</td>
<td>09/2019</td>
</tr>
<tr>
<td>4.3.1.4 Rollout the academic level of the Data Science Analytics Badge</td>
<td>10/2019</td>
<td>12/2020</td>
</tr>
<tr>
<td><strong>4.3.2 A Recognition System for Non-Formal Education in Data Science</strong></td>
<td>06/2019</td>
<td>12/2020</td>
</tr>
<tr>
<td>4.3.2.1 Prepare a first version of the logistics and criteria for the recognition of non-formal training in data science</td>
<td>06/2019</td>
<td>01/2020</td>
</tr>
<tr>
<td>4.3.2.2 Evaluation and revision of the proposal to recognise non-formal training in data science</td>
<td>02/2020</td>
<td>05/2020</td>
</tr>
<tr>
<td>4.3.2.4 Rollout the recognition of non-formal training in data science</td>
<td>06/2020</td>
<td>12/2020</td>
</tr>
</tbody>
</table>

**Table 10: Timeline of Task 3.4 For Periods 2 and 3**

Figure 15 graphically represents a summary of the task’s main outcomes after the reformulation of the work plan to take into consideration the reviewer’s comments. As can be seen in this figure, during Period 2 (month 19-36), the UPM worked both on formal education in data science as well as began working on a new framework to recognise non-formal education in data science.
During Period 2, we conducted two rounds of feedback and revision of the requirements of the Data Science Analytics Badge, with the result being version 1.0 of its requirements. Then we ran a pilot of the entire application to issue badges process in order to further validate the requirements of the badge and to test all steps of the workflow with a limited number of institutions. Three applications to issue badges were received, and two were accepted. Upon the successful completion of the pilot and minor changes from lessons learned, we published the first open call for applicants to issue the Academic Level of the Data Science Analytics Badge (v.1.0). This process is illustrated in Figure 16.
In Period 2, Task 4.3 also started working on a recognition strategy for non-formal education in data science. During this period, the work focused on identifying the needs of non-formal training programmes, students and employers. An initial proposal for the program was presented.

The next two sections provide further detail regarding the activities performed in Task 4.3 during Period 2.

### 6.5 A Badge-Based Framework for Recognising Data Science Skills in Formal Education

The specific tasks related to the recognition of Data Science Skills in formal education given in Table 10 during Period 2 are further detailed in Table 11.

<table>
<thead>
<tr>
<th>Activity Description</th>
<th>Start Date</th>
<th>End Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3.1 A Badge-Based Framework for Recognising Data Science Skills in Formal Education</td>
<td>07/2018</td>
<td>12/2020</td>
</tr>
<tr>
<td>4.3.1.1 Analyse the results of the First Pre-Pilot of the Data Science Analytics Badge</td>
<td>07/2018</td>
<td>12/2018</td>
</tr>
<tr>
<td>a. Analyse the results gathered from industry, academia and the BDVA Activity Group (AG) on the Data Science Badges</td>
<td>07/2018</td>
<td>11/2018</td>
</tr>
<tr>
<td>Activity Description</td>
<td>Start Date</td>
<td>End Date</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>b. Propose version 0.2 of the Data Science Analytics Badge</td>
<td>11/2018</td>
<td>12/2018</td>
</tr>
<tr>
<td>4.3.1.2. Conduct a Second Pre-Pilot of the Data Science Analytics Badge</td>
<td>12/2018</td>
<td>03/2019</td>
</tr>
<tr>
<td>a. Distribute version 0.2 of the Data Science Analytics Badge to industry and academia</td>
<td>12/2018</td>
<td>01/2019</td>
</tr>
<tr>
<td>b. Analyse the results gathered from industry, academia and the BDVA AG on version 0.2 of the Data Science Analytics Badge</td>
<td>02/2019</td>
<td>02/2019</td>
</tr>
<tr>
<td>c. Propose version 1.0 of the Data Science Analytics Badge</td>
<td>03/2019</td>
<td>03/2019</td>
</tr>
<tr>
<td>4.3.1.3 Launch and run the Pilot of the Data Science Analytics Badge</td>
<td>04/2019</td>
<td>09/2019</td>
</tr>
<tr>
<td>a. Propose the logistics of the application to issue badges and generate instructions and guidelines to be followed by the applicants and the reviewers</td>
<td>04/2019</td>
<td>06/2019</td>
</tr>
<tr>
<td>b. Prepare dissemination materials for the pilot (a new web page, a press release, a short academic paper, and various presentations)</td>
<td>05/2019</td>
<td>07/2019</td>
</tr>
<tr>
<td>c. Analyse and evaluate available badge issuing platforms</td>
<td>05/2019</td>
<td>07/2019</td>
</tr>
<tr>
<td>d. Review the applications for issuing the Data Science Analytics Badge</td>
<td>09/2019</td>
<td>10/2019</td>
</tr>
<tr>
<td>e. Finalise the setup of the badge issuing platform</td>
<td>10/2019</td>
<td>10/2019</td>
</tr>
<tr>
<td>f. First badges can be issued by accepted universities</td>
<td>10/2019</td>
<td></td>
</tr>
<tr>
<td>4.3.1.4 1st Open Call to issue the Academic Level of the Data Science Analytics Badge</td>
<td>10/2019</td>
<td>06/2020</td>
</tr>
</tbody>
</table>

Table 11: Detailed Timeline for activities related to Formal Education

Below we discuss the work conducted during Period 2.
6.5.1 The First Phase of the Pre-Pilot of the Data Science Analytics Badge

D4.2 [11] presented an initial proposal for the levels and requirements of the data science badges. At the end of Period 1, the first pre-pilot survey of v0.1 of the levels and requirements of the Data Science Analytics Badge was launched to get feedback from academia and industry. The requirements of this badge (v0.1) at the three levels (basic, intermediate and expert) are shown in Table 12.

<table>
<thead>
<tr>
<th>Data Science Analytics (v 0.1)</th>
<th>Basic Level</th>
<th>Intermediate Level</th>
<th>Expert Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSA.1</td>
<td>Choose and execute existing data analytics and predictive analytics tools.</td>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td>Design and evaluate predictive analysis tools to discover new relations.</td>
</tr>
<tr>
<td>DSA.2</td>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td>Select most appropriate statistical techniques and model available data to deliver insights</td>
<td>Assess and optimise organisation processes using statistical techniques.</td>
</tr>
<tr>
<td>DSA.3</td>
<td>Design and evaluate predictive analysis tools to discover new relations.</td>
<td>Analyse available data sources and develop a tool that works with complex datasets.</td>
<td>Assess, adapt, and combine data sources to improve analytics.</td>
</tr>
<tr>
<td>DSA.4</td>
<td>Name and use basic performance assessment metrics and tools.</td>
<td>Use multiple performance and accuracy metrics; select and use most appropriate for specific type of data analytics application.</td>
<td>Evaluate and recommend the most appropriate metrics, propose new for new applications.</td>
</tr>
<tr>
<td>DSA.5</td>
<td>Define data elements necessary to develop specified data analytics.</td>
<td>Develop specialised analytics to enable decision-making.</td>
<td>Design specialised analytics to improve decision-making.</td>
</tr>
<tr>
<td>DSA.5</td>
<td>Choose and execute standard visualisation.</td>
<td>Build visualisations for complex and variable data.</td>
<td>Create and optimise visualisations to influence executive decisions.</td>
</tr>
</tbody>
</table>

Table 12: Requirements of the Data Science Analytics Badge (v 0.1)

The feedback gathered from industry and academia, was discussed during the BDV Activity Group Meeting (AG28) in September 2018. The input gathered from these three sources allowed us to generate version 0.2 of the BDVe Data Science Analytics badge. Below we discuss each of these steps.
6.5.1.1 Feedback from Academia

As discussed in D4.2 [11], 10 universities were contacted to get feedback about v0.1 of the Data Science Analytics badge (Appendix C shows the questionnaires distributed to the academic institutions). From them, five replied with feedback about the badge. We will now present and analyse this feedback.

6.5.1.1.1 Academic Reply #1

This university answered “yes” to all of the questions regarding whether the mentioned skills were included in their program except for the skill “Create and optimise visualisations to influence executive decisions,” which they said is not included. They also provided answers regarding where evidence could be obtained for the acquisition of all of the required skills. It is interesting to note that:

- Their evidence for the related skills at the three different levels all come from the same source. I.e., “Choose and execute existing data analytics and predictive analytics tools”, “Identify existing requirements and develop predictive analysis tools”, “Design and evaluate predictive analysis tools to discover new relations.” all come from an assignment of the module on Machine Learning. This seems to imply that students obtain Expert Level experience with five of the six skills upon their first contact with the subject matter. As previously mentioned, the expert level of the sixth skill is not included in their program.

- The evidence for two of the six skills (#4 and #5) at all three levels come from their end of the year project. This seems to imply that at least two of the five skills are obtained through conducting a “hands-on” project.

They also provided the following feedback:

- Q: If your program does not include training for any of the required skills, do you think that it is unreasonable to expect that it does?
  A: The only skill that the MSc Computer Science – Data Analytics does not support is “Create and optimise visualisations to influence executive decisions”, as the current module on Data Visualisation does not support it. This is not unreasonable, as this is quite a complex skill that could be addressed at a higher level of education.

- Q: How much effort was needed to analyse your program and provide the above information for each badge?
  A: This was a complicated procedure, as each requirement is covered by different modules in the MSc Computer Science – Data Analytics. Modules are centrally described in terms of the topics covered, but not in different levels of expertise gained.

- Q: Do you have any other comments/suggestions regarding the badge groups or their required skills?
A: There was no evidence for each requirement that was taken from the MSc overall. Each requirement referred to a specific module, as this MSc contains a variety of modules that cater for different skillset needs concerning Data Science. What is also interesting is that to get this badge, a student would need to register for multiple models. From the initial analysis, no single module gives the required skills. That is not to say that a module could include an assignment to cover all of these points. For example, I teach a module on case studies for data analytics. I could, in theory, design some assessments to give the above badges and to cover all the levels. In fact, I would be willing to do this as a pilot.

6.5.1.1.2 Academic Reply #2

This university replied that they could not answer the questionnaire without first clarifying the content and wording of the required skills. They proceeded to comment on the content of each required skill as given below:

**Basic level**

- Q: Identify existing requirements and develop predictive analysis tools.
  A: These are two different types of skills that should be separated. Furthermore, developing predictive analysis tools is hardly a basic level skill.

- Q: Design and evaluate predictive analysis tools to discover new relations.
  A: Again, I think predictive analysis tools is misleading in this context (do you mean "predictive models"?). "Discover new relations" is too informal. What about descriptive models?

- Q: Name and use basic performance assessment metrics and tools.
  A: "Basic performance assessment metrics" should be "performance metrics," and I am not sure what you mean by "tools" in this context.

- Q: Define data elements necessary to develop specified data analytics.
  A: This is both unclear and incomplete (something appears to be missing after "data analytics").

- Q: Choose and execute standard visualisation.
  A: How do you execute visualisation? Do you mean to execute a tool (again)? What about knowledge about basic concepts and techniques?

**Intermediate level**

- Q: Identify existing requirements and develop predictive analysis tools.
  A: This was already present at the basic level.

- Q: Select most appropriate statistical techniques and model available data to deliver insights.
  A: "Most appropriate" is hard to assess. I suggest removing "most". Who creates/delivers insights? Reformulate.
Q: Analyse available data sources and develop a tool that works with complex datasets.
A: Two separate (but not clearly specified) types of skill again. Clarify "complex".

Q: Use multiple performance and accuracy metrics; select and use most appropriate for specific type of data analytics application.
A: Accuracy is a performance metric. The metrics are typically complementary, i.e., gives different views, and it is often not meaningful to say what is "most appropriate" (cf. comparing apples and pears).

Q: Develop specialised analytics to enable decision-making.
A: Not clear what differentiates this from the previous points.

Q: Build visualisations for complex and variable data.
A: Again, the terms ("complex" and "variable") have to be clearly defined

Expert level

Q: Design and evaluate predictive analysis tools to discover new relations.
A: This was already mentioned at the basic level.

Q: Assess and optimise organisation processes using statistical techniques.
A: I am not sure why this application area has to be considered at the expert level only. Is it really meaningful to point out this application area?

Q: Assess, adapt, and combine data sources to improve analytics.
A: Although the terms have to be clarified, I consider this to be a basic level skill.

Q: Evaluate and recommend the most appropriate metrics, propose new for new applications.
A: This needs to be reformulated, as part of it is already covered, and the new part is too informal. (To be able to propose is not sufficient.)

Q: Design specialised analytics to improve decision-making.
A: This is too shallow and does not consider any obvious additional skill.

Q: Create and optimise visualisations to influence executive decisions.
A: What is the point of separating out "executive" decisions, and exactly what are the criteria for these?

6.5.1.1.3 Academic Reply #3
This university replied that they do teach all of the required skills at all three levels in their program. They made some observations:

Skills at the basic level are all covered by assignments and questions on exams, with the exception being the first skill which is only covered by an assignment.

At the intermediate level, all skills are covered by assignments and questions on exams, while three could also be covered in group projects and an oral defence.

At the expert level three skills are covered by elective/optional classes, two (#3 and #5) could be covered in group projects and an oral defence and the last
(#6) in addition to the possibility of it being covered in group projects and an oral defence it could also be covered by an individual project. This seems to imply that not all students who graduate their program would obtain the expert level.

Finally, they provided the following relevant feedback:

- Q: Do you have any other comments/suggestions regarding the badge groups or their required skills?
  A: Sometimes, the description of the skill is a bit too vague, e.g. between these two at expert level: “Design and evaluate predictive analysis tools to discover new relations”, and “Design specialised analytics to improve decision-making” can be understood as being the somehow same, because if you design a predictive analysis tool to discover new relations, this can really help you to improve the decision making I believe. However, I feel that the efforts BDVe is doing here by classifying badges in the 5 competence groups are very relevant. It will be very helpful indeed to better qualify and position a curriculum track. Thanks!

6.5.1.1.4 Academic Reply #4

This university provides training for all of the required skills at three levels. Additionally, they identify the specific subjects and assignments that provide evidence regarding such knowledge.

With respect to the open questions, this university addressed the last two.

- Q: How much effort was needed to analyse your program and provide the above information for each badge?
  A: 1 month
- Q: Do you have any other comments/suggestions regarding the badge groups or their required skills?
  A: Big data infrastructures

6.5.1.1.5 Conclusions from Academic Results

The main issues which emerged from the feedback were summarised as follows (to be able to refer to them later in the document they have been numbered A.X)

- A.1. The skills were too centred on predictive models; they should also include skills related to descriptive models. (academic 2 response 2)
- A.2. The skills related to data preparation and visualisation should be made clearer. (academic 1 and 2 response 6)
- A.3. There appeared to be a mixing of concepts related to tools and models. (academic 2 response 2 and 3)
- A.4. Skills regarding visualisation seem to be centred on higher-level positions. (academic 2 response 6)
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

- A.5. Words such as “complex” were not sufficiently specific and should be removed or clarified. (academic 2 response 2 inter)
- A.6. In general, the process of evaluating whether a program teaches the required skills could require a significant effort. (academic 1 general comments)
- A.7. Providing evidence was not always straightforward. (academic 1 general comments)
- A.8. The description of the skills was vague. (academic 3 general comments)

We note that university #3 also mentioned the lack of skills related to Big Data Infrastructures (general comments). However, these skills belong to the badge related to Big Data Engineering, which was not part of this pre-pilot evaluation.

6.5.1.2 Feedback from Industry
As shown in D4.2 [11], eight industries were contacted to get feedback about the Data Science Analytics badge v0.1 (Appendix D shows the questionnaire distributed). Below we discuss the input provided by the five organisations that provided feedback.

6.5.1.2.1 Industry Reply #1
They defined three job titles:
1. Data scientist for big data solutions
2. Researcher on Big Data
3. Big data and AI developer

The next figures summarise the information provided by Industry reply #1 regarding the relevance of the required skills to the three positions.

![Figure 17: The relevance of basic level skills for Industry 1](image-url)
In general, in all levels, all of the skills had a relevance higher than three for the different job titles. Although there was not a clear relationship between job titles and levels, it seems that the Basic and Intermediate levels were more relevant for Big Data and AI Developer, while the Expert level was more relevant for Researcher on Big Data. Industry 1 also provided the following comments regarding the descriptions of the skills:

**Basic Level**
- Q: Design and evaluate predictive analysis tools to discover new relations.
A: Not entirely clear. Mix between design (which suggests further implementation) and evaluate (which may either suggest reuse of existing tools or evaluation, but with no specific implementation stage is hard to understand). I would suggest saying, “Design, implement and evaluate...” if you aim at the development of the solution, or alternatively “Evaluate and, if needed, design and implement...”

▪ Q: Name and use basic performance assessment metrics and tools
  A: perhaps use “benchmarking” instead of “performance.”

▪ Q: Define data elements necessary to develop specified data analytics.
  A: Not entirely clear. I’m not sure if you mean data preparation tasks here, such as data pre-processing, cleansing, etc. “Define data elements” is a bit confusing

**Expert Level**

▪ Q: Design and evaluate predictive analysis tools to discover new relations.
  A: Not entirely clear. See my comments above for the basic level.

▪ Q: Assess and optimise organisation processes using statistical techniques.
  A: Yes, but why focus only on statistical techniques?

▪ Q: Assess, adapt, and combine data sources to improve analytics.
  A: Yes, but I guess this is related to data preparation and preprocessing, which is a more well-known way to phrase it.

They also provided the following feedback:

▪ Q: Are there data science skills in the area of data analytics which your company typically requires which are not included in the requirements for the badges?
  A: There is no reference to horizontal domains. Sometimes we ask for data scientists with knowledge in some verticals (i.e. banking, energy...), which is a skill that is also required from our markets and sectors. We don’t focus specifically on predictive models. I miss skills related to descriptive analytics, prescriptive analytics, diagnosis, etc. In fact, we don’t usually mention this is our job positions, but rather we ask of specific techniques and tools, rather than focusing on types of data analytics.

▪ Q: Do you have any other comments/suggestions regarding the badge groups or their required skills?
  A: We don’t divide the candidates into three categories, but rather into two: junior and senior. So, three categories sound a bit weird from the point of view of our way we hire people. We do have more than three category levels once people are hired, but in the selection process, we don’t use this.

**6.5.1.2.2 Industry Reply #2**

This company defined the following three roles:
1. Data Scientist: analyses and enriches data by deploying various kinds of data analytics methods, models and algorithm; extracts and visualises business-relevant information, explores new data analytics-enabled business opportunities and discusses those with stakeholders.

2. Data Analyst: conducts an analysis of data using various methods and tools to extract information as a basis for decision making.

3. Data Steward: defines and implements strategies to enable end-to-end access to data.

They did not distinguish between the three levels of the skills when hiring, so they provided answers only for the basic level, as shown in Figure 20. Furthermore, they redefined the scale of answers as follows:

1. (not used)
2. Starter
3. Basic
4. Advanced Level
5. Expert Level

In general, from this survey, we conclude that:

- For the Data Scientist position, all the skills were considered to be advanced except for the first which was considered to be an expert level skill.
- For the Data Analysis position, all the skills were considered to be basic, except for the first, which was considered to be advanced.

![Industry 2: Basic Level](image-url)
For the Data Steward position, the first and the last skills were considered to be at the basic level while the others were all at the starter level. They did not provide any feedback regarding the skills themselves.

6.5.1.2.3 Industry Reply #3

This company did not define different job titles. Figure 21 shows the information they provided regarding the relevance of the required skills to their hiring, in general, using only the basic level. In this case, all of the skills had the highest relevance for them.

![Figure 21: The relevance of Basic Level skills for Industry 3](image)

Regarding the skills themselves, they provided the following feedback:

- **Q:** Define data elements necessary to develop specified data analytics.
  
  **A:** Not clear what “specified data analytics” is meant here.

- **Q:** Choose and execute standard visualisation.
  
  **A:** What is “standard visualisation”, and what will the “non-standard visualisation” be?

- **Q:** Develop specialised analytics to enable decision-making.
  
  **A:** Analytics what ....? Analytics tools or analytics methods .... Something is missing here. What is specialised analytics? Every analytics method is kind of specialised.

- **Q:** Assess and optimise organisation processes using statistical techniques.
  
  **A:** What are “organisation processes" in this context?

- **Q:** Design specialised analytics to improve decision-making.
  
  **A:** Analytics what? What is “specialised analytics”? See the same comment above.
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

Other comments:
- A: I’m not in favour of defining 3 different categories (basic, intermediate, and expert). I would more like to see a basic level and then several specific extensions like one extension for deep dive in visualisation or deep dive in tools, ....
- A: For me, the topics mentioned in this line and the 2 lines above are quite similar. I would combine them to just one line.

<table>
<thead>
<tr>
<th>Basic Level</th>
<th>Intermediate Level</th>
<th>Expert Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose and execute existing data analytics and predictive analytics tools.</td>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td>Design and evaluate predictive analysis tools to discover new relations.</td>
</tr>
<tr>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td>Select most appropriate statistical techniques and model available data to deliver insights</td>
<td>Assess and optimise organisation processes using statistical techniques.</td>
</tr>
<tr>
<td>Design and evaluate predictive analysis tools to discover new relations.</td>
<td>Analyse available data sources and develop a tool that works with complex datasets.</td>
<td>Assess, adapt, and combine data sources to improve analytics.</td>
</tr>
</tbody>
</table>

Table 13: Replies from people with Basic, Intermediate and Expert levels.

6.5.1.2.4 Industry Reply #4

They defined the following three job titles:
- Data Engineer: a person whose primary job responsibilities involve preparing data for analytical or operational uses
- Data Scientist: a person employed to analyse and interpret complex digital data, such as the usage statistics of a website, especially in order to assist a business in its decision-making
- Data Architect: a person whose primary job responsibilities involve defining how the data will be stored, consumed, integrated and managed by different data entities and IT systems

Figure 22, Figure 23 and Figure 24 show respectively the information regarding the relevance of the required skills to the three positions, and the three levels. Notice that for the basic level, there were two skills that they did not consider relevant.
Figure 22: The relevance of Basic Level skills for Industry 4

Figure 23: The relevance of Intermediate Level skills for Industry 4
According to this information, the Data Scientist position was very related to all of the skills in all of the three levels, but the best fit was with the Expert level. Most of the skills had a low relevance for Data Architect and Data Engineer jobs (most skills in the three levels are valued under 2). However, this company did not suggest that any skills were missing in the badges for these positions.

They provided the following feedback regarding the skills:

**Basic level**
- They eliminated the following two skills from this level:
  - Identify existing requirements and develop predictive analysis tools.
  - Design and evaluate predictive analysis tools to discover new relations.

**Expert level**
- In the skill: Assess and optimise organisation processes using statistical techniques, they proposed that the word “organisation” be removed.
- In the skill: Create and optimise visualisations to influence executive decisions, they proposed that “influence executive” be changed to “support”.

Furthermore, they commented:
- Q: Are the descriptions of the required skills for each of the badges easily understood?
  A: In general, please, link with concrete knowledge, methods, languages, tools, platforms, or infrastructure to ease its understanding.
- Q: Do you have any other comments/suggestions regarding the badge groups or their required skills?
A: To simplify in three groups.

6.5.1.2.5  Industry Reply #5

This company worked with three job titles:

- Business understanding
- Data exploration and analysis
- Data modelling

The relevance of the different skills for each job is presented in Figure 25, Figure 26, and Figure 27 for each of the three levels.

![Figure 25: The relevance of Basic Level skills for Industry 5](image)

![Figure 26: The relevance of Intermediate Level skills for Industry 5](image)
As we can see, most skills at the three levels were highly relevant to the Data Modelling and Data Exploration and Analysis job profiles. However, for Business Understanding positions, relevance was no higher than 2. This company provided some detailed comments about their Business Understanding positions that are presented below.

Regarding the skills, they only provided one comment for the basic level:

**Basic level**

In the skill: Define data elements necessary to develop specified data analytics, they mentioned: “What kinds of data elements? Provide some examples.”

Furthermore, they commented:

- **Q:** Are there data science skills in the area of data analytics which your company typically requires which are not included in the requirements for the badges?
  **A:** For the data science job title 1 (Business understanding), there are no specific required skills mentioned in the tables. The business understanding step is essential to understand what is the exact business goal to develop any data analysis and optimisation in an organisation.

- **Q:** Do you have any other comments/suggestions regarding the badge groups or their required skills?
  **A:** The required skills associated with the business understanding job would be the following:
have a firm grasp of the company's business goals and objectives as well as an understanding of the KPIs which let them know if they are heading in the right direction
- understand how the problem you solve can impact the business
- Communication ability. Able to clearly and fluently translate the technical findings to a non-technical team, such as the Marketing or Sales departments (for example, know how to create a storyline around the data to make it easy for anyone to understand)
- Critical thinking

6.5.1.2.6 Conclusions from Industrial Replies

According to the information provided by the industrial replies, we identified the following relevant issues, which have been number I.X:

- I.1. It was not necessary to provide three levels, maybe 2 would be enough. (all, except industry 5, in general comments)
- I.2. It was clear that the badge is for data scientist (in the sense of analytics). (all according to the definition of roles and importance of the skills)
- I.3. The separation between data preparation tasks and data analytics was not clear. (industry 1 reply 2, 3, 4 and industry 5 replies 1)
- I.4. Data analytics was too focused on “predictive” what about “descriptive”? (industry 1 and 3 replies 2 all levels)
- I.5. No skills in particular domains, this would be interesting for senior-level positions. (industry 1, and industry 5 general comments)
- I.6. No skills in particular tools. (industry 1 in general in all the replies makes this point)
- I.7. It looks like tools, techniques, and evaluation metrics were mixed together. (industry 1 and 2 replies to questions 3 and 4 all levels)
- I.8. The part regarding visualisation was not clear. (industry 1, 2 and 4 replies 6 all levels)
- I.9. Remove the “organisation” processes part. This was a mixture of domain and functionality in general. (industry 1 expert level question 6)
- I.10. The problem of having the same competence evaluated at different levels was not clear and should be clarified. (industry 1 general comments)

6.5.1.3 The September 2018 BDVA Activity Group Meeting (AG28)

The proposed badge program, as well as the previous results provided by industry and academy, were discussed during the BDVA Activity Group meeting (AG28) held in Brussels on September 12, 2018.

The objectives of the meeting were as follows:

- Give an overview of the proposed data science skills recognition framework
Present the initial version of the types and requirements for the recognitions
Conduct a working session to gather comments regarding the proposal
Discuss the survey results and reach a consensus regarding modifications to the initial proposal

Here is a summary of the general discussion regarding the badge program. In order to be able to make reference to these conclusions and comments later in the document, they have been labelled AG.X:

AG.1. Concerns were expressed regarding how we could ensure the quality of the education the badge holders receive. At this point, the focus of the upcoming pilot is on formal education programmes, and as such, we can assume that the accreditation required by the degrees in question would ensure the quality of the training. Obviously, there are quality differences between different accredited degree programmes but addressing this problem seems to be outside the scope of the BDVe project. It was suggested that the evidences included as part of the badge’s meta-data could also provide additional information regarding the quality of the training. We discussed two possibilities for the contents of these evidences:

- A copy of the actual work submitted by each student - The advantage of this option is that employers would be able to see real examples of the job candidate’s work. These examples would also give information regarding the level of the skill acquired. The main disadvantages include legal problems related to DP and the administrative overhead of gathering and publishing this evidence.

- A copy of the assignment, exam question, laboratory description etc. - The advantage of this option is that some information regarding the level of the training of the candidate would be provided while at the same time less work would be involved than in publishing each submission. The main disadvantage would be the possible unwillingness on the part of the educators to publish documents, including things like exam questions (which they might want to reuse in the future).

Given that the first version of the badge requirements had six required skills for each level of each badge, and that each of the six could come from six different evidences, it is possible that a significant amount of information would need to be organised and presented. Maybe it would be possible to select one or two to serve as representative examples to simplify the work of the educators.
It was suggested that in the future, a more ‘certificate’ type model could be implemented with common standardised exams. This would provide a common baseline that could be used to directly compare candidates. But this suggestion is rather complicated; exams would need to be written and evaluated, also proctoring facilities would need to be arranged (especially if the exams were administrated outside the universities/training facilities). Furthermore, it was not clear that a standardised exam would provide evidence of the depth of skills required. Candidates should not just know how to do something but should also be able to do it. It was agreed that we would revisit this suggestion in the future if the program is successful.

AG.2. Concerns were expressed regarding the quantity of administrative overhead on the part of the educators in the process of issuing badges. It was felt that faculty would be unwilling to commit to the additional work without compensation and/or clear evidence of the benefits of the programme.

We discussed some of the feedback received from the submitted surveys regarding the three levels of the Data Science Analytics Badge requirements (the badges which we will pilot in the spring) and reached the following conclusions:

AG.3 It was agreed that three levels for each badge were excessive, especially if students completing MSc-level studies would most typically be eligible to apply for the third level. It was suggested that we provide two levels, one which was more knowledge and academic training based and a second level, which would require a certain amount of real professional experience. This is somewhat related to EDISON’s notions of competence (the academic level) and skill (the professional level). This change was agreed upon.

AG.4. It was agreed that the absence of descriptive models was a problem (only predictive models are mentioned in the requirements). It was agreed that ‘descriptive’ should be added wherever predictive models are mentioned.

AG.5. The question of whether specific tools should be mentioned in the requirements was discussed. The observation was made that different organisations encourage their employees to use different tools and that even if very popular tools were selected, they would not necessarily satisfy everyone’s needs. Furthermore, once a data scientist knows how to solve a certain problem with a particular tool, training them to use another tool to solve the same problem is not difficult. Thus, it was concluded that mentioning specific tools was not necessary.
AG.6. The question of mentioning specific domains of knowledge was also discussed, and it was agreed that there were too many domains to be addressed and that, in any case, it was too specific.

AG.7. It was agreed that the separation between data preparation tasks and data analytics should be clarified. The importance of data preparation was emphasised.

6.5.1.4 Version 0.2 of the Data Science Analytics Badge: Levels and Requirements

The results of the pre-pilot survey provided two sets of recommendations regarding the badges and their requirements:

- Recommendations provided by universities (see section Error! No se encuentra el origen de la referencia.).
- Recommendations provided by industry (see section Error! No se encuentra el origen de la referencia.).

Those recommendations, as well as other issues, were discussed during the September 2018 BDVA Skills AG meeting (see section 6.5.1.3).

Below we provide a consolidated list of all the recommendations to ensure that they are all addressed in the levels and requirements of the new version (0.2) of the Data Science Analytics badge.

To identify overlaps between academic and industrial recommendations, we begin with Table 14, which shows the recommendations provided by industry (rows) and by universities (columns) with crosses in places where the recommendations coincide. As can be seen, there were four very similar suggestions. This is the case of:

- (A2 and I3)
  - The skills related to data preparation and visualisation should be made clearer
  - The separation between data preparation tasks and data analytics was not clear

- (A1 and I4)
  - The skills were too centred on predictive models; they should also include skills related to descriptive models
  - Data analytics was too focused on “predictive” models. What about “descriptive” models?

- (A3 and I7)
  - There appeared to be a mixing of concepts related to tools and models
  - It looks like tools, techniques, and evaluation metrics were mixed together

- (A2 and I8)
  - The skills related to data preparation and visualisation should be made clearer
  - The part regarding visualisation was not clear
Table 14: Consolidation of Recommendations provided by academia and industry

Table 15 summarises the consolidated recommendations provided by industry and universities, along with the recommendations regarding their inclusion or not from the AG. Note that in this table, we have removed the general comments related to the process itself (A7 and A8). These recommendations will be considered at a later date. In designing the new version of the levels and requirements of the Data Analytics Badge, the consolidated recommendations for which the AG did not explicitly provide feedback, as well as those for which the AG provided agreement, will be considered. These recommendations are marked with a check in the last column. (As such, I5 and I6, will not be considered based upon the disagreement of the AG.)

<table>
<thead>
<tr>
<th>Academic/Industrial Recommendations</th>
<th>BDVA Activity Group Recommendations</th>
<th>Will it be considered in the revision?</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.1 It is not necessary to provide three levels, maybe 2 would be enough</td>
<td>AG.3. It was agreed that three levels for each badge were excessive, especially if students completing MSc-level studies would most typically be eligible able to apply for the third level. It was suggested that we provide two levels, one which was more knowledge and academic training based and a second level, which would require a certain amount of real professional experience.</td>
<td>✓</td>
</tr>
<tr>
<td>I.2. It was clear that the badge is for data scientist (in the sense of analytics).</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Academic/Industrial Recommendations</td>
<td>BDVA Activity Group Recommendations</td>
<td>Will it be considered in the revision?</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>I.3.-A.2. The separation between data preparation tasks and data analytics was not clear.</td>
<td>AG.7. It was agreed that the separation between data preparation tasks and data analytics should be clarified. The importance of data preparation was emphasised.</td>
<td>✓</td>
</tr>
<tr>
<td>I.4.-A.1. Data analytics was too focused on &quot;predictive&quot; what about &quot;descriptive&quot;?</td>
<td>AG.4. It was agreed that the absence of descriptive models was a problem (only predictive models are mentioned in the requirements). It was agreed that 'descriptive' should be added (or predictive removed) wherever predictive models are mentioned.</td>
<td>✓</td>
</tr>
<tr>
<td>I.5. No skills on particular domains, this would be interesting for senior-level positions</td>
<td>AG.6. The question of mentioning specific domains of knowledge was also discussed, and it was agreed that there were too many domains to be addressed and that, in any case, it was too specific.</td>
<td>✗</td>
</tr>
<tr>
<td>I.6. No skills on particular tools</td>
<td>AG.5. Thus it was concluded that mentioning specific tools was not necessary.</td>
<td>✗</td>
</tr>
<tr>
<td>I.7. It looks like tools, techniques, and evaluation metrics were mixed together</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>I.8. The part regarding visualisation was not clear</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>I.9. Remove the &quot;organisation&quot; processes part. This was a mixture of domain and functionality in general</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>I.10. The problem of having the same competence evaluated in different levels was not clear and should be clarified</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>AG.4. Skills regarding visualisation seem to be centred on higher-level positions.</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>A.5. Words such as &quot;complex&quot; were not sufficiently specific and should be removed or clarified</td>
<td>✓</td>
</tr>
</tbody>
</table>
Table 15: The final list of recommendations to be considered in the revision of the Data Science Analytics badge

### 6.5.1.4.1 Description of the levels and requirements for version 0.2 of the Data Analytics Badge

Table 16 shows the initial proposal (version 0.1) of the levels and requirements of Data Science Analytics Badge, along with the new version (0.2) of the requirements according to the recommendations shown in Table 15. Notice that Table 16 provides the same skills for the Academic Level and the Professional Level. Both share the same description but will require different evidences, from an academic point of view or professional experience, respectively (See AG.3). Skills for the new badge have been numbered for later reference and discussion (DSA.X)

<table>
<thead>
<tr>
<th>Required Skills: Data Science Analytics</th>
<th>Version 0.1 (From the Edison EDSF)</th>
<th>Version 0.2 (based upon feedback)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Level</td>
<td>Intermediate Level</td>
<td>Expert Level</td>
</tr>
<tr>
<td>Choose and execute existing data analytics and predictive analytics tools.</td>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td>Design and evaluate predictive analysis tools to discover new relations.</td>
</tr>
<tr>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td>Select most appropriate statistical techniques and model available data to deliver insights</td>
<td>Assess and optimise organisation processes using statistical techniques.</td>
</tr>
<tr>
<td>Design and evaluate predictive analysis tools to discover new relations.</td>
<td>Analyse available data sources and develop a tool that works with complex datasets.</td>
<td>Assess, adapt, and combine data sources to improve analytics.</td>
</tr>
</tbody>
</table>
6.5.2 The Second Phase of the Pre-Pilot of the Data Science Analytics Badge

With the aim of assessing version 0.2 of the badge, a new evaluation process was planned. This process took place from December 2018 till March 2019. Our aim was that this process provides rapid feedback about the changes in the badge. So unlike in the first phase of the pre-pilot, this second phase of the pre-pilot took only three months, and the number of institutions contacted was much smaller. In this second phase of the pre-pilot, we considered two main stakeholders:

- Universities, as they will need to train their students to acquire the skills needed for each badge
- Members of the BDVA Task Force 9 on Skills (representatives of industries that need to hire professionals with data science skills).

First, universities assessed the proposal for the badge, and the results of this evaluation were discussed with the Task Force. With the information gathered from

---

### Table 16: New Formulation of Data Science Analytics Badge (v 0.2)

<table>
<thead>
<tr>
<th>Required Skills: Data Science Analytics</th>
<th>Version 0.1 (From the Edison EDSF)</th>
<th>Version 0.2 (based upon feedback)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basic Level</strong></td>
<td><strong>Intermediate Level</strong></td>
<td><strong>Expert Level</strong></td>
</tr>
<tr>
<td>Name and use basic performance assessment metrics and</td>
<td>Use multiple performance and</td>
<td>Evaluate and recommend the most</td>
</tr>
<tr>
<td>tools.</td>
<td>accuracy metrics; select and use</td>
<td>appropriate metrics, propose new</td>
</tr>
<tr>
<td></td>
<td>most appropriate for specific type of data analytics application.</td>
<td>for new applications.</td>
</tr>
<tr>
<td>Define data elements necessary to develop specified</td>
<td>Develop specialised analytics to</td>
<td>Design specialised analytics to</td>
</tr>
<tr>
<td>data analytics.</td>
<td>enable decision-making.</td>
<td>improve decision-making.</td>
</tr>
<tr>
<td>Choose and execute standard visualisation.</td>
<td>Build visualisations for complex and variable data.</td>
<td>Create and optimise visualisations to influence executive decisions.</td>
</tr>
</tbody>
</table>

---

95
this second phase of the pre-pilot, we produced revisions of the requirements for the badge, generating version 1.0 of the Data Science Analytics badge. Table 17 shows a summary of the schedule of the second phase of the pre-pilot:

<table>
<thead>
<tr>
<th>Task</th>
<th>December 2018</th>
<th>January 2019</th>
<th>February 2019</th>
<th>March 2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>University Survey – phase 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis of results</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDVA Task Force Discussion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version 1.0 Data Science Analytics Badge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 17: Schedule for the Second Phase of the Pre-Pilot

6.5.2.1 Feedback from Academia

A contact person from each university received a brief summary of the goals of the badge program along with the most recent version of the Data Science Analytics badge. This evaluation package is shown in Appendix E: BDVe’s Data Science Badges – 2nd Pre-Pilot Feedback from Academia.

The universities that participated in this second phase of the pre-pilot were those included in EIT Digital’s Data Science MSc’s Program:

- The University of Southern Denmark (SUD)
- Eötvös Loránd University (ELTE)
- Liepājas Universitāte (LIEPU)
- Deggendorf Institute of Technology (DIT)
- University of Agder (UIA)
- Warsaw University of Technology (PW)

The aim of those assessments was three-fold:

- Increase the awareness of the program, beginning with the first pre-pilot, to include the partners of one of the most noted data science MSc’s consortia in Europe
- Promote the incorporation of new learning outcomes in those programmes, in case some relevant learning outcomes are not satisfied
- Get feedback about the review process regarding the kinds of material to be requested of badges applicants, and about the requirements of the review process.

6.5.2.1.1 Academic Reply #1

This university answered “yes” to all of the questions regarding whether the mentioned skills were included in their program. They also provided detailed answers
regarding the evidences for the acquisition of all of the required skills. It is interesting to note that:

- In most of the skills, evidences come from different modules or subjects. All of them are required to get the MSc certification.
- Most evidences were generated through practical projects that students have to prepare. The form contains detailed information about how these projects are run; however, it was not indicated whether those projects are individual or group assignments.

They also provided the following feedback:

- Q: If your program does not include training for any of the required skills, do you think that it is unreasonable to expect that it does?
  A: N/A.
- Q: How much effort was needed to analyse your program and provide the above information for each badge?
  A: The effort was reasonable and seemed justified to us.
- Q: Do you have any other comments/suggestions regarding the badge groups or their required skills?
  A: I think it could be of potential value to include one-two more skills or make more precise existing skill questions clarify whether the course being analysed includes Big Data analytics and machine learning. For individual objectives, we could imagine providing the set of columns such as module(s) related to the objective, the form of assessment, the status of the modules (some of the modules could be elective). Hence, the table could have more than two columns to promote more structured responses.

6.5.2.1.2 Academic Reply #2

This university completed the questionnaire for two programmes, a bachelor and an MSc’s degree. However, instead of providing concrete details about the evidences they provided the percentages of the different courses that covered the corresponding skills. It is interesting to mention that:

- For the bachelor’s program, all skills were covered, with the exception of DSA.3. (Assess, adapt, and combine data sources to improve analytics). Most skills were covered in courses related to statistics and mathematics. Only a small percentage of those courses were devoted to the skills, only about 20%.
- For the MSc’s program, all the skills are covered. Courses that address them are related to quality management and business intelligence; always the same courses for all skills and the corresponding percentages are not high, around 10%.

No extra comments were provided for either program.
6.5.2.1.3 Conclusions from the Academic Replies

From the results gathered, we can conclude the following:

- Apart from one suggestion to include the specific names of data science techniques in the skills (general comments, University 1), in this second assessment, we did not get any other comments about the descriptions of the skills. From this, we conclude that the academics that reviewed them did not find special difficulties in understanding their content or meaning.
- We should provide some guidelines to help applicants select the required evidence in order to avoid different universities providing different results (like the details of the courses (University 1) or the percentages (University 2)).

6.5.2.2 Feedback from the BDVA Skills Task Force

In order to validate the comments received from the universities, we contacted two members of the BDVA Task Force 9 on Skills. The members of the Task Force that participated in the discussion had participated in the first phase of the pre-pilot, so they were already aware of the proposed data science skills recognition framework, as well as our work during the first phase of the pre-pilot. So, the main goal of this discussion was to present the comments received from the universities and to reach a consensus regarding modifications to version 0.2 of the badge.

We provided them with version 0.2 of the badge with two possible descriptions for D.A.1 and D.A.2 that included changes resulting from the comments from universities, as given in Section 6.5.2.1. Table 18 shows these two options for each of the descriptions.

<table>
<thead>
<tr>
<th>Required Skills: Data Science Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DSA.1.</strong> Identify existing requirements to choose and execute the most appropriate data discovery technique to solve a problem depending on the nature of the data and the goals to be achieved. Or:</td>
</tr>
<tr>
<td><strong>DSA.1.</strong> Identify existing requirements to choose and execute the most appropriate data discovery technique (machine learning, statistics, databases, ...) to solve a problem depending on the nature of the data and the goals to be achieved.</td>
</tr>
<tr>
<td><strong>DSA.2.</strong> Select the most appropriate statistical techniques to understand and prepare data prior to modelling to deliver insights. Or:</td>
</tr>
<tr>
<td><strong>DSA.2.</strong> Select the most appropriate techniques (machine learning, statistics, databases, ...) to understand and prepare data prior to modelling to deliver insights.</td>
</tr>
<tr>
<td><strong>DSA.3.</strong> Assess, adapt, and combine data sources to improve analytics.</td>
</tr>
</tbody>
</table>
Required Skills: Data Science Analytics

| DSA.4. | Use the most appropriate metrics to evaluate and validate results, proposing new metrics for new applications if required. |
| DSA.5. | Design and evaluate analysis tools to discover new relations in order to improve decision-making. |
| DSA.6. | Use visualisation techniques to improve the presentation of the results of a data science project in any of its phases. |

Table 18: Alternatives for DSA.1 and DSA.2 of the Data Science Analytics Badge

The conclusions of the discussion held with the Task Force members regarding DSA.1 and DSA.2 are as follows:

- Regarding DSA.1, the TF members contacted preferred "DSA.1. Identify existing requirements to choose and execute the most appropriate data discovery technique to solve a problem depending on the nature of the data and the goals to be achieved." They argued that adding "(machine learning, statistics, databases, ...)" might be limiting, and new methods could evolve over time, and that would mean that the text would need to be revised.

- Regarding DSA.2, they preferred a mix between the two options presented, "DSA.2. Select the most appropriate techniques (machine learning, statistics, databases, ...) to understand and prepare data prior to modelling to deliver insights." They suggested removing "(machine learning, statistics, databases, ...)". They did not like the first option for this skill; they wanted to emphasise the fact that this skill could refer to any technique not only "statistical techniques".

6.5.2.3 Version 1.0 of the Data Science Analytics Badge

Table 19 shows version 1.0 of the required skills for the Data Science Analytics badge.

Required Skills: Data Science Analytics (version 1.0)

| DSA.1. | Identify existing requirements to choose and execute the most appropriate data discovery technique to solve a problem depending on the nature of the data and the goals to be achieved. |
| DSA.2. | Select the most appropriate techniques to understand and prepare data prior to modelling to deliver insights. |
| DSA.3. | Assess, adapt, and combine data sources to improve analytics. |
| DSA.4. | Use the most appropriate metrics to evaluate and validate results, proposing new metrics for new applications if required. |
| DSA.5. | Design and evaluate analysis tools to discover new relations in order to improve decision-making. |
Required Skills: Data Science Analytics (version 1.0)

DSA.6. Use visualisation techniques to improve the presentation of the results of a data science project in any of its phases.

Table 19: Data Science Analytics Badge Version 1.0

Table 20 describes how the different recommendations presented in Table 15 were considered when defining version 1.0 of the badge.

<table>
<thead>
<tr>
<th>Academy/Industry Recommendations</th>
<th>Incorporation in the new badges</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.1 It is not necessary to provide three levels, maybe 2 would be enough</td>
<td>Only two levels, the academic and professional, were considered in the new version of the badge</td>
</tr>
<tr>
<td>I.2. It is clear that the badge is for data scientist (in the sense of analytics)</td>
<td>No changes needed</td>
</tr>
<tr>
<td>I.3.-A.2. The separation between data preparation tasks and data analytics is not clear</td>
<td>DSA.2 was restated in order to clarify this separation</td>
</tr>
<tr>
<td>I.4 - A.1. Data analytics is too focused on “predictive” what about “descriptive”?</td>
<td>The use of “analytics” was made more general by removing the references to “predictive.”</td>
</tr>
<tr>
<td>I.7 – A.3. It looks like tools, techniques, and evaluation metrics are mixed together</td>
<td>A distinction was clearly made: DSA.1 and DSA.2 refer to techniques, while DSA.4 refers to metrics and DSA.5 refers to tools</td>
</tr>
<tr>
<td>I.8. The part regarding visualisation is not clear</td>
<td>DSA.6 was rewritten to clarify this skill</td>
</tr>
<tr>
<td>I.9. Remove the “organisation” processes part. This is a mixture of domain and functionality in general</td>
<td>All references to application domains were removed</td>
</tr>
<tr>
<td>I.10 The problem of having the same competence evaluated in different levels is not clear and should be clarified</td>
<td>Using two levels with the same skills but with different evidences addresses this issue</td>
</tr>
<tr>
<td>A.4. Skills regarding visualisation seem to be centred on higher-level positions.</td>
<td>DSA.6 now includes skills related to visualisation that are generally required for all data scientists</td>
</tr>
<tr>
<td>A.5. Words such as “complex” are not sufficiently specific and should be removed or clarified</td>
<td>All of the skills were rewritten to avoid this word and to replace it with more concrete terms</td>
</tr>
<tr>
<td>A. 6 The description of the skills is vague</td>
<td>All of the skills were rewritten to make them clearer and more concise.</td>
</tr>
</tbody>
</table>

Table 20: Matching Between Recommendations and v1.0 of the Badge Requirements

At this time, the design of the Data Science Badges was also modified, with the text “academic” and “professional” replacing the stars (see Figure 28).
6.5.3 The Pilot of the Academic Level of the Data Science Analytics Badge

Once version 1.0 of the Data Science Analytics badge was completed, we prepared a pilot of the process for applying to issue this badge at the academic level. Figure 29 shows the steps and roles involved in the process of issuing the badges.

- A committee of people from both industry and academia collaborated to design the badges. In our case, this is the process followed for the definition of v1.0 of the Data Science Analytics Badge, as described in the last section.
- Interested institutions/educators can apply to issue the badge. This application includes submitting evidence that shows that they provide their students with the skills required by the badge. If their application is approved, the institutions can issue the badge directly to their students.
- Students can display the recognitions online: in their CV, in social networks, etc.
- Interested employers can: verify that the badge is valid and use its metadata to access relevant information. The metadata includes requirements to earn the badge and evidence, which shows that the candidate has the required skills.

Before starting the pilot, several preliminary activities were necessary. First, the process by which institutions can apply for the badge needed to be defined in detail.
and the assessment criteria used to evaluate whether the institutions provide sufficient evidence to show the acquisition of the skills needed to be established. Apart from this, we followed the review process as defined in D4.6. *A Framework for the Recognition of Data Science Skills in Europe* [12], summarised in Figure 30.

![Figure 30](image_url)

**Figure 30: Summary of the process for applying and issuing academic badges**

The next sections give further details of these activities.

### 6.5.3.1 Applying to Issue the Badge

Only approved programmes can issue the Academic Level of the BDV Data Science Analytics Badge. To obtain this approval, a program must apply to issue the badge by providing evidence to show how they can ensure that their students have acquired the corresponding skills.

To simplify the application process, we prepared a template, *BDV Data Science Analytics Badge Academic Level Template for Badge Issuing Applications (Appendix G)*, to help institutions to collect specific details about how each of the skills required by the badge is acquired by the program’s students. Table 21 shows the information that institutions must provide regarding the first skill of the badge.
### DSA.1. Identify existing requirements to choose and execute the most appropriate data discovery techniques to solve a problem depending on the nature of the data and the goals to be achieved

<table>
<thead>
<tr>
<th>Name of the course or activity</th>
<th>Optional/ Mandatory</th>
<th>The average number of hours that students spend in the course/ activity acquiring the skill</th>
<th>Evidence used to evaluate the acquisition of the skill (exam, assignment, …)</th>
<th>Name of the evidence file</th>
<th>The minimum level required to demonstrate the acquisition of the skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course 1</td>
<td></td>
<td>X hours</td>
<td>Exam 1, questions 2-3, Assignment 4, …</td>
<td>Add as many rows as needed</td>
<td></td>
</tr>
<tr>
<td>Course 2</td>
<td></td>
<td>Y hours</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Add as many rows as needed*

**Minimum number of hours that the student dedicates to acquiring the skill**

**Comments (optional)**

---

**Table 21: Extract of the template to collect evidences related to the Data Science Analytics Badge, first skill**

As can be seen in Appendix G, this document also requires that institutions provide some general information about the applicant’s program. Institutions send this completed application, along with the different evidence files referenced by the application to bdv.badges@big-data-value.eu.

Specific guidelines to help an institution to complete the application were also provided in a document named the *BDV Data Science Analytics Badge Academic Level Guideline for Applicants* (Appendix H).
6.5.3.2 Reviewing Applications for Issuing the Badge
Each application is independently reviewed by two experts that assess the submitted application as well as the evidence files following the Guidelines for Reviewers document (Appendix K).

Based on the information presented in an application, the acquisition of each of the skills of the Data Science Analytics Badge is scored as follows by each reviewer:

▪ 0 — Fail: The application fails to address the skill, or the acquisition of the skill cannot be assessed due to missing or incomplete information. (‘Obvious clerical errors’ are not considered to result in incomplete information.)
▪ 1 — Poor: the skill is inadequately addressed, or there are serious inherent weaknesses.
▪ 2 — Fair: the application broadly addresses the skill, but there are significant weaknesses.
▪ 3 — Good: the application addresses the skill well but with a number of shortcomings.
▪ 4 — Very good: the application addresses the skill very well but with a small number of shortcomings.
▪ 5 — Excellent: the application successfully addresses all relevant aspects of the skill; any shortcomings are minor.

Those scores are assigned based upon the following criteria:

▪ Is the average total number of hours that a student dedicates to the skill sufficient?
▪ Are the activities chosen to evaluate the skill appropriate for determining whether a student has acquired the skill?
▪ Is the level of achievement expected of students in line with the industry’s expectations for new graduates?
▪ In each evaluation activity, is there a clear description of the tasks that will be performed for students to obtain the skill?

Each reviewer provides a Final Score for the application by summing up the scores of each of the skills. The reviewer recommends that the applicant program be able to issue the badge if the program is awarded a total of at least 20 points, and no skill is scored as less than 3. If the final score of an application is between 20 and 24, the reviewer has the option to recommend that the badge issuing period of the program be limited and that the program be required to resubmit another application to issue badges in the following year.

Finally, the reviewer provides an Overall Recommendation for Improvement with a summary of his/her evaluation. This overall recommendation should be consistent with the final score provided and will be given to the applicant.
In order to provide this evaluation, reviewers are asked to complete the *BDV Data Science Analytics Badge, Academic Level, Evaluation Template* (Appendix J).

A final decision is made if the recommendations of the two reviewers coincide. If the two reviewers are not able to reach a consensus, a third reviewer is asked to participate in the process. Reviewers participating in this process must agree to the *Code of Conduct for Badge Issuing Application Reviewers*, as given in Appendix I. 

BDVe project members and two members of the BDVA Task Force 9 on Skills were asked to review the evaluation criteria along with all the documents provided for both the reviewers and the applicants. Their comments were taken into account, and a consensus was reached regarding the final version of all of these materials.

### 6.5.3.3 Web Site for the Pilot for Issuing the Data Science Analytics Badge

In the BDVe web site, we publicly provided all the information related to applying for badges as well as the review of those applications. This web page’s URL is [http://www.big-data-value.eu/skills/skills-recognition-program/](http://www.big-data-value.eu/skills/skills-recognition-program/). Figure 31 shows the main page, where a brief description of the background of the badges is presented, as well as a link to the application page.

![Figure 31: Screenshot of the skills recognition program web site with the home](image)

Information about the deadline and eligible institutions was also presented, as shown in Figure 32.
CALL FOR ACADEMIC LEVEL DATA SCIENCE ANALYTICS BADGE ISSUERS

HOW TO APPLY?
Application forms and related documents can be found [here].
Organizations are requested to contact us to express their interest in participating in the pilot.

DEADLINE FOR THE SUBMISSION OF APPLICATIONS
MAY 30, 2019

ELEGIBILITY
Any public or private educational institution active in the field of data science education and located in any of the following countries can apply:

- Member States of the European Union,
- EFTA/EEA countries: Iceland, Liechtenstein and Norway,
- EU candidate countries: Turkey, the former Yugoslav Republic of Macedonia and Serbia.

FULL DETAILS
Documents to help programs applying to issue badges

Figure 32: Deadline and eligibility details in the Skills Recognition Program
Both applicants and reviewers can download the corresponding documents as shown in Figure 33.
The web page also includes a FAQ with questions and their corresponding answers to clarify potential doubts that might arise during the application process. Figure 34 shows a screenshot of the FAQ, while the whole set of questions is given in Appendix L.

---

**Figure 33:** Information about the documents for applicants and reviewers

Documents to help programs applying to issue badges

- Guidelines for applicants — [Download](#)
- Template for issuing applications — [Download](#)

Documents for badge application reviewers

- Guidelines for reviewers — [Download](#)
- Evaluation template — [Download](#)
- Code of Conduct — [Download](#)

Further details of the badge program can be found [here](#).

Please go to the [FAQ](#) for answers to common questions.

---

**Figure 34:** Screenshot of the first part of the FAQ

Who can apply to issue BDV badges?

Any EU based university level educational institution can apply to issue BDV badges. Applicant programs will most typically be at the Master’s level but undergraduate programs can also apply. Approval to issue a badge is given to a particular degree program. Different programs at the same institution must apply separately.

Can non-EU institutions apply to issue badges?

At the moment, only institutions within the EU can issue BDV badges.

Do separate programs in multiple degrees (e.g., EIT Digital) need to apply separately?

Yes, as each individual program can evaluate skills in different ways, each program needs to apply separately.

Must all students of an approved badge issuing program be eligible to apply for a badge?

Badges are issued on a student by student basis, thus the required skills for a badge can be acquired in elective courses or in optional activities within a program. However, in this case only students who
6.5.3.4 Running the Pilot for Issuing the Academic Level of the Data Science Analytics Badge

Once the infrastructure for the pilot was complete, we began the process. In order to attract as many institutions as possible for the pilot, institutions applying during the pilot were evaluated at no cost, and if their applications were accepted, they were provided with access to the badge issuing platform at no cost (these costs were assumed by the project).

Table 22 shows the timeline of the pilot process. The deadline for sending applications was June 15th, 2019.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Applications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notifications sent to Applicants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appeal Period</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparation of the platform for Issuing the Badges</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accepted Universities can begin to Issue the Badge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 22: The Pilot Timeline

6.5.3.4.1 Dissemination of the Pilot

The pilot was open to all, but individual invitations were sent to 20 institutions that teach MSc programmes in Data Science or related fields. In particular, the following institutions were individually invited to participate in the pilot (this list contains participants in EIT Digital’s Data Science MSc’s program and interested participants from ICT 2018 and EDBVF 2018):

- Aristotle University of Thessaloniki, Greece
- TU Berlin, Germany
- University of Southern Denmark, Denmark
- Eötvös Loránd, Hungary
- Deggenendorf Institute of Technology, Germany
- Liepaja University, Latvia
Universidad Politécnica de Valencia (Spain)
• Royal Institute of Technology (Sweden)
• Eindhoven Institute of Technology (The Netherlands)
• Univeristé Nice Sophia Antipolis (France)
• Université Paul Sabatier (France)

Also, the following institutions from the EduHub were contacted by Paluno and invited to participate in the pilot:
• Athens University of Economics and Business (Greece)
• Charles University (Czech Republic)
• University of Tartu (Estonia)
• University College Dublin (Ireland)
• Budapest University of Technology and Economics (Hungary)
• Universidad de Nova de Lisboa (Portugal)
• University of Eastern Piedmont Amedeo Avogadro (Italy)
• Vytautas Magnus University (Lithuania)
• Varna Free University (Bulgaria)

A press release of the pilot was included in the BDVA Newsletter, and messages were posted on Twitter
(https://twitter.com/BDVA_PPP/status/113074335270596608) and LinkedIn
(https://www.linkedin.com/feed/update/urn:li:activity:6536511685258056). One of the tweets sent is shown in Figure 35.

Figure 35: Twitter message regarding the BDVe Pilot

Finally, the UPM distributed information or gave presentations introducing the BDVe badge program and pilot process in several forums:
• BDVA Skills Task Force meeting – January 2019
• Expert Workshop Supporting Specialised Skills Development: Big Data, IoT and Cybersecurity for SMEs (Workshop 3 February 2019, and Workshop 4 – April 2019)
6.5.3.5 Results of the Pilot for Issuing the Academic Level of the Data Science Analytics Badge

During the pilot, three institutions submitted applications to issue the Academic Level of the Data Science Analytics Badge. All three applications were complete (the universities used the template and included all the required evidences as described in Section 6.5.3.1). The review process described in 6.5.3.2 was followed, and the result was that: one application was accepted, another application was accepted with comments, and the last application was rejected. Due to privacy and confidentiality issues, the particular details of the institution that was evaluated negatively was kept confidential. The institutions and programmes that were granted the right to issue the badge were:

- MSc on Big Data Analytics. Universidad Politécnica de Valencia (Spain)
- Data and Web Science MSC Program. Aristotle University of Thessaloniki (Thessaloniki, Greece)

Although the number of applications participating in the pilot was not large, it was enough to perform a feasibility study and a proof of concept of the program. It allowed us to check that both parts of the process, the application side, and the review side were mature enough to proceed to organise the first open call for institutions to issue the Academic Level of the Data Science Analytics Badge.

6.5.4 The Platform for Issuing BDV Data Science Badges

In this section, we will consider issues related to the issuing of the Open Badges. First, we will look at some details of v2.0 of the Open Badge Standard. From this standard and other considerations specific to the BDV Badge program, we will produce two lists of requirements for the badge issuing platform. Then we will evaluate all of the certified v2.0 badge issuing platforms to select the one which we will use for the BDV badges.

6.5.4.1 A Quick Introduction to v2.0 of the Open Badge Standard

The most recent version of the technical specifications for Open Badges (v2.0) was published on 12 April 2018 (https://www.imsglobal.org/sites/default/files/Badges/OBv2p0Final/index.html) This section will provide a brief overview of that specification.

An Open Badge must contain three pieces of linked meta-data in JSON-LD:

- Issuer Profile – This resource describes who issued the badge. Usually, one profile is created for each organisation, but it is possible to have multiple issuers (for example, different departments within the same university). (The creator of a badge can grant others the right to issue the badge, see The Original Creator Open Badge Community Extension section below.)
### BadgeClass

This resource contains information regarding the badge itself and must include things like the issuer, a description of the badge, and the criteria used to issue the badge.

### Assertion

This represents one particular badge (a BadgeClass) issued to one particular person. People can be identified in a number of ways (Telephone number, URL), but many badge platforms only accept email identifiers.

#### 6.5.4.1.1 A Simple Badge Example

The simplest way of thinking about this meta-data (Issuer Profile, BadgeClass, and Assertion) is in the form of three separate JSON files, each published at a stable URL. A minimal example is as follows (this example contains only the properties required to issue a valid Open Badge and is taken partially from [https://github.com/mozilla/openbadges-backpack/wiki/New-Issuers:-Give-Yourself-a-Badge](https://github.com/mozilla/openbadges-backpack/wiki/New-Issuers:-Give-Yourself-a-Badge)):

**Figure 36: A simple badge example**

In that example, links, shown as magenta arrows, are provided from the award (Assertion) to the badge class (the URL of the badge class is referenced in the badge property of the award) and from the badge class to the issuer (the URL of the issuer is referenced in the issuer property of the badge class). This allows many individually awarded badges to reference the same badge class and many different badge classes...
to reference the same issuer. Also, links are made to other resources such as the image of the badge and a web page, which describes the criteria for issuing the badge.

6.5.4.1.2 Embedding the BadgeClass and Issuer into the Assertion

Rather than publishing the information in three separate JSON files, it is possible to combine or embed the linked data into the badge assertion. An example is taken from the v2.0 of the specification (with some simplifications):
```json
{
  "@context": "https://w3id.org/openbadges/v2",
  "id": "https://example.org/assertions/123",
  "type": "Assertion",
  "recipient": {
    "type": "email",
    "identity": "alice@example.org"
  },
  "issuedOn": "2016-12-31T23:59:59+00:00",
  "verification": {
    "type": "hosted"
  },
  "badge": {
    "type": "BadgeClass",
    "id": "https://example.org/badges/5",
    "name": "3-D Printmaster",
    "description": "This badge is awarded for passing the 3-D printing knowledge and safety test.,
    "image": "https://example.org/badges/5/image",
    "criteria": {
      "narrative": "Students are tested on knowledge and safety, both through a paper test and a supervised performance evaluation on live equipment"
    },
    "issuer": {
      "id": "https://example.org/issuer",
      "type": "Profile",
      "name": "Example Maker Society",
      "url": "https://example.org",
      "email": "contact@example.org",
      "verification": {
        "allowedOrigins": "example.org"
      }
    }
  }
}
```
That last example still includes some external resources, like the image and the criteria, which can also be embedded into the assertion if desired.

A badge hosting platform does not need to accept the embedded BadgeClass and Issuer; in fact, it is likely that it will discard this information and choose to get ‘reliable’ copies of that JSON by downloading them using the URL specified in the respective id properties.

6.5.4.1.3 Signed Badges

Both of the previous examples included hosted assertions. That means that in order to check the validity of the instance of the badge, the verifier needs to contact the URL of the assertion and check that the contents of the hosted URL agree with the given assertion. Another possibility of verifying the contents of a badge is to rely on a cryptographically produced signature of the assertion. In that case, the issuing organisation needs to publish a public key at a stable URL. Then, as part of the verification process, that public key is used to verify the signature of the assertion or instance of the badge. A JSON Web Signature (JWS) is used to cryptographically sign the badge. A JWS contains: a header, a claim, and a signature, all encoded in base 64. Here is a decoded example:

Header:

```
{
  "alg": "RS265"
}
```

Claim:
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

```
{
    "@context": "https://w3id.org/openbadges/v2",
    "id": "urn:uuid:a953081a-4bbd-4927-9653-7219bca00e3b",
    "type": "Assertion",
    "recipient": {
        "type": "email",
        "hashed": true,
        "salt": "deadsea",
        "identity": "sha256$c7ef86405ba71b85acd8e2e95166c4b111448089f2e1599f42fe1bba46e865c5"
    },
    "issuedOn": "2016-12-31T23:59:59+00:00",
    "verification": {
        "type": "SignedBadge",
        "creator": "https://example.org/publicKey.json"
    },
    "badge": {
        "type": "BadgeClass",
        "id": "https://example.org/badges/5",
        "name": "3-D Printmaster",
        "description": "This badge is awarded for passing the 3-D printing knowledge and safety test."
    },
    "issuer": {
        "id": "https://example.org/issuer",
        "type": "Profile",
        "name": "Example Maker Society",
        "url": "https://example.org",
        "email": "contact@example.org",
        "publicKey": "https://example.org/publicKey.json"
    }
}
```
Some things to note from the previous example:

- A hashed recipient was used to prevent the leaking of the recipient’s email address. Here in the (possibly public) JSON file, we have three pieces of information: the hash algorithm used (in this case sha256), the salt string (basically just a random string), and finally, the hash produced from an email with that algorithm and salt. Then without needing to include the user’s email in the assertion itself, a badge hosting service can verify that the badge was issued to a certain user’s email address by generating the hash of that email using the specified algorithm and salt and then checking to see if the newly generated hash is equal to the hash in the assertion. Salt is used to prevent the usefulness pre-calculating of a large number of hashes from a dictionary of emails. It is important to remember that additional verification is also required: that the email address really belongs to the person in question, but that verification is outside the scope of the badge system. (Normally, the badge hosting service checks that the email belongs to a user by asking for a token sent to that email.)

- The assertion is not hosted at a stable URL; thus, rather than using an IRI/URI (Internationalised Resource Identifier/Uniform Resource Identifier), it uses a UUID (universally unique identifier) as its id property. The UUID is generated locally and independently (it should be unique!) for each such assertion. When correctly generated, a UUID has a strong probability of being globally (worldwide) unique (see https://en.wikipedia.org/wiki/Universally_unique_identifier).

- “In the case of signed-verification Assertions, an embedded BadgeClass or issuer Profile can be interpreted to be the value claimed at the time of issue, though publicKeys referenced in an embedded issuer Profile should not be trusted to belong to the issuer without checking the hosted Profile.” (http://www.imsglobal.org/sites/default/files/Badges/OBv2p0Final/examples/index.html)
6.5.4.1.4 Revoking Badges
Revoking a hosted assertion is simply done by deleting the hosted assertion file. Revoking signed assertions is also possible by the issuer publishing a list of revoked UUID’s.

6.5.4.1.5 The Original Creator Open Badge Community Extension
One organisation can create a badge and then give permission to others to issue it. In this case, the original organisation can be identified as the OriginalCreator in the Badge class, and then each issuing organisation can use its own issuer profile.

```
{
  "extensions:OriginalCreator": {
    "@context": "https://openbadgespec.org/extensions/originalCreatorExtension/context.json",
    "type": ["Extension", "extensions:OriginalCreator"],
    "url": "https://example.org/creator-organisation.json"
  }
}
```

(Example is taken from http://www.imsglobal.org/sites/default/files/Badges/OBv2p0Final/extensions/index.html#OriginalCreator)

6.5.4.1.6 Baking Badges
Once all the meta-data required for a badge has been prepared and published (or signed), the meta-data can be baked into a PNG or SVG file to make a badge. Baking a badge is the process of taking a PNG or SVG file and adding the JSON of the badge assertion or a URL of the badge assertion into the image. If this is done properly, the baked badge is still a valid PNG or SVG file and now contains all the information required to verify the validity of the badge (with the exception of the issuer’s profile and public key if it is signed). A baked badge is highly portable; for example, you can email it as an attachment, put it on a web page or upload the file to a badge backpack.

6.5.4.2 Recommendations for the production of BDV Badges
As described earlier, there are two main options for hosting the BDV Badges:

- Hosting the badges at a stable URL (a web page or block-chain for example)
  - Advantages
### Advantages
- The badge instance is permanently published on the web and thus cannot be lost or corrupted. Thus, maintaining the information needed to verify a badge does not depend on the user.
- Simple to maintain statistics regarding issued badges.

### Disadvantages
- Long term storage on the web is not free and puts a financial burden on the issuing organisation(s). This burden is directly proportional to the number of badges issued. There are professional hosting services that claim permanent hosting for one-time fees, but obviously that guarantee depends upon the continued existence of the hosting service.
- There could be legal issues related to publishing personal information. Also, facilities to allow users to delete or rectify their data need to be provided.

### Producing signed badges

#### Advantages
- Long term hosting of the issuer profile and a public key is required, but that is a very small amount of data.

#### Disadvantages
- Slightly more complicated to produce the badges.
- Separate bookkeeping is required to maintain a history of issued badges.
- The user can lose or corrupt the badge file.

As neither of these two options has a clear advantage, we will consider both possibilities when evaluating badge issuing platforms.

#### 6.5.4.3 Description of how the badge issuing platform will be used

In order to better understand the role that the badge issuing platform will play we will briefly summarise the entire process.

A committee of representatives from both academia and industry will formulate a common set of badges and their requirements (BDV badges).

An educational institution can apply to issue certain BDV badges (this application process will be managed outside the badge issuing platform). Once approved to issue a particular BDV badge, the institution will be able to independently issue badges to its students (no intervention on the part of the committee is required). Approved institutions are not allowed to change the BDV badges, nor can they design new badges with BDV branding. Furthermore, they can only issue the badges for which
they have been approved. Each institution only can access information regarding that institution’s issued badges.

Students at an approved institution can apply for a badge once they have completed its requirements. Students submit personal data and evidence files that are stored in the hosting platform. When applying for a badge, the student agrees to the legal requirements related to the use of their personal data (including usage data, see below) and later through the same platform they can exercise their right to have that information modified or deleted (as per the General DP and Regulation (GDPR)). Students can store their badges in any compatible platform and share them on the internet and in social media.

The committee will be able to view statistics regarding all issued BDV badges (but not individual personal data).

The committee can introduce changes in the BDV badges, which are then applied to all institutions who issue those badges in the system in the form of new versions of existing badges. That committee can also decide to discontinue a particular badge or introduce new badges. Periodic reviews of issuing institutions will be conducted (again, these reviews will be conducted outside the badge issuing platform), and the committee’s approval for an education institution to issue badges can be revoked.

6.5.4.4 The badge issuing platform’s functional requirements

Functional requirements:

▪ FR1 – Badges are designed by a central body and then shared only with approved issuers (who cannot modify the shared badges).
▪ FR2 – Approved institutions can issue instances of the badges without the intervention of the central body.
▪ FR3 – Approved institutions can only access the personal information of their own students.
▪ FR4 – The hosting platform offers a badge application process for students at an approved institution. This application includes the upload of evidence.
▪ FR5 – The hosting platform provides space to store student evidence files.
▪ FR6 – The central body can view statistics regarding all issued BDV badges (but not individual personal data).
▪ FR7 – The central body can modify or revoke existing badges as well as design new badges.
▪ FR8 – The central body can revoke permission to issue a badge.

6.5.4.5 Basic technical requirements for the badge issuing platform

Platform requirements:

▪ TR1 - Open Badge 2.0 certified
  o https://www.imsglobal.org/cc/statuschart/openbadges
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

- https://www.imsglobal.org/sites/default/files/Badges/OBv2p0Final/cert/index.html
- TR2 - Meets requirements of GDPR
- TR3 - Company and/or Infrastructure located in Europe
- TR4 - Perpetual badge hosting agreement (badges are not deleted if issuing contract is discontinued)
- TR5 – Issued badges hashed emails and are signed by each approved institution.

6.5.4.6 A review of all Open Badge v2.0 certified badge issuing platforms
This list includes all certified services, taken from https://www.imsglobal.org/cc/statuschart/openbadges on February 1, 2019:

1. Accredible (San Francisco, CA, US)
   - Name: Accredible v1.1
   - Web: https://www.accredible.com/digital-badges/?utm_source=ims
   - GDPR: Yes https://www.accredible.com/accredible-security/
   - Perpetual hosting? Yes https://www.accredible.com/pricing/
   - Costs (per year):
     - $960 - 1000 recipients/year (multiple badges issued to the same email count as one recipient)
     - $1920 – 2000 recipients/year
     - $3840 – 4000 recipients/year
     - $7680 – 8000 recipients/year
     - Custom plans for 10,000+ recipients

2. Concentric Sky (Oregon, US)
   - Name: Badgr v3.1
   - Web: https://info.badgr.io
   - GDPR: “We self-certify with the U.S. Department of Commerce that Badgr complies with EU-US Privacy Shield and Swiss-US Privacy Shield in relation to the GDPR framework as set forth by the EU regarding the collection, use, and retention of personal data from Registered Users based in the European Economic Area; please refer to our Privacy Shield Statement.”
   - Costs: Free
   - Other: Code is open source

3. Italian Quality Company IQC Srl (Bologna, Italy)
   - Name: C-box v1.1
   - Web: http://iqcbox.com/
http://www.itaqua.it/
  o Other: Badges which its users provide: https://www.iqcbox.com/BadgeCCity.aspx

Web only provides information in Italian.

4. Credly (New York, NY)
  o Name: Credly v2018-02-20
  o Web: https://credly.com
  o Infrastructure: Amazon Web Services
  o GDPR: https://info.credly.com/data-security-privacy?hsCtaTracking=72fc39b2-b9cb-4384-a449-2b1898b8958%7C51050119-8f62-44e2-90f1-7c77508d236e
  "Credly complies with the GDPR by maintaining the ongoing confidentiality, integrity, availability, and resilience of our systems that process personal data of badge earners; by ensuring we can restore data in a timely manner in the event of a physical or technical incident; and by regularly testing, assessing and evaluating the effectiveness of our technical and organisational measures."

  “Credly is certified to the US-EU privacy shield.”

5. Pearson Education (bought by Credly in 2018 so now New York, NY)
  o Name: Acclaim v20180221
  o Web: https://www.youracclaim.com/
  o GDPR: https://www.youracclaim.com/privacy
  "If you reside in the UK or a country within the EU, or your Personal Information is processed in the UK or a country in the EU, Credly will, and will ensure that its data processors will process your Personal Information in accordance with all applicable DP legislation, including the GDPR.”

6. Digitalme (Leeds, UK)
  o Name: Digitalme Credly v2018-02-19
  o Web: https://credly.digitalme.co.uk/
  o GDPR: “Credly has built a dedicated European-based data centre, which extends its secure credential cloud infrastructure to support Digitalme operations and help organisations based in Europe comply with strict EU DP regulations.”

A second platform provided by Digitalme
  o Name: Open Badge Academy v2018-02-19
  o Web: https://www.openbadgeacademy.com/
Closing August 31, 2019
7. iQualify
   - Name: iQualify LMS v2018-02-21
   - Web: https://www.iqualify.com/get-started
   - GDPR: “Choose between our global zone or a number of regional zones for storage of your learners' personally identifiable information. We'll help you ensure compliance with your compliance requirements for storing personal data.”
   - Costs (per year):
     - 1800 € Starter
     - 7200 € Partner

This system is primarily an online learning platform through which an institution can issue badges
8. Learning Machine
   - Name: Learning Machine Issuer App v1.0
   - Web: https://www.learningmachine.com/product
   - Infrastructure: Amazon web services
   - Other: Delegation is permitted for Federated Tier customers

This system is a generic credential issuing system which seems to be more geared towards diplomas.
9. Engage Learning (Washington DC, USA)
   - Name: Learning Objects v2018-02
   - Web: https://learningobjects.com/#/what-we-do
   - Infrastructure: Amazon web services

This is an educational consultant firm. They evaluate and suggest improvements to curricula. They offer badges as part of their services to an institution.
10. Collective Shift/LRNG (Chicago IL, merged with Southern New Hampshire University (SNHU))
    - Name: LRNG v1.33
    - Web: www.lrng.org
    - GDPR: NO “By using our sites or giving us your personal information, you are directly transferring your information to us in the United States. The United States may not have the same level of DP as your jurisdiction.” https://www.lrng.org/privacy

LRNG is a learning platform that focuses on youth programmes in different cities in the USA. Participation is by invitation only.
11. Campus Labs (Buffalo NY and Ontario, Canada)
    - Name: MyMantl v1.0
    - Web: https://mymantl.com
    - GDPR: NO “You are responsible for complying with any local laws in its jurisdiction which might impact its right to import, export or use the
Service, and you represent that it has complied with any regulations or registration procedures required by applicable laws to make this license enforceable.” https://mymantl.com/terms/view/site-terms/

12. Discendum (Finland)
   o Name: Open Badge Factory v2018.08
   o Web: https://openbadgefactory.com/
   o Infrastructure: hosted in Finland
   o GDPR:
   o Costs (per year):
     ▪ Free – Max 2 badges, 5000 issued/year, 1 user account
     ▪ 200 € - Max 10 badges, 5000 issued/year, unlimited user accounts
     ▪ 680 € - unlimited badges, 15000 issues/year, sub-organisations (200 € each)
     ▪ 1260 € - unlimited badges, 50000 issues/year, sub-organisations (200 € each), restrict right to issue badges or to review applications
   o Other: reference from NUI Galway
     https://openbadgefactory.com/references/?open=modal-nui
   o Other: provides facilities for a student to apply for a badge and for that application to be reviewed by the organisation.
   o Other: in addition to sub-organisations badges can also be shared with partner organisations.
   o Other: an evidence file of up to 10GB can be stored with each issued badge.

A second service also from Discendum:
   o Name: Open Badge Passport v2.6.0
   o Web: https://openbadgepassport.com/
   o Infrastructure: hosted in Finland.
   o GDPR:
     https://openbadgepassport.com/faq/
     As a European service provider, we will fully comply with GDPR before the transition time ends on May 25, 2018.
   o Costs: Free
   o Other: **Only for hosting badges, not for Issuing**
   o Other: Company details
     https://openbadgepassport.com/faq/
     “Discendum, a successful Finnish educational technology company

2014 by Tekes, the Finnish Funding Agency for Innovation. The MacArthur Foundation funded it further in 2015 as part of the Digital
Media and Learning Competition’s “Trust Challenge,” to foster trust in online learning environments.

- Other: Code
  https://openbadgepassport.com/faq/
  open source (core will be, in progress)

13. MyKnowledgeMap Limited (York, UK)
   - Name: Openbadges.me v1.0
   - Web: https://www.openbadges.me/openbadges
   - GDPR: Yes “Openbadges.me is compliant with GDPR, and we fully respect the privacy of your data and maintaining strong DP standards.”
     https://openbadges.zendesk.com/hc/en-us/articles/360003791174-Data-Protection-including-GDPR-
   - Infrastructure: Microsoft Windows Azure servers
   - Costs (per year):
     - Free – 200 recipients or 1000 badges
     - €400 – 500 recipients or 5000 badges
     - €1200 – 5000 recipients or 50000 badges
   - Other: can create multiple organisations and control user access

6.5.4.7 The evaluation of the platforms

All the platforms considered in Section 6.5.4.6 are Open Badge 2.0 certified (TR1), so we begin by eliminating any who do not openly claim to meet the requirements of the GDPR or are not located in Europe. The following services meet these requirements:

- Italian Quality Company IQC Srl (but they only provide information in Italian)
- Digitalme Credly
- Digitalme (closing August 31, 2019)
- iQualify (mainly an online learning platform)
- Discendum
- MyKnowledgeMap Limited

The first, third, and fourth options in the list are eliminated for the reasons mentioned in the parenthesis. Due to concerns regarding the exit of the UK from the EU, Discendum’s Open Badge Factory (https://openbadgefactory.com/) was considered to be the best available option.

When considering the use of the Open Badge Factory in the context of the functional (FR1-FR8) and basic technical (TR1-TR5) requirements, the following issues arise:

- Should approved institutions be classified as “sub-organisations” or ‘Partners” in the Open Badge Factory?
  - Using sub-organisations:
    Advantages
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

- One contract with an additional 200 € annual fee for each sub-organisation.
- The central body can prevent sub-organisations from issuing badges other than their approved BDV badges with the platform.

Disadvantages
- Not possible to limit access to personal information to only sub-organisations. All personal data is visible to the BDV and all sub-organisations.

Using partners:

Advantages
- Each partner can issue 5,000 badges per year.
- Each partner can issue non BDV badges if they wish.
- The BDV can see statistics regarding issued badges of the approved institutions without having access to personal data.

Disadvantages
- Separate contract agreements for each partner, but also a 200 € annual fee for each.

Conclusion – to avoid DP issues, partners will be used.

- TR5 - Issued badges have hashed emails and are signed by each approved institution.
  The open badge factory does use hashed emails, but they do not sign their badges. However, as each approved institution will have its own issuer profile, the issuer can be verified online.

6.5.4.8 Instructions for Using the Badge Issuing Platform

Instructions for using the Open Badge Factory were written, distributed to institutions granted permission to issue the badges. Those instructions are included in Appendix M.

6.5.5 First Open Call for Issuing the Academic Level of the Data Science Analytics Badge

In October 2019, the first Open Call for Institutions Interested in Issuing the Data Science Analytics Badge at Academic Level was published. The information about the call was provided in the BDVe web site along with the material described in 6.5.3.3 regarding the process description, application forms, FAQ, etc.
As with the pilot, during the BDVe project, institutions applying to the different calls to issue badges will be evaluated at no cost, and if their evaluations are accepted, they will be able to provide up to 5000 badges to their students each year for the duration of their acceptance.

Some of the dissemination activities to publicise the first open call are described below:

- Notifications sent to universities participating in the EIT Digital Data Science MSc
- Notifications sent to selected universities in the BDVe Education-Hub
- Distribution of the call to YERUN (Young European Universities)
- An announcement in the BDVA Newsletter, Twitter, and LinkedIn.
- Webinars targeted on universities in the BDVA to raise awareness regarding the program and to encourage them to participate (November and December 2019)

### 6.5.6 Meetings with External Stakeholders

The primary meetings are described in 3.3, below are described some additional meetings that were held with stakeholders to get feedback about the badge framework and/or to disseminate the work done.

#### 6.5.6.1 Virtual Meeting with Warsaw University of Technology

**When:** 15/11/2018

**Summary:**

- The UPM explained in detail the logistics of the proposal. Warsaw University’s initial feedback was very positive. Warsaw University also said that it was likely that they would participate in the pilot.

- They liked the fact that it could be possible that students who had not finished their final projects could receive badges. They said that many students begin to work and then find it difficult to finish their final projects, so this option would be interesting for them.

- The additional workload required to issue the badges was also discussed, as well as the trade-offs between information and submission effort.

- Warsaw University asked about including their MSc’s program in the BDVe Hub and said that they would submit information using the online form.

#### 6.5.6.2 Virtual Meeting with the University of Adger, Norway (18/01/2019)

**When:** 18/01/2019
Summary:

- The UPM presented the current state of the badge (v.02) proposal to get feedback and to encourage the university to participate in the pilot.
- The University of Adger showed a lot of interest in the process, quality issues, and benefits for industry and students.
- The UPM agreed to send them more information and the deliverable 6.4, so they could have a broader view of the work.
- The UPM also introduced them to the BDVA and encouraged them to contact the association to explore possible benefits for their institution.

6.5.6.3 BDVA Tasks force on Skills
When: 22/01/2019
Location: Brussels
Summary:

- The UPM presented to the other TF leaders the state of the data science recognition framework to get feedback and comments.
- How to promote and get more involvement of the BDVA members was also discussed.

6.5.6.4 Workshop SME Alliance: Boosting Security, Data and IoT Skills – Workshop #3
When: 08/02/2019
Location: Brussels
Organised by: European DIGITAL SME Alliance with funding: Supporting Specialised Skills Development: Big Data, Internet of Things and Cybersecurity for SMEs (EASME/COSME/2017/007)
Web: https://www.digitalsme.eu/skills-home/
Summary:

- The UPM distributed both the badge and education hub flyers to all participants of the workshop.
- During break times, and in individual discussions, a general summary of the badge project and other WP4 work was discussed with different participants.
The UPM gathered information about the Swedish Knowledge Foundation, which offered software-related skills courses, including two online courses on data science.

6.5.6.5 Virtual Meeting with Politecnico di Milano
When: 21/02/2019
Summary:
- PoliMi gave an overview of their study/report on the data science skills needed in the manufacturing domain.
- The UPM presented the current state of the badge program to get feedback and study the possibility that their university participates in the pilot. However, PoliMi was specifically interested in manufacturing. After discussing the BDVe’s proposal, both institutions agreed that the scope of the badge program was too broad to be included in their report on manufacturing.

6.5.6.6 Virtual meeting with Young Universities Research Network (YERUN)
When: 09/07/2019
Summary
- The badge system was presented and discussed. The aim was to raise awareness regarding the badge program among university chancellors in YERUN.
- The Secretary-General of YERUN proposed the possibility of organising a formal presentation of our badge program in a future YERUN meeting so more specific and concrete collaborations can be defined.

6.6 Endorsements for Non-Formal Training in Data Science

Although different definitions for non-formal education can be found in the literature, one of the most relevant is[^2]:

[^2]: COUNCIL RECOMMENDATION of 20 December 2012 on the validation of non-formal and informal learning (2012/C 398/01)
structured on-line learning (e.g. by making use of open educational resources), and courses organised by civil society organisations for their members, their target group or the general public.

In recent years the offer of non-formal training in data science in the form of online courses, MOOCs, in-company training, etc., offered both by official academic institutions and other non-academic institutions, has greatly increased.

Both formal and non-formal learning can contribute to the lifelong learning of data scientists – the continual training of data scientists throughout their careers. While formal systems are often focused on initial training, a lifelong learning system must include a variety of formal and non-formal learning together. This is necessary to meet the individual's need for continuous and varied renewal of knowledge and industry's need for a constantly changing array of knowledge and competences.

This need of knowledge, and therefore of well-trained professionals, is particularly important in the data science domain:

- The past three years have seen “massive growth — 15 times, 20 times growth” in data science-based jobs in sectors like education, marketing, and manufacturing (Wharton, What’s Driving the Demand for Data Scientists?, May 2019)
- Companies expect to triple the size of their big data staff in the next three years (ESADE, Adoption and Impact of Big Data And Advances Analytics in Spain, May 2018)

Therefore, in data science, non-formal education plays a crucial role, and complements formal training, by allowing practitioners to gain new skills and to adapt to new data science requirements.

Here we will consider non-formal education to include any organised training activity outside of formal education (undergraduate or graduate university degrees). Non-formal training includes both e-learning and traditional classroom courses. These courses can be of widely different durations and include training provided by employers, traditional educational institutions, and other third parties.

6.6.1 Challenges when Considering Non-Formal Education

Though the needs of stakeholders in the data science ecosystem when considering non-formal education are similar to those of formal education, there are a few issues which we would like to highlight:

- Students interested in data science training can quickly find a huge variety of options and therefore face difficulties when trying to pick from this overwhelming supply. Which courses are more highly valued by industry, or
what is the right course for their experience and expectations are just a couple of issues that arise.

- Employers that need to evaluate non-formal training also face the problem of how to compare the wide variety of different types of courses. For example, how rigorous are the different programmes in terms of duration, quality, evaluation of the students, identity verification during assessment activities, ...
- Educators offering those courses also face difficulties related, for example, with how to stand out from other courses. That is, how to clearly communicate their offer, attract students, ensure the quality of their training, ...

In other contexts, standardised labelling systems are used to systematically provide information to help to characterise and compare different products in the same category. For example, Figure 37 shows the UK guidelines for Front of Pack Labels, which could be used to, for example, compare different kinds of breakfast cereal.

![Front of Pack Labels](https://www.gov.uk/government/publications)

Figure 37: Example of the Application of the UK guidelines for Front of Pack Labels

With this idea of a standardised labelling system as an inspiration, we have been working on an Endorsement System for recognising non-formal training in data science.

**6.6.2 Timeline for the Implementation of an Endorsement System for Non-formal Education in Data Science**

Below we summarise the different activities required to propose and test the endorsement system for non-formal education in data science.

---

3 Image taken from Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets by the UK Department of Health (https://www.gov.uk/government/publications).
### Activity Description

<table>
<thead>
<tr>
<th>Activity Description</th>
<th>Start Date</th>
<th>End Date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>4.3.2</strong> A Recognition System for Non-Formal Education in Data Science</td>
<td>06/2019</td>
<td>12/2020</td>
</tr>
<tr>
<td>4.3.2.1. Prepare a first version of the logistics and criteria for endorsement of non-formal training in data science</td>
<td>06/2019</td>
<td>01/2020</td>
</tr>
<tr>
<td>4.3.2.2. Evaluation and revision of the endorsement system for non-formal training in data science</td>
<td>02/2020</td>
<td>05/2020</td>
</tr>
<tr>
<td>4.3.2.3. The 1st call for the recognition of non-formal training in data science</td>
<td>6/2020</td>
<td>12/2020</td>
</tr>
</tbody>
</table>

**Table 23: Task 4.3. Timeline for activities related to Non-Formal education**

### 6.6.3 A Starting Point for the Endorsement System

Educational endorsement initiatives have the aim of guaranteeing quality standards and criteria. The European Association for Quality Assurance in Higher Education[^4] has developed an extensive guide for quality assurance of e-learning education. Rather than starting from scratch, we would like to take advantage of this existing work as the basis for the program. Even though our definition of non-formal education covers both e-learning and on-site education, the criteria defined in this document do serve as a good starting point. Their quality criteria are summarised in Table 24.

<table>
<thead>
<tr>
<th>Quality Criteria</th>
<th>Description</th>
<th>Example of Indicators</th>
</tr>
</thead>
</table>
| Policy for quality assurance | Institutions should have a policy for quality assurance that is made public and forms part of their strategic management. | - Institutional policies, structures, processes, and resources are in place to guarantee the successful teaching and learning process of students, including those with special educational needs.  
- Electronic security measures are considered by the institution’s policy/code of practice. |

<table>
<thead>
<tr>
<th>Quality Criteria</th>
<th>Description</th>
<th>Example of Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design and approval of programmes</td>
<td>Institutions should have processes for the design and approval of their programmes.</td>
<td>• Curricula design reflects pedagogical practices and innovation, if applicable.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Teaching staff involved in designing/developing/evaluating programmes are familiar with the advantages/disadvantages of using e-learning in particular course contexts.</td>
</tr>
<tr>
<td>Student-centred learning, teaching, and assessment</td>
<td>Institutions should ensure that the programmes are delivered in a way that encourages students to take an active role in creating the learning process, and that the assessment of students reflects this approach.</td>
<td>• Learning materials fit the pedagogical model and facilitate student learning.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• The technical infrastructure is aligned with the teaching methodology, learning activities, and e-assessment methods, and it eases the teaching and learning process.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Students are clearly informed about the e-assessment.</td>
</tr>
<tr>
<td>Student admission, progression, recognition, and certification</td>
<td>Institutions should consistently apply pre-defined and published regulations covering all phases of the student “life cycle”.</td>
<td>• Students/prospective students are informed about requirements concerning equipment, e-learning and digital skills, pre-knowledge and prerequisite subjects, and attendance.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Students are informed about the workload and pedagogical model of the e-learning programme.</td>
</tr>
<tr>
<td>Teaching staff</td>
<td>Institutions should assure themselves of the competence of their teachers.</td>
<td>• The institution has defined the structure, profile, and role of the teaching staff that is aligned with the pedagogical model.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• The teaching staff is coordinated effectively.</td>
</tr>
<tr>
<td>Learning resources and student support</td>
<td>Institutions should have appropriate funding for learning and teaching activities and ensure that</td>
<td>• The institution defines the electronic security measures that guarantee standards of quality and</td>
</tr>
</tbody>
</table>
Institutions applying for endorsement must provide evidence about the fulfilment of the different criteria in the previous table.

At first glance, the complexity of the system does not seem to be appropriate for the wide variety of programmes which we would like to consider. For example, the organiser of a short five-hour MOOC on data visualisation would be overwhelmed by the work required to apply for the endorsement. Thus, our first task was to distil this list of quality indicators to a minimum based upon feedback from stakeholders in the data science ecosystem.

### 6.6.4 A Preliminary Proposal for an Endorsement System for Non-Formal Education in Data Science

After discussing the ideas behind the quality indicators from the European Association for Quality Assurance in Higher Education with a number of stakeholders (see section 6.6.7), we began by establishing three main objectives for the endorsement system:

- **Quality Control**
Establish a common set of minimum requirements that must be met for a program to be endorsed

- Transparency
  - Require the publication of essential data regarding the course in a publicly accessible location

- Ease of comparison
  - Define a common format to simplify side-by-side comparisons of the essential data
  - Define a common set of data science skills to allow the comparison of content.

Based upon those same discussions with stakeholders, we defined the following as a starting point for criteria to characterise non-formal courses in data science:

- The audience of the course – for example, programmers, analysts, managers.
- Level of course – for example, introductory, ... expert
- Content – number of hours dedicated to the different data science skills
- Teaching types – number of hours: in person, online interactive, online video, online other static content
- Type of assessment – no assessment, online without verification, online with verification, in person with verification
- Institutional Data – link to web page
- Instructors Profiles - link to web page

Regarding the content of the courses, we plan to take advantage of the existing work to define the skills for the Data Science Analytics badge, DS1, ... , DS6, as given in Table 19.

6.6.5 The logistics of the Endorsement System

A summary of the endorsement process is shown in Figure 38. Course providers interested in endorsement can apply by submitting the corresponding information regarding the criteria along with evidences. The application will be reviewed by a group of experts, and if approved, the corresponding program will be endorsed. The information about the quality criteria of the corresponding program will be accessible in the credentials provided to students. In the example below of a badge-based credential, that information would be provided in the badge’s metadata.
6.6.6 Work Planned in Period 3

This initial proposal will be evaluated by gathering feedback from a number of online course providers. (The Politecnico de Milano was contacted and expressed interest in collaborating.) The feedback received will be used to improve both the quality requirements as well as the application process. Feedback will also be sought from industry to get their input and improve again both the quality criteria and the process. (Preliminary meetings have already been held with a representative from Nokia to gather feedback.) Then a pilot of the process will be run before an open call for endorsement in non-formal training in data science is published.

6.6.7 Meetings with External Stakeholders

Below we detail the meetings held with stakeholders in order to get feedback regarding the non-formal education endorsement system in data science.

6.6.7.1 EBDVF 2018 – Meeting with Nokia

When: 12/11/2018

Location: Vienna

Summary:

- The goal of the meeting was to receive industrial feedback regarding the comments from the BDVe project reviewers during the 1st Review, which suggested that the data science skills recognition system include online courses in the.
- How to measure of quality for online courses was discussed and how Nokia could use those measures to evaluate and compare different online
training experiences. How to categorise or at least carefully classify the contents of the online training was also discussed.

- A multi-dimensional endorsement program was suggested with more of a focus on logistics and transparency than on content.

### 6.6.7.2 Virtual Meeting with Big Data Lab – Corsi

**When:** 13/06/2019

**Summary:**

- The UPM and the Big Data Lab discussed a possible collaboration to use the new non-formal recognition system in Big Data Lab – Corsi’s training programmes.

- Corsi’s training program ([http://www.bigdata-lab.it/corsi/](http://www.bigdata-lab.it/corsi/)) was described as follows:
  - They provide Big Data training throughout the region of Emilia Romagna.
  - In 2019 they had their first round of training and had more than 400 students.
  - The program is looking for a system like the BDV badges to recognise their training.

- It was agreed that their training could be a great opportunity to pilot the new non-formal recognition system.

### 6.6.7.3 Virtual Meeting with EIT Digital Professional School

**When:** 24/07/2019

**Summary:**

- The UPM explained the existing BDV badge program and sketched plans for the future non-formal recognition system.

- The EIT Digital Professional School is working on a training program entitled “Data Science for Decision Making.” At the moment, it will contain:
  - Nano-MOOC in Coursera ([https://www.coursera.org/](https://www.coursera.org/))
  - MOOC (much longer)
  - Face-To-Face Workshops (2-3 planned for Q3 and Q4 of 2019)

- They have been looking into how to ‘certify’ their students involved in that program and are interested in using the future non-formal learning recognition system, which we are planning. We agreed to study the collaboration in a pilot during the first semester of 2020.
- EIT Digital Professional School Manager discussed their work in cybersecurity. UPM mentioned that our focus is on data science but felt that the framework of the recognition system could be applied to other domains.

6.6.7.4 Meeting with Big Data Lab – Corsi
When: 16/10/2019
Location: Madrid, Spain
Summary:
- Follow up meeting with the representative of the Big Data Lab to identify potential courses that could collaborate with the UPM endorsement program and definition of a preliminary timeline for the collaboration.

6.6.7.5 EBDVF 2019 - Meeting with Nokia
When: 15/10/2019
Location: Helsinki
Summary:
- The main purpose of the meeting was to gather feedback regarding the proposed criteria for the new non-formal endorsement system from the point of view of the industry.
- We reviewed each of the proposed criteria for the endorsement and discussed them in detail.

6.6.7.6 Virtual Meeting with Big Data Lab – Corsi
When: 4/11/2019
Summary:
- The main purpose of the meeting was to present the existing list of criteria for the new non-formal endorsement system and to gather feedback from an online course provider.
- We reviewed each of the criteria and discussed them in detail.

6.6.8 Summary
In period 2, Task 4.3 focused on two main projects: finalise the definition of the Academic Level of the Data Science Analytics Badge and begin work on a second recognition framework for non-formal educational programmes in data science. To complete the definition of the Academic Level of the Data Science Analytics Badge feedback regarding two previous versions of the badge was gathered and analysed and a pilot of the complete application process was conducted. Currently, two institutions can issue the badge and the first open call for applications to issue the badge is currently open. Also, an initial proposal for the non-formal recognition system was presented. In the next period, we plan to continue to support the
application process for the Academic Level of the Data Science Analytics Badge and implement the new recognition system for non-formal learning in data science.
7 Data Scientists Mobility Programme (T4.4-EIT DIGITAL)

The goal of Task 4.4 is to support a Data Scientist Mobility Programme through the following actions:

- synchronise with other European and national mobility programmes and initiatives
- provide a marketplace/portal for internship positions and mobility opportunities targeting Data scientists, students, and professionals
- define and support the implementation of a mobility framework with two objectives:
  - facilitate the mobility of researchers, students, and professionals within the BDV PPP funded projects
  - facilitate the formation of co-located teams from the different organisations executing those projects

During the reporting period (M19-M36), the work has focused on the following two actions:

1. **Define, Design, and Implement the portal** for Big Data internship positions available within the BDV PPP and the ecosystem outside.
2. **Mobility framework concept definition and implementation plan.**

7.1 Define, Design and Implement the Internship portal

In the period M19-M36, we moved from the concept definition to the analysis of the functional specification and implementation of an Internship Portal (Appendix P.A). The mock-ups have been transformed into an operational portal, not yet populated at the moment ([www.big-data-value.eu/internship](http://www.big-data-value.eu/internship)). The Big Data Category uses the Big Data Taxonomy defined in the Edison Project and the BDVA Task Force 9. The portal provides, as agreed for budget reasons, only the supply side, hosting all open internship positions of the different companies involved in BDV PPP projects as well as BDVA members. As already mentioned in the deliverable D4.2 [11], we have decided to implement a simple version of the internship portal and not a full matchmaking portal. In this case, we do not host the demand side (the CV of the students and/or professionals), and so the companies cannot search within a student’s CV dataset. Several selection criteria are defined so that the students can filter by market, country, Big Data skill, etc. (Figure 39). When selecting a specific company,
more detailed information about the internship position and the skill requirements that the applicant should satisfy will be shown. (Figure 40)
In order to validate what was defined at a functional level and raise the awareness about the internship portal and check the willingness to adopt it, we have launched a survey among the BDV PPP partners and the BDVA members under the EC portal ([https://ec.europa.eu/eusurvey/BDVe-Internship_portal_survey/management/edit](https://ec.europa.eu/eusurvey/BDVe-Internship_portal_survey/management/edit)), see Figure 41. Under the following link, it is possible to verify which are, at the current date, the results coming from the survey.

[https://ec.europa.eu/eusurvey/BDVe-Internship_portal_survey/management/results](https://ec.europa.eu/eusurvey/BDVe-Internship_portal_survey/management/results) (Figure 42 shows just a subset of the survey results)
The Big Data Value project (BDVe), an H2020 Coordination and Support Action of the Big Data PPP, is exploring the opportunity of developing an internship portal collecting in one unique portal all open internship positions in Big Data, being different from existing generalist internship portals or mixed internship/job portals.

The purpose of the Internship portal would be to group together all offers coming from the BDV-PPP members.

Having a specialized Big-data internship portal is, in our view, beneficial for both the companies posting the internship positions but also the potential applicants.

This survey is coordinated by EIT Digital. The data collected will not be made public and will be used only for statistical purposes. For more information please contact Angelo Giuliana at Angelo.Giuliana@eitdigital.eu

*The name of the person filling the survey

*Email of the person filling the survey

Figure 41: Internship survey questionnaire
Following the survey results, we have started to contact directly all the BDV PPP partners that have expressed their interest in the BDVe Internship Portal but also external intermediaries (business incubators, innovation clusters, Big Data DIHs, etc.). At the time of reporting, the percentage of the interviews done is about 60% of the number of people that have answered the survey.

The aim is to collect all the open internship positions that deal with requests in the year 2020 in the next months and upload them to the portal beginning 2020. This will then be launched publicly immediately. Additionally, we are analysing the possibility of using the BDVA PPP social channels to repost or re-tweet the BDV PPP companies internship positions on Big Data, amplifying the audience reached branding the initiative as supported and promoted by BDVe. A way to automatically grab the information from the BDV PPP companies’ social channels and re-used in the BDVA PPP social channels will be analysed both in terms of feasibility and impact reached.

One open issue is the sustainability of the internship portal and who will take care of the internship portal. At the time of reporting, the upload of all open internship positions is done by the BDVe consortium; it could be foreseen to update the internship portal in order to provide the updating directly by the companies that want...
to post their internship Big Data position. This will be decided depending on the number of companies that will use the BDVe Internship Portal.

### 7.2 Mobility framework concept definition and implementation plan

In the last report D4.2 [11] Skills Education CoE Report, it was reported the need to do further investigation to identify potential possibilities of interaction between all actors of the BDV PPP, because in several discussions we had with the BDV PPP partners we got different reactions on the proposed concept; a few considered it as an opportunity, while the majority had concerns about IPR issues, budget problems and the risk of delaying their principal activities.

Therefore, we decided to approach the projects at the BDV PPP Riga Summit held in June 2019, providing them during the Technical and Steering Committee a survey to be filled (Appendix O.A). The results were quite good, more than 90% of the projects gave their availability to verify the possibility of making a pilot setting a co-location team exploring some common challenges among the selected projects (see Figure 43). Despite this, some of them had some concerns under which conditions this pilot could be set, and the main doubts concerned the budget to be allocated and who will provide it, make synergies with the other H2020 projects such as the AI4EU, which are the obligations, what is the effort needed, etc.

In order to start the process and identify which are the projects that could be candidates to such mobility framework pilots, we have selected the companies that are present in more than one BDV PPP project. The reason is that being in more than one project, they are more aware which are the possible common challenges on which starting the pilot and setting the co-location team; they could act as internal evangelists.

In Appendix O.B is shown, which are the candidate companies to whom we have addressed our request to start the process: Demokritos Research Institute and ATB. The cells highlighted in the picture show the projects where the companies are involved, the colours represent if the project is a candidate or not depending on their ending date.

Starting from the beginning of 2020, the pilots will be launched (if the projects confirm their willingness), and the results, in terms of impact and challenges solved or addressed, will be measured.
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

BOVie Survey on Mobility Framework - TC at Riga meetup 26/6/20

Question 1: Would you be interested in taking part in a pilot setting up a co-located team between projects representing partners from two or more H2020 work on challenge to be solved?

Question 2: Would you be interested in joining EU Brng Data Mobility Program if available?

<table>
<thead>
<tr>
<th>BOVie-PP projects</th>
<th>Question 1</th>
<th>Question 2</th>
<th>BOVie-PP projects</th>
<th>Question 1</th>
<th>Question 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. DIGIS</td>
<td>Yes</td>
<td>Yes</td>
<td>21. FANDANIMO</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>2. Big Data Geneva</td>
<td>Yes (4)</td>
<td>Yes (5)-No</td>
<td>24. Fastandbrain</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>3. BigDataCreen</td>
<td>Yes</td>
<td>Yes</td>
<td>25. MiData</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>4. BigDataChannel</td>
<td>Yes</td>
<td>Yes</td>
<td>26. Ikarus</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>5. BigDataSync</td>
<td>Yes (2)</td>
<td>Yes (4)</td>
<td>27. LIMK</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>6. BodyPass</td>
<td>Yes</td>
<td>No</td>
<td>28. MyHealthMyData</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>7. Boost4.0</td>
<td>Yes</td>
<td>Yes</td>
<td>29. OBDCNO</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>8. CLASS</td>
<td>Yes</td>
<td>Yes</td>
<td>30. SLIFO</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>9. Cross-CPP</td>
<td>Yes</td>
<td>Yes</td>
<td>31. SODA</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>10. DataBench</td>
<td>Yes (3)</td>
<td>Yes (5)</td>
<td>32. Special</td>
<td>Yes</td>
<td>maybe</td>
</tr>
<tr>
<td>11. DataBlox</td>
<td>Yes (6)</td>
<td>Yes (5)-No</td>
<td>33. TrustButForYou</td>
<td>Yes (3)</td>
<td>Yes (3)</td>
</tr>
<tr>
<td>12. DataPlot</td>
<td>Yes</td>
<td>Yes</td>
<td>34. TrustAndKnow</td>
<td>Yes (3)</td>
<td>Yes (3)</td>
</tr>
<tr>
<td>13. e-Sides</td>
<td>Yes</td>
<td>Yes</td>
<td>35. TEI</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>14. Edutyo</td>
<td>Yes</td>
<td>Yes</td>
<td>36. TYPOON</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>15. eBusinessGraph</td>
<td>Yes (2)</td>
<td>Yes (5)</td>
<td>37. CloudButton</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>16. EUO</td>
<td>Yes</td>
<td>Yes</td>
<td>38. Muskeer</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>17. JX-Shop</td>
<td>Yes (2)</td>
<td>Yes (2)</td>
<td>39. DeepseeData</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>18. Examinge</td>
<td>Yes</td>
<td>Yes</td>
<td>40. INFOIRE</td>
<td>Yes(2)-No</td>
<td>Yes (3)</td>
</tr>
<tr>
<td>19. MoseCry</td>
<td>No</td>
<td>maybe</td>
<td>41. SmartDataLake</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>20. PrestoCloud</td>
<td>Yes</td>
<td>Yes</td>
<td>42. Innovole</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>21. SafeDrive</td>
<td>Yes</td>
<td>Yes</td>
<td>43. ASHEL</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>22. ASK</td>
<td>Yes</td>
<td>Yes</td>
<td>44.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Projects that provide an answer: 44

Question 1: 58 Yes
3 No
Question 2: 52 Yes
4 No
2 maybe

Condition Q1: Clarity on the budget for co-location; Scientific Interest and Potential Impact; Related to synergies in the BOVie SARl and/or Future AI PPP; limited time to dedicate
After internal discussion with relevant group; if cost are covered; Happy to Investigate; Internal check with CEO

Condition Q2: What is requested in order to make a decision; has to target the career of young researchers; For short period of cross project collaboration; More information on the Mobility Program
After internal discussion with the relevant group; needed more information; unsupervised effort

Figure 43: Mobility Framework Survey results
8 Conclusion

This work package tackles the important issue of Big Data Skills and Education within Europe. BDVe recognises that it cannot transform the Skills and Education landscape of Europe alone; thus, a key principle of this WP is a strong emphasis on collaboration with the existing initiatives to build on their progress and work in partnership. WP4 links with National BDV Centres of Excellence to foster collaboration and support the establishment of new centres. The current offerings concerning support for exchange of knowledge on the universities' data scientist and data engineer programmes across the EU Member States are rather limited. WP4 therefore established an EduHub to facilitate knowledge exchange on educational programmes. In addition, WP worked with industry to ensure Data Skills education is meeting industrial needs, providing particularly relevant input for the BDV Badge Programme. WP4 leveraged existing efforts by EIT Digital on its MSc and professional courses and collaborates with the EDSA and EDISON projects for Data Science Curriculum. Finally, a mobility framework and internship portal for data scientists and professionals are being established, leveraging existing European and national mobility programmes and initiatives, as well as internship positions.

This deliverable D4.3 provides the WP4 progress of activities from M19 to M36 of the BDVe project according to the objectives set for this period.

- **T4.1 “Network of National BDV Centres of Excellence”** led by INSIGHT, in the first period, conducted a survey on CoEs and defined a best practice guide based on interviews from different CoEs. In this period, the best practice guide was refined to reflect the increasing importance of Artificial Intelligence as a topic for Big Data CoEs. The guide was extended with additional knowledge and has been successfully tested on GATE, the first Big Data CoE in Bulgaria. Further work focused on a cross-case analysis for the identification of common best practices and critical success factors which have been captured in a white paper. The Network of CoE have collaborated together in a number of areas including site visits for best practice sharing and in the development of a number of funding proposals involving multiple members of the network. Finally, we have developed an Engagement Lifecycle to support a systematic approach to supporting CoEs at different stages of development (pre-start-up, Early Stage, and Mature).

- **T4.2 “Big Data Value Education Hub”** led by UDE, in the first period, created an EduHub with 166 MSc courses in BDV. In this period, the EduHub was enhanced with 181 MSc programmes, 12 PhD programmes, and 62 on-site/online professional training programmes. The EduHub is hosted as part of the BDV PPP web portal.
T4.3 “The Recognition of Skills for Big Data Professionals” led by the UPM, in the first period, conducted a review of existing recognition frameworks for skills in data science and proceeded with Open Badges. After considering different options, the initial version of the types and requirements for the badges was taken from work from the EDISON project. In this period, it conducted a review of the plan for the badges gathering feedback from academia and industry, it designed its first pilot of the Academic level of the Data Science Analytics Badge, and it is currently offering that badge to academia. Also, it began work on a second skills recognition system focusing on non-formal training. Details are reported in D4.6 [12].

T4.4 “Data Scientist Mobility Programme” led by EIT DIGITAL, in the first period, conducted a survey on existing mobility programmes and it underwent an initial design for portal for internship positions and mobility programmes at the industry. After considering different options for the internship portal, it was decided to focus only on the supply side of the open Big Data internship position. In this period, a survey was conducted that led to a high interest in a mobility framework and internship portal, and initial plans were made on the future pilot programme. Additional solutions for the internship portal are investigated using social channels.
9 References


https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5b76527b1&appId=PPGMS


Appendix A: – CoE Deep Dives

In period 2, we focused on a more detailed elaboration of the best practices within the BDAICoE model. This involved identifying specific practices within each part of our framework. This was achieved by performing a detailed analysis of the case studies and interviews with subject experts. The model has been extended with key practices for the core model capabilities (Strategy, Governance, Structure, Funding, People and Culture).

Key practices were also identified for each of the capabilities for the centres:

- Business Development
- Collaboration
- Research Support Services
- Technical Infrastructure
- Intellectual Property (IP) and Data Protection (DP)
- Education and Public Engagement (EPE)
- Policy Outreach
- Technology and Knowledge Transfer
- Performance and Impact Assessment

Core Organisational Model

A. Strategy

The definition of strategy in the context of BDAICoE framework is:

“The strategy represents the means by which a centre of excellence intends to achieve its overall mission and goals.”

Within the studies, examples of good practice are captured in Table 25.

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clearly stated research strategy with goals, objectives, and research areas.</td>
<td>Strategy and objective</td>
</tr>
<tr>
<td>Aligned with National and European research and innovation priorities.</td>
<td>Aligned with national and European priorities</td>
</tr>
<tr>
<td>Seeks to influence national and European research and innovation priorities.</td>
<td>Influence national and European priorities</td>
</tr>
</tbody>
</table>
### Practice

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Senior Management formulates strategy and objectives.</td>
<td>Strategy/objectives formulation</td>
</tr>
<tr>
<td>Widespread consultation for the formulation of strategies including</td>
<td>Widespread consultation</td>
</tr>
<tr>
<td>research topics dialogue with stakeholders in the research ecosystem,</td>
<td></td>
</tr>
<tr>
<td>including industry and funders.</td>
<td></td>
</tr>
<tr>
<td>The centre uses KPIs for measurement of its performances regarding</td>
<td>Strategy / KPIs alignment</td>
</tr>
<tr>
<td>objectives, goals, mission, and vision.</td>
<td></td>
</tr>
<tr>
<td>The centre’s KPIs cover impact areas, including economic, commercialisation and academic.</td>
<td>KPIs aligned with funders’ agenda</td>
</tr>
<tr>
<td>Operationalised plans are aligned with KPIs.</td>
<td>other issues KPIs</td>
</tr>
<tr>
<td>Decision making through consensus: a structured process to enable</td>
<td>Decision Making</td>
</tr>
<tr>
<td>achievement of consensus.</td>
<td></td>
</tr>
</tbody>
</table>

**Table 25: Summary of Strategy Practices**

### B. Governance & Structure

#### I. Governance

The definition of governance in the context of BDAICoE framework is:

“Governance in Centres of Excellence refers to the level of decision-making and operations.”

Within the studies, examples of good practice depend on the type of institutions and the level of decision making. The practices are captured in Table 26.

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keyword</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governance that emphasises accountability and openness.</td>
<td>Accountability /openness.</td>
</tr>
<tr>
<td>The Governance Board chaired by an external representative. The Board</td>
<td>GC’s responsibilities</td>
</tr>
<tr>
<td>meets on a quarterly basis for strategic oversight of all activities.</td>
<td></td>
</tr>
<tr>
<td>Bi-annual reporting on the strategy, operations, financial &amp; risk</td>
<td>Bi-annual reporting arrangement plan</td>
</tr>
<tr>
<td>management, research programme, industry partners, funding, technology</td>
<td></td>
</tr>
<tr>
<td>transfer, and communications.</td>
<td></td>
</tr>
<tr>
<td>Centre has an academic advisory committee and industry advisory</td>
<td>IAC, SAC &amp; IIC support GC</td>
</tr>
<tr>
<td>committee.</td>
<td></td>
</tr>
<tr>
<td>Reporting demonstrates to the stakeholders the fulfilment of</td>
<td>Reporting</td>
</tr>
<tr>
<td>accountability, openness, and transparency. Regular feedback meetings</td>
<td></td>
</tr>
<tr>
<td>are hosted.</td>
<td></td>
</tr>
</tbody>
</table>
The Centre’s director is supported by a Centre Steering Committee composed of: principal Investigators and industry representatives, representatives of the Host institute, and representatives of funders.

The Centre’s Steering Committee actively monitors the KPIs.

The executive management team comprises the centre’s Director, Senior Staff, and Principle Investigators, and operations managers who are responsible for the day-to-day operations of the centre and its research programmes.

The General Assembly (GA) consists of high-level officials of the centre and its centre partners
- The GA makes the final decisions for the centre.
- The members are made up of a representative from each partner of the centre, including academics and industry.

The Strategy Board (SB) draws the strategic plan, which is used in approving projects at the centre. The SB is composed of the centre’s Leadership, including the Operations Manager, Scientific Coordinator, Faculty Research Strategist, Pilot Strategy Coordinator, Intellectual Property Manager, and Education Coordinator.

Table 26: Summary of Governance

II. Structure
The definition of structure in the context of BDAICoE framework is:

“The structure is how a Centre of Excellence is designed (i.e., levels, roles, units, decisions rights, and accountability).”

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>The structure provides for the centralised coordination and integration of the activities (i.e. business development) across sites for consistency and synergy across the centre sites.</td>
<td>Centralisation, coordination, integration</td>
</tr>
<tr>
<td>At each site, the structure facilitates collaboration with the creation of research-specific groups (i.e. Natural Language Processing, Distributed Systems, etc.)</td>
<td>Research groups</td>
</tr>
<tr>
<td>The structure allows the independent organisation of each site as well as enables decentralisation of site governance.</td>
<td>Decentralisation</td>
</tr>
<tr>
<td>The centre’s Director and PIs form a flat team structure.</td>
<td>Structure</td>
</tr>
<tr>
<td>The operations support personnel include:</td>
<td>Roles</td>
</tr>
<tr>
<td>▪ A Centre Director</td>
<td></td>
</tr>
<tr>
<td>▪ Principal Investigators</td>
<td></td>
</tr>
</tbody>
</table>
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

### Practice

<table>
<thead>
<tr>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>An Office Manager</td>
</tr>
<tr>
<td>Collaborators</td>
</tr>
<tr>
<td>Researchers</td>
</tr>
</tbody>
</table>

The team members are flexibly assigned to project teams.

The Operation Committee takes care of the daily operations of the centre from planning to implementation.

The operation committee is composed of the centre’s Leader, Administration Manager, and Education Coordinator.

The centre is structured comprising two categories of teams:

- Research groups focusing on a specific topic with the goal of attaining scientific excellence,
- Industry Focused Groups representing industry-focused researchers interfacing with business partners and producing their needs with the help of the Research group.

### Table 27: Summary of Structure

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>The team members are flexibly assigned to project teams.</td>
<td>Research Team</td>
</tr>
<tr>
<td>The Operation Committee takes care of the daily operations of the centre from planning to implementation.</td>
<td>Operation board’s role</td>
</tr>
<tr>
<td>The operation committee is composed of the centre’s Leader, Administration Manager, and Education Coordinator.</td>
<td>Operation board’s composition</td>
</tr>
<tr>
<td>The centre is structured comprising two categories of teams: Research groups focusing on a specific topic with the goal of attaining scientific excellence, Industry Focused Groups representing industry-focused researchers interfacing with business partners and producing their needs with the help of the Research group.</td>
<td>Mixed Teams</td>
</tr>
</tbody>
</table>

### C. Funding

The definition of funding in the context of BDAICoE framework is:

“Funding refers to the availability, diversity, and sustainability of the monetary support for carrying out research and educational activities in the centre of excellence.”

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>The funding model is for a long-term (5-8 years) cycle and addresses long-term objectives, e.g. Core research services that provide industry partners’ technology and intellectual property needs.</td>
<td>Long-term funding</td>
</tr>
<tr>
<td>Targets a mixed model of diverse funding sources: national, industrial, and European funding.</td>
<td>Funding diversification</td>
</tr>
<tr>
<td>Pursues research commercialisation opportunities, including licensing and spinouts.</td>
<td>Commercialisation opportunities</td>
</tr>
<tr>
<td>Collaborative projects with industry sponsored by the partner.</td>
<td>Collaborative contract</td>
</tr>
<tr>
<td>Funding comes from consultancy services.</td>
<td>Other funds</td>
</tr>
<tr>
<td>Industry membership model.</td>
<td>Other funds</td>
</tr>
<tr>
<td>Directly funded contract research with individual member companies.</td>
<td>Other funds</td>
</tr>
<tr>
<td>Funding through in-kind arrangements (People, Infrastructure, Hardware).</td>
<td>In-kind funding</td>
</tr>
</tbody>
</table>
The funding policy demands the centre to provide a percentage (20-30%) of its funding needs from industry partners.

**Table 28: Summary of Funding**

### D. People and Culture

#### III. People

The definition of people in the context of BDAICoE framework is:

“People are the human capital required to carry out specific tasks towards the goals of the organisation.”

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core HR policies are provided by the host institutes.</td>
<td>People</td>
</tr>
<tr>
<td>Staff diversity – catering for good gender mix among researchers and students and local/international personnel proportions.</td>
<td>People</td>
</tr>
<tr>
<td>Young researchers are given opportunities to work on research and industry projects. They are encouraged to apply for grants and become leaders in project work packages.</td>
<td>People support</td>
</tr>
<tr>
<td>PhD members are offered financial support during maternity leave.</td>
<td>Support</td>
</tr>
<tr>
<td>Attract international researchers and develop the careers of young researchers.</td>
<td>People</td>
</tr>
<tr>
<td>Personal skills development, e.g. public speaking, person-to-person communication, language training, unconscious bias, and cultural awareness.</td>
<td>Training</td>
</tr>
<tr>
<td>Advertisements of vacant positions in both local and international forums.</td>
<td>HR sourcing</td>
</tr>
<tr>
<td>Industry partners have opportunities to place their postgrads on internship programmes with the centre.</td>
<td>People</td>
</tr>
<tr>
<td>The centre is building a common team identity by the creation of an enabling environment using informal meetings such as team lunches meetings, kick-off meetings, and poster sessions.</td>
<td>People</td>
</tr>
<tr>
<td>A mentorship programme is personalised for each intake of students and researchers:</td>
<td>People development</td>
</tr>
<tr>
<td>▪ Mentor-mentee matching is based on their profiles.</td>
<td></td>
</tr>
<tr>
<td>▪ Mentors are industry leaders.</td>
<td></td>
</tr>
<tr>
<td>▪ Mentor/mentee gains from the professional/personal goals of the mentorship programme.</td>
<td></td>
</tr>
<tr>
<td>▪ The programme promotes mutual understanding, collaboration, etc.</td>
<td></td>
</tr>
<tr>
<td>between the centre, partners, and staff.</td>
<td></td>
</tr>
</tbody>
</table>
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>A mentorship programme to support female careers.</td>
<td>People</td>
</tr>
</tbody>
</table>

**Table 29: Summary of People**

**IV. Culture**

The definition of culture in the context of BDAICoE framework is:

“Culture represents the underlying values, beliefs, and norms that drive the teams and the centre of excellence as a whole.”

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distributed geographically centres maintain a culture that enables collaboration on works that mutually benefit all sites.</td>
<td>Collaboration</td>
</tr>
<tr>
<td>Cultural practices promote awareness, regular contact, and teamwork, communication, and dissemination activities (meetings, symposia &amp; centre communication systems, annual intercultural day).</td>
<td>Workforce inclusivity, communication, and dissemination</td>
</tr>
<tr>
<td>Staff and student welfare programmes are employed, i.e. <em>Listening lunches</em>, giving members an opportunity to voice concerns, talks on mental health, and other issues.</td>
<td>Welfare programmes</td>
</tr>
<tr>
<td><strong>An ‘Unconscious Bias‘ training is done in the centre to eliminate preferential treatment.</strong></td>
<td>HR practices</td>
</tr>
<tr>
<td>Encouragement of social events such as cycling or walking tours.</td>
<td>Socials for well-being</td>
</tr>
<tr>
<td>Support for ‘inclusivity‘ in recognition of the benefit of diversity and support gender equality.</td>
<td>Inclusivity &amp; Equality</td>
</tr>
<tr>
<td>The weekly centre meeting is used to promote the spirit of togetherness.</td>
<td>Meetings</td>
</tr>
<tr>
<td>The community volunteering may overlap with Education and Public Engagement.</td>
<td>Volunteering</td>
</tr>
<tr>
<td>A result-orientated environment. People are always conscious of the deadline, and they must work hard to deliver.</td>
<td>Culture</td>
</tr>
<tr>
<td>The new member programme is a cultural practice aimed at providing new in-takes detailed orientation experience and support to adjust to the working environment.</td>
<td>Onboarding culture</td>
</tr>
<tr>
<td>The new members are assigned a ‘buddy‘ (a Post-doc or a PhD student) that offers orientation and mentor support to a new intake.</td>
<td>Onboarding culture</td>
</tr>
<tr>
<td>The centre provides researchers with a mechanism for understanding social forums and building a good rapport with the company representatives.</td>
<td>Mentoring</td>
</tr>
</tbody>
</table>
Informal non-project related meetings are held, and they include kick-off meetings, poster sessions, lunches.

An office space is designed to facilitate cultural evolvement of the people in the centre and make them feel part of the centre.

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informal non-project related meetings are held, and they include kick-off meetings, poster sessions, lunches.</td>
<td>Informal gathering</td>
</tr>
<tr>
<td>An office space is designed to facilitate cultural evolvement of the people in the centre and make them feel part of the centre.</td>
<td>Cultural support</td>
</tr>
</tbody>
</table>

Table 30: Summary of Cultural Practices

Capabilities

A. Business Development

Business Development (BD) is a separate role that is gradually becoming a common role in research centres. The basic duty is to help in the creation of opportunities for the application of the inventions of the centre and to attract collaborators for joint research, particularly with the industry partners as a way to establish a technology transfer from the research institution to the industry for economic exploitation. The expert in charge of BD often possesses a mix of business and negotiation expertise with a good understanding of the research environment and themes. These skills enable the BD officer to play key roles in negotiation and arrangements that lead to the signing of collaborative research contracts.

Results from the BDAICoE case study indicate that collaboration is important and regarded as the lifeline of a research centre in existence. For this reason, BD activities that increase the chances of collaboration are now being considered with a good budget to support the role. In the case study, various arrangements were seen to be in place for BD. Some centres have elaborate BD functions and related roles such as a separate office that has been created for the role of BD with related roles of Strategic Partnership Manager and Applied Innovation Officer. Within smaller to medium-size centres, the BD role is performed mainly by the centre’s Director and/or the centre’s Manager. The research centre funding agencies and other agencies particularly in the policy-making environment, play roles in BD activities to support the research institutions in their mandate. National funding and development agencies can give the opportunity to centres to attend international events and to present its innovative outcomes to industry members.

In many of the centres studied, the BD activity begins with an inquiry by contacting interested partners on projects among the industry stakeholders and showcasing research areas to company representatives who provide feedback for input for planning research direction. BD can raise awareness of the centre’s activities by making presentations at various science and industry conferences where potential users and their solutions are present. Often a centre’s BD process takes the following
steps: understand the Company and introduce centre capability and develop a proposal for decision-making (GO/NO-GO). Then, introduce available funding models and articulate the goals of the project proposal. The final stage is to decide to sign the contract.

**Recommendation for New Centres**

BD is closely linked to collaboration; in fact, the former leads to the latter. Many organisations create BD roles to deal with existing opportunities or create new ones with potential customers. In the research environment, the BD role helps to do the same thing, and important is the fact that the creation of the role enables the core researchers to focus on the research activities while other activities leading to collaboration are left in the hands of another professional. The BD Officer, however, must liaise with the research Unit Leaders to understand their needs and the status and features of their inventions and innovative products to be able to negotiate a contract on their behalf with a potential collaborator. Besides, the main goals of the BD role are usually designed to achieve the strategic goals of the research centre including collaboration. For example, one centre provides a model for the use of the BD functions to improve success because it uses business development activities to achieve the following goals:

- As a single point of contact for the centre; the start of inquiry and interface between the centre and the industry partner
- A mechanism to collect feedback on emerging trends from the industry to the centre.

In addition to the above, another advantage of the BD role is that it is used to seek the right partnership to push the centre to develop enough bandwidth to stimulate the right challenges for technological advancement.

### Practice

<table>
<thead>
<tr>
<th>Practice</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercialisation and licensing aspects are handled by the host institute TTO.</td>
<td></td>
</tr>
<tr>
<td>A Business development capability to engage industry proactively and support principal investigators to pursue research goals (national and international) and support the transfer of research outputs to industry.</td>
<td></td>
</tr>
<tr>
<td>Engages industry stakeholders and researchers working with agencies.</td>
<td></td>
</tr>
<tr>
<td>Works with national investment agencies to support the attraction of investments.</td>
<td></td>
</tr>
<tr>
<td>Organises industrial days to showcase research and innovation outputs in the centre.</td>
<td></td>
</tr>
<tr>
<td>Business Development gathers industrial feedback from companies as input for research direction.</td>
<td></td>
</tr>
<tr>
<td>The Business Developments act as an intermediary between academics and companies while providing a consistent interface of the centre to a company.</td>
<td></td>
</tr>
<tr>
<td>A staged business development process.</td>
<td></td>
</tr>
<tr>
<td>A Strategic Partnership Manager role was created for the overall centre. The Strategic Projects</td>
<td></td>
</tr>
</tbody>
</table>
Practice

Manager’s role is

▪ to initiate and develop engagements with key Multinational Corporations (MNCs) and SMEs, centring on agreed common focus areas,
▪ to identify and finalise with the Industry Partners some goals in connection with focus areas, e.g. the recruitment and development of Human Capital, Connected Health, Discovery Economy, etc.
▪ The Strategic Projects Manager engagement process reflects a multi-year-collaborative research partnership to deliver projects through funding programme.

Builds awareness in the Data Analytics ecosystem through seminars, conferences, consultancy, and members’ networking events.

Leverages local and international media agencies to support the centre in terms of outreach.

Seeks partnership development to stimulate the right conditions for technological advancement.

| Table 31: Summary of Business Development |

B. Collaboration

The definition of collaboration in the context of BDAICOE framework is:

"UIC refers to the formal and informal engagement and interaction between a higher educational institution and an industry partner with the aim to facilitate knowledge and technology exchange as well as to provide ad hoc advice and networking opportunity for the professionals. This can be done through the establishment of activities such as collaborative and contract research and the provision of consulting services."

The collaboration capability is considered a matter of surviving or dying in the research environment. Findings from the case study show that collaboration is given a significant consideration and the practice is common to all the research centres studied in this project. The fundamental aim of the collaboration is to facilitate the leveraging of cross-disciplinary skills from partners in the national and international environment including the European Union. Research centres devote considerable resources to the development of industrial collaborations for network and strategic partnerships. Regarding the scope, the research institution’s management teams strive to ensure a broad scope of the collaborative partner’s domain, sector, and geography. For example, many centres seek a cross-disciplinary collaboration from both local and international environments as well as from the industry and academic sectors. The goals are to arrive at a cross-fertilisation of ideas which is good in the development of state-of-the-art and innovative products and services. Despite the above intention, a centre-specific focus also plays key roles in the kind of the
partnership maintained by a research centre. For example, some centres actively seek collaborative research partners from among other academic and research centres, companies, and other entities (e.g. Municipality government) on purely scientific inventions and EU projects covering many areas such as Manufacturing, Agriculture, Financial Services, Smart Cities, etc. To engage partners easily and sign contracts, all the centres seek a well-developed Business Development role to create opportunities for collaboration partners, particularly with industry participants, because this leads to the technology transfer through the development of spin-outs and IP licences. Within the centres, with an applied focus, this went further, with the centre devoting its resources to the development of demonstrators and services design to solve partners’ problems and through these industry-driven research activities. In these situations, industries could influence the choice of demonstrator projects.

**Recommendation for New Centres**

Whatever drives collaboration in a research centre, the key thing is to ensure cross-fertilisation of ideas from various disciplines, and this has to be aligned with the strategic goals of the institutions. Understanding the industry trends may help direct management attention to the type of collaboration and partnership to seek, on the one hand, while on the other, it will make sense to follow the need of the end-users of the final product or outcomes.

<table>
<thead>
<tr>
<th>Practice</th>
<th>The centre provides Master classes to industrial partners to stimulate collaborative opportunities.</th>
</tr>
</thead>
</table>
| Examples of collaboration benefits include: | ▪ A collaboration capability is leveraged in the strategy implementation for success  
▪ A collaboration capability helps in attracting the industry partners for cash revenues. |
| Demonstrators’ activities allow partners to influence the choice of demonstrator projects. Leverages its core funds for the development of analytics demonstrators. |  |
| The centre collaborates with both tech developer & technology user organisations under two methods: | ▪ Reliance on demands and requirements arising from the partner’s dissatisfaction with the existing solution  
▪ use prototypes to generate feedback from the actual use in order to understand how to upgrade the solution. |
| The centre collaborates with industry partners to identify new & relevant research problems to influence the international research agenda. |  |

**Table 32: Summary of Collaboration**
C. Technical Infrastructure

The definition of infrastructure in the context of BDAICoE framework is:

“Infrastructure is the systems, practices, and tools that facilitate and reinforce the work within the organisation.”

<table>
<thead>
<tr>
<th>Practice</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>On-going development of research infrastructures to ensure the infrastructure and services, facilitating the delivery of world-class research and professional industry collaborations (including Testbeds, Living labs, etc.)</td>
<td></td>
</tr>
<tr>
<td>Experimentation test labs and living labs including a smart building and partnership with cities.</td>
<td></td>
</tr>
<tr>
<td>Internal data centres within the centre.</td>
<td></td>
</tr>
<tr>
<td>Cloud Services are leveraged on an as-needed basis for specific projects.</td>
<td></td>
</tr>
<tr>
<td>The centre’s researchers can work on the premises of partners.</td>
<td></td>
</tr>
<tr>
<td>Dedicated premises for the centre.</td>
<td></td>
</tr>
<tr>
<td>Adopts project-centric infrastructural provision. In some projects, it works on the infrastructure provided by the industrial partner.</td>
<td></td>
</tr>
</tbody>
</table>

Table 33: Summary of Technical Infrastructure

D. Research Support Services

Context and Definition

The University College Cork (UCC) explains the components support activities which together make up the research support services. As the name implies, the UCC’s Research Support Services Office seeks to provide the university community with research-oriented services to aid research activities through the provision of services such as information, assistance, guidance, and advice on all aspects of the planning, education, sustaining and application of research. The Research Information Network (RIN), a UK-based organisation investigates what kinds of information-related services are available to support researchers’ activities through the research lifecycle and how those services are used and valued by the researchers.

---

The RIN found that the four most research-oriented universities in the UK provide information-based support services for researchers across the four stages of the research lifecycle. However, the supports available are mainly focused on the initial (discovery) stage of the cycle, e.g. identifying grant opportunities and then the final (result dissemination) stages of the cycle, e.g. knowledge transfer and commercialisation.

**Good Practices – Research Support Services (RSS)**

With the aim to discover the tools and services researchers make use of in research activities, how these tools and services are effectively supplying the needs of the researchers, RIN (ibid) uncovers a list of tools and services used by researchers:

- An alert system for researchers to know about forthcoming grant opportunities from various bodies
- A system that locates and connects potential collaborators and facilitates a collaborative management of documents, including large text and data analysis to curate and preserve research data
- A system for skills development in information handling
- An effective information dissemination tool – e.g. to publish research results and to provide advice on how to protect IP rights
- A tool for the storage and management of pre-print, publications, and post-print research results and researchers’ institutions investigate their standing within their field, including the management of citations and citation analysis

---

6 The Research Information Network (RIN) named the four most research-intensive universities in the UK as Leicester, University College London (UCL), Warwick and York.
Based on the above list of services utilised by researchers, RIN (ibid) offers four areas that should be covered by good research support services, and these address:

- How to generate and develop new ideas, research proposals, and projects
- How to seek, apply for, secure and manage funding opportunities
- Adequate and appropriate support for actual research activities
- Effective means of research information dissemination, including publishing findings.

**Study Findings**

Research support services are already known to be a necessary input to the success of research activities; hence, research institutions provide a series of support services financed by a reasonable level of budget. For example, larger centres maintained research support services through the following two areas among others:

- **Proposal Support:** Grant Development Managers provide support at all stages of the project lifecycle from the proposal submission, negotiation, implementation, and completion. Broken down, these services include email alerts from the Project Management Office (PMO), providing information on available research funding opportunities, information on the publication of calls for project proposals, etc.

- **Project Management:** A Project Manager, usually the Head of each research unit/group, takes responsibility for each project and works closely with other consortium partners (in case of large projects involving external partners) and the centre’s Principal Investigator. All projects have allocated resources, clear objectives, an agreed approach as well as the defined timelines. The Project Manager plays roles in the allocation of resources to various projects according to the availability and need.

- Research support services are provided primarily by the host institutions, with the centre providing more specific services where needed. The level of local support is dependent on the size of the centre.

<table>
<thead>
<tr>
<th>Practice</th>
<th>Keyword</th>
</tr>
</thead>
<tbody>
<tr>
<td>An alert system for available grant opportunities.</td>
<td>Grant alert</td>
</tr>
<tr>
<td>A tool to locate, connect and collaborate on documents.</td>
<td>Collaboration tool</td>
</tr>
<tr>
<td>Information and skill development tool.</td>
<td>Skill development</td>
</tr>
<tr>
<td>Research management support systems include source control, business process, issues trending, helpdesk, project management, collaboration tools.</td>
<td>Support tools</td>
</tr>
</tbody>
</table>

**Table 34: Good practices in research support services**
**Recommendations for new Centres**

Information gathering and documentation skills are critically important to researchers because these activities are important to the development of ideas and proposal preparation at the initial stage of research work when seeking, handling, and managing information resources of different kinds and from different sources [8]. In order to support researchers on a project information generation, research institutions and universities should provide researchers with tools that fetch material from sources to which they are also provided access (ibid). It is equally important to alert researchers about available funding opportunities and sources as well as provide tools to enable them to collaborate effectively with co-researchers for cross-fertilisation of ideas necessary for the development of innovative products. With adequate support, researchers cannot disseminate their end results to user communities and cannot share with fellow researchers to improve the body of science. Furthermore, researchers need support services to be able to arrange the technology transfer and IP contracts as well as related issues of spin-offs.

**E. Intellectual Property (IP) and Data Protection (DP)**

**Context and Definition**

**IP:** The World IP Organisation (WIPO) defines IP as follows [15]:

"IP refers to the creations of the mind: inventions, literary and artistic works, and symbols, names, and images used in commerce."

The organisation categorises IP into an industrial property, including sub-divisions such as patents, trademarks, industrial designs, and geographical locations, and copyright, which covers literary works, films or movies, music, and artistic works (ibid).

**DP:** DP simply defined is the “legal control over access to and use of data stored in computers” [7] and in organised physical filing systems. This definition is further explained by the DP Act[8] referring to DP “the ways in which information about living people may be legally used and handled. The main intent is to protect individuals against misuse or abuse of information about them”. This kind of protection is also afforded to organisations as legal persons to cover the use, processing, and sharing of their data.

**Good Practices – IP and DP**

---


Recently, the EC issued the GDPR document that came into force on 25th May 2018. This legal document provides for the collection, processing, use, sharing, storage, and retrieval of personal data by any organisation or individual. The practices recommended by this body of laws are enough guides on the ways researchers and research institutions should deal with research data, whether as data input into research processes or as data output from the processes. Thus, with regards to data gathering from literature reviews, surveys, and interviews and the processing of the data and sharing, researchers need to be careful about how they deal with the handling of third-party data. Permission must be obtained from data owners before starting to collect data and processing it. Similarly, appropriate referencing has to be provided in line with research ethics for all third-party materials used for research purposes. In order to meet with some of these obligations, the research institute or university needs to provide some facilitating services – commonly referred to as the research support services in form of relevant training, provision of appropriate training on relevant regulations (e.g. GDPR), tools and platforms to enable researchers comply with applicable legislation such as GDPR with regards to DP and privacy compliance. DP and security needs should also be considered in case of the protection of research output so that not to allow it to fall into wrong hands or misused.

**Study Findings**

Every research centre is very careful about the issue of data and IP protection issues. Evidence suggests that institutions do not want to be caught in the troubles of DP infringement. For example, the research centres’ and universities' main rules about compliance with plagiarism rules, adoption of a specific citation and referencing style, as well as maintenance of offices of DP and IP arrangement and related matters. All the centres have implemented a series of processes and procedures that mandate the comprehensive recording and documentation of research plans, results achieved, data created, data distribution, IP creation and management, Non-Disclosure Agreements (NDAs), IP assignment agreements, publications protocols, invention disclosure protocols, and other aspects of good research governance. Larger centres maintain offices for IP and DP and created the IP and DP Committee respectively to oversee the development, implementation, and update of policies and procedures relating to IP and DP, including Responsible Research and Innovation (RRI). **Data Collection Guidelines** are provided by each of the parent universities and are drawn in compliance with data privacy and security regulation, which has now been updated to follow the recently implemented GDPR.

One of the smaller centres saw an advantage in having an IP protocol that facilitated a smooth and fast process that encourages partners in a collaboration. The process is completely simplified to facilitate the process completion process in one day to attract potential partners to collaboration.

Other centres pursued an IP model with the aim to support open knowledge on which commercial solutions can be built. The guiding principles for managing IP-results from
laboratory activities are automatically defined as laboratory results of which the generic components are funded by the National funding Agency, and they are therefore released using one of two suitable open-source licenses (LGPL or Apache). Using any of these licenses permits free integration, as-is, with proprietary software.

Table 35 shows a summary of IP and DP good practices.

<table>
<thead>
<tr>
<th>Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create awareness about the regulations (e.g. GDPR) and provide training and tools for researchers to facilitate their compliance.</td>
</tr>
<tr>
<td>Simplify IP and DP compliance processes to facilitate the adoption and achievement of goals.</td>
</tr>
<tr>
<td>IP and Data Protection Policies are overseen by a dedicated coordinator.</td>
</tr>
<tr>
<td>Research Integrity is supported by the Host Universities’ policies and is implemented as a package that includes multi-actors and public engagement, enabling easier access to scientific results, the consideration of gender and ethics in innovation concerning content and process, and formal and informal science education.</td>
</tr>
<tr>
<td>A Data Protection Manager role helps strengthen the implementation of data protection policies</td>
</tr>
<tr>
<td>Highly simplify the IP process into a one-page document and same day process.</td>
</tr>
<tr>
<td>An open knowledge intellectual property model supports commercial solutions development.</td>
</tr>
<tr>
<td>Two licences – the LGPL or Apache, and the use of any of these licences permits free integration of ‘as-is’ situation with proprietary software.</td>
</tr>
</tbody>
</table>

Table 35: Good practices in IP and DP

F. Policy Outreach

Context and Definition

These are outreach activities designed to create a community impact sufficient to influence the government’s decision to create, change, or adopt a policy or directive that supports certain practices in the society. EPE practices are primarily designed to take research results to the community and to make them become role players in the use and innovation cycle of scientific inventions. However, it is also aimed at producing a long-term impact for which the government is compelled to use policy enactment for support so that the desired outcome can be achieved and sustained.

Study Findings

Influence National Policy: Through EPE and outreach activities, many centres have succeeded in influencing policymaking. For example, through data collection, analytics and interpretation capabilities deployed to support optimisation of the allocation of healthcare resources, the government has been influenced to make provisions for this need in the national body of policies. Similarly, a policy backing has been created to support the introduction of the computer science subject into the secondary school
curriculum as a result of a centre’s coding programmes designed for young students in secondary schools. Many of the centres play a key role in supporting the national and European policies for Digital Economy, Industry 4.0 and the Data Economy.

**Research Integrity:** Typically, centres uphold national policy demands such as gender balance in workplace and research domain-specific policies such as research and innovation ethics on content and process as well as formal and informal science education. Some centres employ a multi-actor arrangement for public engagement in research and innovation which enables easier access to scientific results. One example of a national policy on Research Integrity is the Irish Universities Association (IUA) National Policy Statement on Ensuring Research Integrity in Ireland [1]. Four commitment areas provided by the National Forum on Research Integrity in Ireland [5] underpin research integrity in the country:

1. **Standard** – a commitment to ensure the highest standards of integrity via enforcement of basic principles of good research practices.
2. **Education** – activities of education and promotion of good research practices.
3. **Collaboration for continuous improvement** – a commitment to teamwork to reinforce, review, and safeguard integrity.
4. **Action to address misconduct** – employment of transparent, fair, and effective processes to deal with research misconduct.

Table 36 provides a summary of the best practices of policy outreach.

<table>
<thead>
<tr>
<th>Practice</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Influences policy areas (i.e. education, health) making at National and European level.</td>
<td></td>
</tr>
<tr>
<td>Influences a research and innovation policy at both national and European levels, including Digital Economy, Digitisation of Industry (Industry 4.0), and the Data Economy.</td>
<td></td>
</tr>
<tr>
<td>The centre is a member of relevant associations for the European level activity (i.e., BDVA, ETP4HPC, AIOTI, etc) H2020.</td>
<td></td>
</tr>
</tbody>
</table>

**Table 36: Good practices in Policy Outreach**

**G. Education and Public Engagement (EPE)**

The definition of EPE in the context of BDAI-CoE framework is:

“Outreach is the collection of information dissemination activities with which a research centre informs the public about the science and technology developments in the centre. The aim is to enable the public to appreciate science and technology.”

Findings from current research reveal that outreach programmes are taken as serious activities in some centres because through these programmes, the research out is publicised to the outside world. It is an important strategy to make the results of research known not just to partners of a research institution,
but also to the funding agencies, an important avenue by the research centre to disseminate information into local and international science and technology research communities as well as to youngsters in colleges and high schools. EPE has been designed by some researcher organisations into a package of events that include voluntary training of high school pupils and community programmes that bring awareness to various demographics levels. Some of the EPE programmes aim to create inclusivity whereby members of the community are motivated to take advantage of technology tools on offer and contribute to the adoption of technology. Other practices aim to create awareness about problems and unwanted outcomes of science so that the use of results of technology tools do not fall prey to the negative impact of science and technology such as cybercriminals. In this context, outreach also discourages science research into technology that supports weapons of mass destruction, environmental degradation, etc., by way of support to give to policy-makers the ability to provide appropriate policy restrictions.

Good Outreach practices should help a research centre to:

- Expose its research results to the wide community like researchers (local and international communities), industry partners and the young people
- Make the centre popular among others
- Should help drive interest in science and technology particularly among young people who should become the following generation of researchers
- Include all demographics in the awareness and use of technology tools
- Bring the outcomes to local communities and help them take advantage of the outcomes of research
- Participate in conferences and workshops organised for researchers including the supply of papers to journals
- Discourage negative science research — research that leads to the development of ‘bad technology’
- Follow the Create, Include and Question approach to outreach practices
- Build outreach activities into KPIs

**Study Findings**

Outreach activities are activities that bring science and innovation development to public awareness. These activities enable researchers to interface directly with science and technology output users and teach them how science is developing innovative products to which they must align with, test, use and adapt in their everyday life. Research centres commonly indulge in presentations of their work in seminars and conferences, the publication of science paper and books. However, some create specific avenues for uniquely disseminating their research outcomes for maximum
impact in the society. The centres typically target their outreach activities to three specific stakeholder groups:

- The academic community
- The industry and
- The national and international environments

It is believed that messages directed at the target audience in each of these communities would be slightly different to achieve the desired goals. One centre developed a structured process for outreach; EPE messages and activities are designed and carried out considering three principles, which include – Create, Include and Question. The create approach holds the view that activities should be designed to get people involved in scientific activities. The inclusive approach means activities should adequately target everyone and involve them in outreach programmes – that is activities should consider reaching out for age groups, gender, educational levels, racial communities, business, and social arenas as well as rural and urban considerations. The question approach enlightens scientists and everyone in the society to query, investigate and interrogate what goes under scientific research and what comes out from scientific research. In other words, is scientific investigation or science research producing useful science and technology for humanity or is it producing dangerous and harmful outcomes for humanity? This is a very laudable principle that is aimed at ensuring science research does not get out of hand into the production of destructive technologies such as technologies dangerous to the environment, support for mass destruction, cyber-attacks, etc. for humanity.

**Recommendation for New Centres**

Good thinking research centres do not only use outreach to educate the public about scientific development, but also they design outreach to popularise their output within the various demographic strands of the society, within the economic, business and educational sectors, social and political environments as well as in the local and international communities where possible. They invest time, money, and expertise in programmes that introduce science awareness and direct leaving certificate students towards the stem subjects to ensure there is a good interest in sciences at college levels. In some cases, the research centre tries to influence national policy-making through recommendations on subjects that should be introduced into the post-secondary school level with the hope that it could drive interest in science and technology subjects. Such initiative helps to put the research centre in the news due to media publicity.

<table>
<thead>
<tr>
<th>Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>A dedicated public outreach team.</td>
</tr>
<tr>
<td>All researchers collaborate with a public outreach expert to bring the EPE programme to target different age groups in public outreach.</td>
</tr>
</tbody>
</table>
The three principles of the public outreach approach:

- **Create** – is an approach to get people involved in scientific activities.
- **Include** – make all sectors of the economy, community, and the society the target for EPE activities.
- **Question** – teaches the idea that science and technology research should be evaluated to understand whether outcomes are beneficial to humanity or not.

Public Outreach performances to be measured using KPI for each researcher or staff member.

To attract collaborators, the centre promotes three key areas: within the academic community, in the industry, and in national and international environments.

Membership in external networks to promote the participation of women in IT, science, and technology.

### Table 37: Summary of Public Outreach

#### H. Technology and Knowledge Transfer

**Context and Definition**

Many definitions of the concept exist without a clear consensus on the nature of the transfer process; thus, the context tends to underpin the process as well as the definition of the term technology and knowledge transfer [7]. An example cited by the above author is that in developed countries, the concept often refers to “the process whereby universities or research centres provide access to technologies created there through a variety of mechanisms for interaction with market operators”. Similarly, the term may also refer to the process by which an across-sectoral or cross-national boundary transfer of technology and know-how could involve commercial and non-commercial activities, movement of technical information, physical material assets, and immaterial elements, persons with specific capabilities [7]. While technology per se has been defined as “the systematic knowledge for the manufacture of a product, for the application of a process or the rendering of a service, including any integrally associated managerial and marketing techniques” [4].

The WIPO⁹ plays key roles in support of a mutually beneficial technology transfer through the arrangement of patent information services, innovation support programmes and tools, projects, and activities. Using its committees, WIPO delivers BDV PPP and dispute resolution services and organises knowledge transfer-related activities, including capacity building and training on transfer of technology⁹.

---

Good Practices

By reference to the definition of technology and knowledge transfer, certain categories of transactions constitute technology transfer [7]. Consequently, the activities leading to the identified transactions can be accepted as facilitators of technology transfer:

- The assignment, sale and licensing of all forms of industrial property, except for trade/service marks and trade names;
- The provision of know-how and technical expertise in the form of feasibility studies, plans, diagrams, models, instructions, guides, formulae, designs, specifications and equipment for training, services involving technical advisory and managerial personnel, and personnel training for the installation, operation and functioning of plant and equipment, and turnkey projects;
- The provision of technological contents of industrial and technical cooperation arrangements.

Based on extensive literature review on Sub-Saharan African economies, authors assert that governments targeting Foreign Direct Investment (FDI) and focusing on implementing development objectives has a set of factors to facilitate technology and knowledge transfer [9]. The factors that influence technology and knowledge transfer include effective industry institutions, education effectiveness, and the joint presence of high government policy incentives, effective industry institutions, and education effectiveness. The substance in this assertion centres around the number of institutions involved, the quality and quantity of local firms, the human resource management, and the general policy framework and regulations designed to influence the quality of technology and knowledge transfer [9].

Study Findings

All centres work with industrial and other partners in various sectors to deliver technology and knowledge transfer services and products. Within industry-focused research centres, they look to marrying technology transfer with IP arrangements by simplifying both the pricing and monitoring of IP processes and activities. As a further way to facilitate technology transfer, the IPC brings together the TTOs and its partner universities into the management of IP created by the centre with the aim of simplifying the IP process.

Industry-focused centres employ use cases and applications to bring their theoretical know-how into tangible, practical solutions for its partners in a joint effort. Theoretical knowledge enables the centre to design customised applications as problem solutions and to validate the solutions as best-fit solutions for the partners' specific business challenges. Many centres are using a staged-gate model within research and innovation activities to support the way they transfer technology and knowledge to end-users. Centres use a range of techniques including processes, experiments,
prototypes, pilots, and IP to transfer knowledge. Some centres include a feedback loop like two-way knowledge transfer of real-world scenarios and real business problems. Table 38 provides a summary of the best practices of technology transfer.

<table>
<thead>
<tr>
<th>Practice</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Delivery of partner-centric solutions, being industry-focused in research methodology promotes technology development with industry relevance.</td>
<td></td>
</tr>
<tr>
<td>Usage of the feedback cycle to understand industry need as a guide to the design and development of effective solutions that can be passed to partners benefits and technology and knowledge transfer.</td>
<td></td>
</tr>
<tr>
<td>Use of the iterative cycle of experiments, prototypes, pilots, and IP with feedback loop in the iteration can promote relevant technology development.</td>
<td></td>
</tr>
<tr>
<td>Simplification of the IP process.</td>
<td></td>
</tr>
<tr>
<td>Innovation Cycle demands the project work be designed to combine research with real-world deployment to meet real business problems.</td>
<td></td>
</tr>
<tr>
<td>Innovate with industrial partners to follow a cycle of short-term use-case-based activities with knowledge feedback into the projects.</td>
<td></td>
</tr>
<tr>
<td>Projects are developed and approved using a standard staged-gate model for innovation through the development of technology roadmaps used to define areas of interest to the centre.</td>
<td></td>
</tr>
<tr>
<td>Catalogue of technology demonstrators, IP and Analytics technology reviews, state-of-the-art reviews of data analytics technologies and tools, experts in data analytics and visualisation, etc., all made available to its members.</td>
<td></td>
</tr>
<tr>
<td>To establish industry collaboration, the centre has calls for demonstrator proposals; selection of the proposal, rating based on criteria and decision on the best choices of proposals.</td>
<td></td>
</tr>
<tr>
<td>Transfer knowledge and expertise via a feedback loop in the innovation cycle</td>
<td></td>
</tr>
<tr>
<td>▪ Implement a prototype in industrial pilots and research results in products of industry partners.</td>
<td></td>
</tr>
<tr>
<td>▪ Identify constraints in existing tools; identify opportunities for changes in work practices. Demonstrate the use of tools of partners in prototypes.</td>
<td></td>
</tr>
</tbody>
</table>

Table 38: Good practices in technology and knowledge transfer

I. Performance and Impact Assessment

Context and Definition
One of the definitions of a CoE in the context of project management is that it creates an environment to deliver a continuous stream of successfully managed projects the success of which is measured by having achieved performance that is in the best interest of the whole company as well as the specific project [13]. The knowledge of the requirement is to be subjected to performance appraisal as the centre’s policy creates an impetus that drives a harder work environment and, consequently, an
improved level of achievement. This applies to the individual staff/researcher level as well as the organisational level from the point of view of funding agencies.

**Good Practices**

Findings from literature and this research suggest that to meet high targets often demanded of research centres, the goals must be broken down into manageable sizes, operationalised into daily activities, and imbedded into KPIs with metrics to measure progress in the centres towards the set goals, periodically. The management must be familiar with their unique set of success factors or enablers and should make an effort to harness them even deeper.

The impact assessment may be improved through practices that also create a direct interaction between the research centre and society. For example, one centre has been given an expanded mandate to the measurement of EPE performances using selected KPI metrics. This directive makes it compulsory for each researcher or staff of the centre to be involved in at least four of EPE programmes per annum, and that the centre must record a 25% personnel involvement by 2019. Some of the recommended activities include the following:

1. Contributions to online communications
2. Development of online resources
3. Communications/Engagement training
4. Collateral for a lay audience
5. Development of school interactions
6. Teaching materials/methods
7. Festivals/events in institutions

**Study Findings**

This study reveals that a couple of challenges affect the performance of research instructions in terms of operations, organisational autonomy, operations. All these challenges and impacts more the research centres’ output, according to findings. Firstly, the Big Data industry directly affects the strategy and performance of the BDAICoE and secondly, another industry status with regards to their relative strengths or weaknesses has the capability to impact negatively on the core elements of the BDAICoE model. In a nutshell, an industry status has been argued to represent a significant influence on the research performance of academics [3]. Besides, the lack of adequate funding, the need to meet extra high-level targets, etc., often put most research centres under pressure of performance. Another constraining challenge is the conflicting interests of stakeholders arising from the fact that funding agencies are usually in the public or academic sector while the users of their research outputs are usually in the industry, the private sector.

- To ameliorate some of these challenges, a few centres create an Industry Advisory and Scientific Advisory Committee to oversee performances in line
with their KPIs. Other strategies have been suggested as possible ways of improving performances and operational efficiency: For example, decision-making based on meritocracy should positively affect funding decisions, performance evaluations, and rewards. For example, one centre measures its performances bi-annually and has operationalised KPIs to cover a lot of the impact areas, such as economic, commercialisation and academics. Another centre monitors its performance and impacts through KPI review on a monthly basis internally and quarterly with external funders. Table 39 shows a summary of good practices in performance and impact assessment for a research centre.

<table>
<thead>
<tr>
<th>Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Break high-target goals into manageable sizes, create KPIs, operationalise and measure periodically.</td>
</tr>
<tr>
<td>Management should familiarise with enablers/success factors and harness adequately.</td>
</tr>
<tr>
<td>Create direct interaction between centre and society through enlightenment programmes and training and use policies to backup participation where necessary.</td>
</tr>
<tr>
<td>To improve performance:</td>
</tr>
<tr>
<td>▪ ameliorate challenges through the use of dedicated committee to oversee performances of centre</td>
</tr>
<tr>
<td>▪ apply meritocracy to decision-making particularly on funding decisions.</td>
</tr>
</tbody>
</table>

Table 39: Good practices in performance and impact assessment
Appendix B: Glossy Templates

Figure 45: Glossy template of Persons of Excellence
Figure 46: Glossy template of BDAICoE framework
Figure 47: Glossy template of Case Studies
Appendix C: BDVe Data Science Badges – 1st
Pre-Pilot Feedback from Academia

A. Introductory Note for BDVe’s Data Science Badges: Academic Feedback

One of the principal goals of the BDVe project is to promote the acquisition of skills in data science in the EU. In order to further this goal, one of the project’s tasks involves the design and implementation of a data science skills recognition system. The proposed system is based upon the definition of a collection of badges to recognise data science skills. Once a consensus regarding the program is reached amongst interested parties in both academia and industry, Universities and other learning institutions will be able to apply to issue these badges to their students. These micro-credentials will serve as a common framework for:
- Students - to help them compare the data science skills acquired in different training programmes in data science
- Employers - to help them assess the skills of candidates for data science jobs
- Universities - to help them design their data science programmes to meet the needs of industry

To take advantage of previous efforts to promote data science skills in the EU, we are using the EDISON Competence Framework (http://edison-project.eu/data-science-competence-framework-cf-ds) as the basis for the program. Following this framework, we propose that the program consists of five sets of badges, one set for each of EDISON’s competence groups:
- Data Science Analytics
- Data Engineering
- Data Science Management
- Business Process Management
- Data Science Research Method and Project Management

Also, as given by EDISON, each badge set will contain three different badges at different levels: basic, intermediate, and expert. (A complete listing of the five badge groups along with the skills required for each of their three levels is included as an Appendix to this document.)

The BDVe plans to run a pilot evaluation of one badge group (the three levels of the Data Science Analytics Badge) in the fall of 2018 with the hope of being able to start the complete program for university issuers in January of 2019.

To help us prepare for the pilot study, we would like to request some feedback from your university regarding the following:
Does your principal data science program impart the required skills to issue the three levels of the Data Analytics Badge? If it does not, do you think that it is reasonable to expect that it does?

Which kind of evidence does your program use to ensure that your students have acquired the required skills?

How much effort was needed to gather the requested information regarding the evidences for the required skills?

To help organise these responses, please complete the following questionnaire and return it by email before June 29th 2018.

Many thanks for your time,

Ernestina Menasalvas, Ana M. Moreno and Nik Swoboda (UPM BDVe team)


B. Academic Questionnaire: Required Skills for the Data Analytics Badges

Please select the most relevant data science program at your institution and evaluate whether it satisfies the required skills for the three levels of the Data Analytics Badge. For each of the required skills, please indicate whether your program includes training to acquire that skill and what evidence your program uses to evaluate the acquisition of the skill on an individual student by student basis. Examples of such evidence could include the score given to a particular question on an exam, the grade given to a part of an assignment, or the evaluation of the relevant part of a rubric for an oral presentation.

Program name:

<table>
<thead>
<tr>
<th>Data Analytics Badge – Basic Level</th>
<th>Included in the program?</th>
<th>Evidence used to support the skill acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required skill</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choose and execute existing data analytics and predictive analytics tools.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Design and evaluate predictive analysis tools to discover new relations.

Name and use basic performance assessment metrics and tools.

Define data elements necessary to develop specified data analytics.

Choose and execute standard visualisation.

### Data Analytics Badge – Intermediate Level

<table>
<thead>
<tr>
<th>Required skill</th>
<th>Included in the program?</th>
<th>Evidence used to support the skill acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Select most appropriate statistical techniques and model available data to deliver insights</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analyse available data sources and develop a tool that works with complex datasets.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use multiple performance and accuracy metrics; select and use most appropriate for specific type of data analytics application.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Develop specialised analytics to enable decision-making.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Build visualisations for complex and variable data.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Data Analytics Badge – Expert Level

<table>
<thead>
<tr>
<th>Required skill</th>
<th>Included in the program?</th>
<th>Evidence used to support the skill acquisition</th>
</tr>
</thead>
</table>
### Design and evaluate predictive analysis tools to discover new relations.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>

### Assess and optimise organisation processes using statistical techniques.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>

### Assess, adapt, and combine data sources to improve analytics.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>

### Evaluate and recommend the most appropriate metrics, propose new for new applications.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>

### Design specialised analytics to improve decision-making.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>

### Create and optimise visualisations to influence executive decisions.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>

If your program does not include training for any of the required skills, do you think that it is unreasonable to expect that it does?

How much effort was needed to analyse your program and provide the above information for each badge?

Do you have any other comments/suggestions regarding the badge groups or their required skills?
Appendix D: BDVe Data Science Badges – 1st
Pre-Pilot Feedback from Industry

A. Introductory Note for BDVe’s Data Science Badges: Industry Feedback

One of the principal goals of the BDVe project is to promote the acquisition of skills in data science in the EU. In order to further this goal, one of the project’s tasks involves the design and implementation of a data science skills recognition system. The proposed system is based upon the definition of a collection of badges to recognise data science skills. Once a consensus regarding the program is reached amongst interested parties in both academia and industry, Universities and other learning institutions will be able to apply to issue these badges to their students. These micro-credentials will serve as a common framework for:

- Students - to help them compare the data science skills acquired in different training programmes in data science
- Employers - to help them assess the skills of candidates for data science jobs
- Universities - to help them design their data science programmes to meet the needs of industry

To take advantage of previous efforts to promote data science skills in the EU, we recommend using the EDISON Competence Framework (http://edison-project.eu/data-science-competence-framework-cf-ds) as the basis for the program. Following this framework, we propose that the program consists of five sets of badges, one set for each of EDISON’s competence groups:

- Data Science Analytics
- Data Engineering
- Data Science Management
- Business Process Management
- Data Science Research Method and Project Management

Also, as given by EDISON, each badge set will contain three different badges at different levels: basic, intermediate, and expert. (A complete listing of the five badge groups along with the skills required for each of their three levels is included as an Appendix to this document.)

The BDVe plans to run a pilot evaluation of one badge group (the three levels of the Data Science Analytics Badge) in the fall of 2018 with the hope of being able to start the complete program for university issuers in January of 2019.

To help us prepare for the pilot study, we would like to request some feedback from the human resources department of your company regarding the following:
Could you please rate the relevance of the required skills for a small number of data science positions in your company?

Are the descriptions of the required skills for each badge easy to understand? If not, could you help us improve them?

Are there data science skills in the area of data analytics which your company typically requires which are not included in the requirements for the badges?

To help organise these responses, please complete the following questionnaire and return it by email before June 29th 2018.

Many thanks for your time,

Ernestina Menasalvas, Ana M. Moreno and Nik Swoboda (UPM BDVe team)


B. Industrial Questionnaire: Required Skills for the Data Analytics Badges

Using the following scale:
1 – Not at all relevant
2 – Slightly relevant
3 – Neutral
4 – Moderately relevant
5 – Extremely relevant

Please rate the relevance of the following required skills to three typical data science positions in your company
Data science job title 1:
Data science job title 2:
Data science job title 3:

<table>
<thead>
<tr>
<th>Data Analytics Badge – Basic Level</th>
<th>Relevance to job title 1?</th>
<th>Relevance to job title 2?</th>
<th>Relevance to job title 3?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose and execute existing data analytics and predictive analytics tools.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Design and evaluate predictive analysis tools to discover new relations.

### Name and use basic performance assessment metrics and tools.

### Define data elements necessary to develop specified data analytics.

### Choose and execute standard visualisation.

#### Data Analytics Badge – Intermediate Level

<table>
<thead>
<tr>
<th>Required skill</th>
<th>Relevance to job title 1?</th>
<th>Relevance to job title 2?</th>
<th>Relevance to job title 3?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Select most appropriate statistical techniques and model available data to deliver insights</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analyse available data sources and develop a tool that works with complex datasets.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use multiple performance and accuracy metrics; select and use most appropriate for specific type of data analytics application.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Develop specialised analytics to enable decision-making.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Build visualisations for complex and variable data.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Data Analytics Badge – Expert Level

<table>
<thead>
<tr>
<th>Required skill</th>
<th>Relevance to job title 1?</th>
<th>Relevance to job title 2?</th>
<th>Relevance to job title 3?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design and evaluate predictive analysis tools to discover new relations.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assess and optimise organisation processes using statistical techniques.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Assess, adapt, and combine data sources to improve analytics.

Evaluate and recommend the most appropriate metrics, propose new for new applications.

Design specialised analytics to improve decision-making.

Create and optimise visualisations to influence executive decisions.

<table>
<thead>
<tr>
<th>Required skill</th>
<th>If not understandable, could you suggest an improvement?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose and execute existing data analytics and predictive analytics tools.</td>
<td></td>
</tr>
<tr>
<td>Identify existing requirements and develop predictive analysis tools.</td>
<td></td>
</tr>
<tr>
<td>Design and evaluate predictive analysis tools to discover new relations.</td>
<td></td>
</tr>
<tr>
<td>Name and use basic performance assessment metrics and tools.</td>
<td></td>
</tr>
<tr>
<td>Define data elements necessary to develop specified data analytics.</td>
<td></td>
</tr>
<tr>
<td>Choose and execute standard visualisation.</td>
<td></td>
</tr>
</tbody>
</table>
Appendix E: BDVe’s Data Science Badges – 2\textsuperscript{nd} Pre-Pilot Feedback from Academia

One of the principal goals of the BDVe project is to promote the acquisition of skills in data science in the EU. In order to further this goal, one of the project’s tasks involves the design and implementation of a data science skills recognition system. The proposed system is based upon the definition of a collection of badges to recognise data science skills. Once a consensus regarding the program is reached amongst interested parties in both academia and industry, Universities and other learning institutions will be able to apply to issue these badges to their students. These micro-credentials will serve as a common framework for:

▪ Students - to help them compare the data science skills acquired in different training programmes in data science
▪ Employers - to help them assess the skills of candidates for data science jobs
▪ Universities - to help them design their data science programmes to meet the needs of industry

To take advantage of previous efforts to promote data science skills in the EU, we are using the EDISON Competence Framework (http://edison-project.eu/data-science-competence-framework-cf-ds) as the basis for the program. Following this framework, we propose that the program consists of five sets of badges, one set for each of EDISON’s competence groups:

▪ Data Science Analytics
▪ Data Engineering
▪ Data Science Management
▪ Business Process Management
▪ Data Science Research Method and Project Management

The EDISON framework includes three levels for each of these competency groups. However, based upon feedback received from an initial survey of academia and industry, we have rewritten the descriptions of the skills and consolidated the three levels into two (an academic and an industrial level). The academic level of the badge will be offered to students who have received training to acquire the skills, and the industrial level of the badge will be offered to data scientists who have professional experience with those skills.

The BDVe plans to run a pilot evaluation of one badge (the academic level of the Data Science Analytics Badge) in the spring of 2019.

To help us prepare for the pilot study, we would like to request some feedback from your university regarding the following:
Does your principal data science program impart the required skills to issue the Academic Level of the Data Analytics Badge? If it does not, do you think that it is reasonable to expect that it does?

Which kind of evidence does your program use to ensure that your students have acquired the required skills?

How much effort was needed to gather the requested information regarding the evidences for the required skills?

To help organise these responses, please complete the following questionnaire and return it by email before January 16st 2019.

Many thanks for your time,

Ernestina Menasalvas, Ana M. Moreno and Nik Swoboda (UPM BDVe team)

**Questionnaire: Required Skills for the Data Analytics Badge**

Please select the most relevant data science program at your institution and evaluate whether it satisfies the required skills for the Academic Level of the Data Analytics Badge. For each of the required skills, please indicate whether your program includes training to acquire that skill and what evidence your program uses to evaluate the acquisition of the skill on an individual student by student basis. Examples of such evidence could include the score given to a particular question on an exam, the grade given to a part of an assignment, or the evaluation of the relevant part of a rubric for an oral presentation.

**Program name:**

<table>
<thead>
<tr>
<th>Required skill</th>
<th>Included in the program?</th>
<th>Evidence used to support the skill acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSA.1. Identify existing requirements and choose and execute the most appropriate data discovery technique to solve a problem depending on the nature of the data and the goals to be achieved.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSA.2. Select the most appropriate statistical techniques to understand and prepare data prior to modelling to deliver insights.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSA.3. Assess, adapt, and combine data sources to improve analytics.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSA.4. Use the most appropriate metrics to evaluate and validate results, proposing new metrics for new applications if required.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSA.5. Design and evaluate analysis tools to discover new relations in order to improve decision-making.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSA.6. Use visualisation techniques to improve the presentation of the results of a data science project in any of its phases.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If your program does not include training for any of the required skills, do you think that it is unreasonable to expect that it does?
How much effort was needed to analyse your program and provide the above information for each badge?

Do you have any other comments/suggestions regarding the required skills?
Appendix F: BDVe’s Data Science Badges – 2nd Pre-Pilot Feedback from Industry

BDVe’s Data Science Badges – Pre-Pilot Feedback from Industry

One of the principal goals of the BDVe project is to promote the acquisition of skills in data science in the EU. In order to further this goal, one of the project’s tasks involves the design and implementation of a data science skills recognition system. The proposed system is based upon the definition of a collection of badges to recognise data science skills. Once a consensus regarding the program is reached amongst interested parties in both academia and industry, Universities and other learning institutions will be able to apply to issue these badges to their students. These micro-credentials will serve as a common framework for:

- Students - to help them compare the data science skills acquired in different training programmes in data science
- Employers - to help them assess the skills of candidates for data science jobs
- Universities - to help them design their data science programmes to meet the needs of industry

To take advantage of previous efforts to promote data science skills in the EU, we are using the EDISON Competence Framework ([http://edison-project.eu/data-science-competence-framework-cf-ds](http://edison-project.eu/data-science-competence-framework-cf-ds)) as the basis for the program. Following this framework, we propose that the program consists of five sets of badges, one set for each of EDISON’s competence groups:

- Data Science Analytics
- Data Engineering
- Data Science Management
- Business Process Management
- Data Science Research Method and Project Management

The EDISON framework includes three levels for each of these competency groups. However, based upon feedback received from an initial survey of academia and industry, we have rewritten the descriptions of the skills and consolidated the three levels into two (an academic and an industrial level). The academic level of the badge will be offered to students who have received training to acquire the skills, and the industrial level of the badge will be offered to data scientists who have professional experience with those skills.

The BDVe plans to run a pilot evaluation of one badge (the academic level of the Data Science Analytics Badge) in the spring of 2019.
To help us prepare for the pilot study, we would like to request some feedback from your organisation regarding the following:

- Could you please rate the relevance of the required skills for a small number of data science positions in your company?
- Are the descriptions of the required skills for each badge easy to understand? If not, could you help us improve them?
- Are there data science skills in the area of data analytics which your company typically requires which are not included in the requirements for the badges?

To help organise these responses, please complete the following questionnaire and return it by email before January 16th 2019.

Many thanks for your time,
Ernestina Menasalvas, Ana M. Moreno and Nik Swoboda (UPM BDVe team)
Questionnaire: Required Skills for the Data Analytics Badges

Using the following scale:

1 – Not at all relevant
2 – Slightly relevant
3 – Neutral
4 – Moderately relevant
5 – Extremely relevant

Please rate the relevance of the following required skills to three typical data science positions in your company

Data science job title 1:
Data science job title 2:
Data science job title 3:

<table>
<thead>
<tr>
<th>Data Analytics Badge – University Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required skill</td>
</tr>
<tr>
<td>Relevance to job title 1?</td>
</tr>
<tr>
<td>DSA.1. Identify existing requirements and choose and execute the most appropriate data discovery technique to solve a problem depending on the nature of the data and the goals to be achieved.</td>
</tr>
<tr>
<td>DSA.2. Select the most appropriate statistical techniques to understand and prepare data prior to modelling to deliver insights.</td>
</tr>
<tr>
<td>DSA.3. Assess, adapt, and combine data sources to improve analytics.</td>
</tr>
<tr>
<td>DSA.4. Use the most appropriate metrics to evaluate and validate results, proposing new metrics for new applications if required.</td>
</tr>
<tr>
<td>DSA.5. Design and evaluate analysis tools to discover new relations in order to improve decision-making.</td>
</tr>
</tbody>
</table>
DSA.6. Use visualisation techniques to improve the presentation of the results of a data science project in any of its phases.

Are the descriptions of the required skills for each of the badges easily understood?

<table>
<thead>
<tr>
<th>Required skill</th>
<th>If not understandable, could you suggest an improvement?</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSA.1. Identify existing requirements and choose and execute the most appropriate data discovery technique to solve a problem depending on the nature of the data and the goals to be achieved.</td>
<td></td>
</tr>
<tr>
<td>DSA.2. Select the most appropriate statistical techniques to understand and prepare data prior to modelling to deliver insights.</td>
<td></td>
</tr>
<tr>
<td>DSA.3. Assess, adapt, and combine data sources to improve analytics.</td>
<td></td>
</tr>
<tr>
<td>DSA.4. Use the most appropriate metrics to evaluate and validate results, proposing new metrics for new applications if required.</td>
<td></td>
</tr>
<tr>
<td>DSA.5. Design and evaluate analysis tools to discover new relations in order to improve decision-making.</td>
<td></td>
</tr>
<tr>
<td>DSA.6. Use visualisation techniques to improve the presentation of the results of a data science project in any of its phases.</td>
<td></td>
</tr>
</tbody>
</table>

Are there data science skills in the area of data analytics which your company typically requires which are not included in the requirements for the badges?
Do you have any other comments/suggestions regarding the badge groups or their required skills?
Appendix G: BDV Data Science Analytics Badge, Academic Level – Template for Evidences

BDV
Data Science Analytics Badge
Academic Level

Badge Issuing Application
for the
<Program Name>
at
<Institution Name>
<Location>
<Date>

CONFIDENTIAL
The information provided in this application is for the confidential use of the BDVe project and its authorised agents, and will not be disclosed to third parties without the authorisation of the institution concerned, except for its use in summary data which is not identifiable to a specific institution.
This document provides a template to present evidence to show that an academic program both teaches and evaluates the skills needed to meet the requirements of the Academic Level of the BDV Data Science Analytics Badge. Once complete, this document will contain some basic information about the program to be evaluated as well as specific details regarding how the skills required by the badge are acquired by the program’s students.

Further information and detailed instructions for completing this application can be found in the BDV Data Science Analytics Badge Academic Level INFORMATION AND GUIDELINES FOR APPLICANTS document available on the badge program’s web page: http://www.big-data-value.eu/skills/skills-recognition-program/

The skills required for the Data Analytics Badge are listed in Table 40.

<table>
<thead>
<tr>
<th>Required skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSA.1. Identify existing requirements to choose and execute the most appropriate</td>
</tr>
<tr>
<td>DSA.2. Select the most appropriate techniques to understand and prepare data</td>
</tr>
<tr>
<td>DSA.3. Assess, adapt, and combine data sources to improve analytics.</td>
</tr>
<tr>
<td>DSA.4. Use the most appropriate metrics to evaluate and validate results,</td>
</tr>
<tr>
<td>proposing new metrics for new applications if required.</td>
</tr>
<tr>
<td>DSA.5. Design and evaluate analysis tools to discover new relations in order</td>
</tr>
<tr>
<td>to improve decision-making.</td>
</tr>
<tr>
<td>DSA.6. Use visualisation techniques to discover the presentation of the</td>
</tr>
<tr>
<td>results of a data science project in any of its phases.</td>
</tr>
</tbody>
</table>

Table 40: BDV Data Science Analytics Badge Skills
GENERAL INFORMATION

A. Contact Information

Please provide the following information for the primary contact person of the program applying for the badge.

- Name:
- Position:
- Mailing address:
- Telephone number:
- E-mail address:

☐ By checking this box, the contact person agrees to the use of this contact information by members of the BDVe project to send communications related to the badge issuing application process, and if the application is accepted, he/she also agrees to its use for future communications by the BDVe related to the badge program. This data will be stored confidentially and will not be used for other purposes or disclosed to others without the previous consent of the contact person. The contact person can withdraw this consent or modify this data at any time by contacting the BDVe project.

B. Program Background

Please provide a brief history of the program. Be sure to include the following information in this section:

- A link to the program’s web page
- A list of all national and international accreditations obtained by the program with the dates received.
- The number of students to successfully complete the program in each of the last five years.
- How this program is taught: traditional lectures/laboratories, online learning, etc.

C. Options

List and briefly describe any specialisations, tracks, concentrations, etc. included in the program.
Complete the following tables with references to and descriptions of the evidence submitted to show the acquisition of the corresponding skills. When describing the evidence, please be specific regarding which part of the evidence file provides evidence of the corresponding skill. For example, if the evidence file contains an exam, please state which question or questions provide evidence. It is possible that the same evidence file provides evidences for several skills. Each evidence file should be included with the submitted application.

In the case of optional courses/activities, indicate possible itineraries in the comments section of each of the tables, stating clearly each of the minimum sets of evidences needed to demonstrate the acquisition of the skill.

<table>
<thead>
<tr>
<th>Name of the course or activity</th>
<th>Optional/ Mandatory</th>
<th>The average number of hours that students spend in the course/activity acquiring the skill</th>
<th>Evidence used to evaluate the acquisition of the skill (exam, assignment, …)</th>
<th>Name of the evidence file</th>
<th>The minimum level required to demonstrate the acquisition of the skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course 1</td>
<td></td>
<td>X hours</td>
<td>Exam 1, questions 2-3</td>
<td></td>
<td>Add as many rows as needed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Assignment 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>…</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Add as many rows as needed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Course 2</td>
<td></td>
<td>Y hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>add as many rows as needed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum number of hours that the student dedicates to acquiring the skill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comments (optional)
## DSA.2. Select the most appropriate techniques to understand and prepare data prior to modelling to deliver insights

<table>
<thead>
<tr>
<th>Name of the course or activity</th>
<th>Optional/ Mandatory</th>
<th>The average number of hours that students spend in the course/activity acquiring the skill</th>
<th>Evidence used to evaluate the acquisition of the skill (exam, assignment, …)</th>
<th>Name of the evidence file</th>
<th>The minimum level required to demonstrate the acquisition of the skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course 1</td>
<td></td>
<td>X hours</td>
<td>Exam 1, questions 2-3, Assignment 4, …</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Add as many rows as needed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Course 2</td>
<td></td>
<td>Y hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>add as many rows as needed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum number of hours that the student dedicates to acquiring the skill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comments (optional)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### DSA.3. Assess, adapt, and combine data sources to improve analytics

<table>
<thead>
<tr>
<th>Name of the course or activity</th>
<th>Optional/Mandatory</th>
<th>The average number of hours that students spend in the course/activity acquiring the skill</th>
<th>Evidence used to evaluate the acquisition of the skill (exam, assignment, …)</th>
<th>Name of the evidence file</th>
<th>The minimum level required to demonstrate the acquisition of the skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course 1</td>
<td></td>
<td>X hours</td>
<td>Exam 1, questions 2-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Assignment 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>…</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Add as many rows as needed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Course 2</td>
<td></td>
<td>Y hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>add as many rows as needed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Minimum number of hours that the student dedicates to acquiring the skill</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comments (optional)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### DSA.4. Use the most appropriate metrics to evaluate and validate results, proposing new metrics for new applications if required

<table>
<thead>
<tr>
<th>Name of the course or activity</th>
<th>Optional/Mandatory</th>
<th>The average number of hours that students spend in the course/activity acquiring the skill</th>
<th>Evidence used to evaluate the acquisition of the skill (exam, assignment, …)</th>
<th>Name of the evidence file</th>
<th>The minimum level required to demonstrate the acquisition of the skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course 1</td>
<td></td>
<td>X hours</td>
<td>Exam 1, questions 2-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Assignment 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>…</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Add as many rows as needed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Course 2</td>
<td></td>
<td>Y hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>add as many rows as needed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum number of hours that the student dedicates to acquiring the skill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comments (optional)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name of the course or activity</td>
<td>Optional/ Mandatory</td>
<td>The average number of hours that students spend in the course/activity acquiring the skill</td>
<td>Evidence used to evaluate the acquisition of the skill (exam, assignment, …)</td>
<td>Name of the evidence file</td>
<td>The minimum level required to demonstrate the acquisition of the skill</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>---------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Course 1</td>
<td></td>
<td>X hours</td>
<td>Exam 1, questions 2-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Assignment 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>…</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Add as many rows as needed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Course 2</td>
<td></td>
<td>Y hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>add as many rows as needed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum number of hours that the student dedicates to acquiring the skill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comments (optional)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DSA.6. Use visualisation techniques to improve the presentation of the results of a data science project in any of its phases

<table>
<thead>
<tr>
<th>Name of the course or activity</th>
<th>Optional/Mandatory</th>
<th>The average number of hours that students spend in the course/activity acquiring the skill</th>
<th>Evidence used to evaluate the acquisition of the skill (exam, assignment, …)</th>
<th>Name of the evidence file</th>
<th>The minimum level required to demonstrate the acquisition of the skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course 1</td>
<td></td>
<td>X hours</td>
<td>Exam 1, questions 2-3</td>
<td>Assignment 4</td>
<td>Add as many rows as needed</td>
</tr>
<tr>
<td>Course 2</td>
<td></td>
<td>Y hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>add as many rows as needed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Minimum number of hours that the student dedicates to acquiring the skill</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comments (optional)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I, the undersigned, hereby state that the information contained in this application is to the best of my knowledge correct and complete.

Place

Date (dd-mm-yyyy)

Name of the applicant organisation
Appendix H: BDV Data Science Analytics Badge, Academic Level – Information and Guidelines for Applicants

BDV
Data Science Analytics Badge
Academic Level

INFORMATION AND GUIDELINES FOR APPLICANTS
INTRODUCTION

This document provides information regarding the process of applying for the Academic Level of the Data Science Analytics Badge. Additionally, this document presents the evaluation criteria that will be used to assess applications so that applicants are aware of how applications will be evaluated.

The skills required for the Data Analytics Badge are listed in Table 41.

<table>
<thead>
<tr>
<th>Data Science Analytics Badge v1-0 – Academic Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required skills</td>
</tr>
<tr>
<td>DSA.1. Identify existing requirements to choose and execute the most appropriate data discovery techniques to solve a problem depending on the nature of the data and the goals to be achieved.</td>
</tr>
<tr>
<td>DSA.2. Select the most appropriate techniques to understand and prepare data prior to modelling to deliver insights.</td>
</tr>
<tr>
<td>DSA.3. Assess, adapt, and combine data sources to improve analytics.</td>
</tr>
<tr>
<td>DSA.4. Use the most appropriate metrics to evaluate and validate results, proposing new metrics for new applications if required.</td>
</tr>
<tr>
<td>DSA.5. Design and evaluate analysis tools to discover new relations in order to improve decision-making.</td>
</tr>
<tr>
<td>DSA.6. Use visualisation techniques to improve the presentation of the results of a data science project in any of its phases.</td>
</tr>
</tbody>
</table>

Table 41: BDV Data Science Analytics Badge Skills

INFORMATION REGARDING THE APPLICATION PROCESS

Only approved programmes can issue the Academic Level of the BDV Data Science Analytics Badge. To obtain this approval, a program must apply to issue the badge by providing evidence to show how they can ensure that their students have acquired the corresponding skills.

To simplify the application process, the BDV Data Science Analytics Badge Academic Level Template for Badge Issuing Applications is provided to help collect specific details about how each of the skills required by the badge is acquired by the program’s students. This document also includes the presentation of some general information about the applicant’s program. This template should be completed and sent via email along with the different files referenced by the application, which contain the corresponding evidences to bdv.badges@big-data-value.eu.
A. How to Complete the Template for Issuing Applications

V. General Information
The first section of the template is to collect general information about the applicant program and the institution responsible for the program.

A. Contact Information
Here you will provide contact information for the program applying for the badge. The contact person need not be the same person that signs the application form. Note: this contact person must agree to the use of their personal data as required by the GDPR.

B. Program Background
This section includes general information about the program, its web page, accreditations, number of graduates, and a general description of how the program is taught.

C. Options
This section describes possible specialisations in the program. In the case of joint programmes, clearly, state this fact as well as any other details needed to understand the subsequent evidences.

VI. Acquisition of Skills
This section contains a list of the evidences that demonstrate how the program can ensure that each of the skills is acquired by the students of the program. For each of the skills, a table with the corresponding information should be completed according to the following guidelines:

- Name of the course or activity: Indicate the name of the different courses or activities in the program in which students acquire the corresponding skill. Courses could be any class included in the program, while activities could be seminars, workshops, or any other activity.
- Optional/Mandatory: Indicate whether the corresponding course or activity is mandatory or optional for the students of the program. In the case of optional courses/activities, please include in the comments section the possible itineraries that could be followed to earn the badge.
- The average number of hours that students spend in the course/activity acquiring the skill: This column contains information about the estimated number of hours that instructors expect that students need to acquire the corresponding skill in each of the courses/activities. This number includes the average number of hours dedicated to acquiring the skill as well as the time required to produce the different evidences associated with each course/activity.
Evidence used to evaluate the acquisition of the skill: Any evaluated activity which is included in a degree program can be considered to be evidence. Typical examples include exam questions, laboratory exercises, sections of oral exams, and parts of final projects. Typical evidence files in these cases would include a copy of the exam’s questions, the instructions, and description of the laboratory’s submitted work, the syllabus of an oral exam, or the description of the requirements for the final projects. If an evidence file includes multiple evidences, please clearly specify which evidence corresponds with each skill. For example, if an evidence file contains an exam please state which question or questions provide the corresponding evidence.

Name of the evidence file: Provide the name of the file containing each of the evidences. We suggest that you name the files so that they refer to the skill, course, and evidence provided. For example, DS1 – Big Data Course – Exam 1 – Questions 2and3.pdf. Each evidence file should be included with the submitted application in pdf format. As mentioned previously, this file can be a copy of the exam’s questions, a description of a project, syllabus, etc. Please mark clearly the particular part of the content that provides the evidence. For example, if in the previous column, it is mentioned that the evidence is provided in questions 2 and 3 of Exam 1, the file containing Exam 1 should include questions 2 and 3 highlighted. Please do not submit examples of student work or any other information considered to be confidential by your institution.

The minimum level required to demonstrate the acquisition of the skill: the minimum grade that the students must earn, in the corresponding part of the evidence file, in order to show that they have acquired the skill. Examples are: 7 points out of 10, 65 points out of 100, B from the range A-B-C-D.

The minimum number of hours that the student dedicates to acquire the skill: this number represents the sum of the average number of hours that the student dedicates to each of the evidences for each particular skill. In case of optional courses/activities, this number should be the minimum number of hours included in all of the possible itineraries. For example, if there are two possible itineraries, Itinerary 1 with X hours and Itinerary 2 with Y hours, you should include here the lesser of X and Y.

Comments: This section is optional and can include any information that the applicant considers relevant for the reviewers. In case of optional courses/activities, this section should include a brief description of the different itineraries that the students can take, stating clearly each of the minimum sets of evidences needed to demonstrate the acquisition of the skill. For example, Itinerary 1: activity 1, activity 2, activity 3; Itinerary 2: activity 1, activity 3, activity 4.
EVALUATION GUIDELINES FOR THE DATA SCIENCE ANALYTICS BADGE

Each application will be independently reviewed by two experts. A final decision will be made if their recommendations coincide. If the two reviewers are not able to reach a consensus, a third reviewer will be asked to participate in the process. All reviewers participating in the process will sign a code of conduct which, amongst other things, requires the non-disclosure of all information contained in the applications as well as the exclusion of reviewers who could have a conflict of interest.

A. Criteria for evaluation

Based on the information presented, the acquisition of each of the skills of the Data Science Analytics Badge will be scored as follows:

- 0 — Fail: The application fails to address the skill, or the acquisition of the skill cannot be assessed due to missing or incomplete information. (‘Obvious clerical errors’ are not considered to result in incomplete information.)
- 1 — Poor: The skill is inadequately addressed, or there are serious inherent weaknesses.
- 2 — Fair: The application broadly addresses the skill, but there are significant weaknesses.
- 3 — Good: The application addresses the skill well but with a number of shortcomings.
- 4 — Very good: The application addresses the skill very well but with a small number of shortcomings.
- 5 — Excellent: The application successfully addresses all relevant aspects of the skill; any shortcomings are minor.

These scores will be assigned based upon the following criteria:

- Is the average total number of hours that a student dedicates to the skill sufficient?
- Are the activities chosen to evaluate the skill appropriate for determining whether a student has acquired the skill?
- Is the level of achievement expected of students in line with the industry’s expectations for new graduates?
- In each evaluation activity, is there a clear description of the tasks that will be performed for students to obtain the skill?
A reviewer will recommend that the applicant program be able to issue the badge if it is awarded at least 20 points and no skill is scored as less than 3.

B. The Final Decision

There are four possible **Final Decisions** for each application:

- **Application Accepted** = there is no doubt that the program provides the expected data science analytics skills. The program is granted permission to issue the badge for four years.
- **Application Accepted with Recommendations** = the program provides the data science analytics skills but needs some improvement. The program is granted permission to issue the badge for one year and needs to resubmit its application within a year.
- **Application Rejected with the possibility of revision** = there is not enough evidence provided in the application to show that the program adequately provides the required data science analytics training, but this could possibly be due to missing or incomplete information. The program is given the opportunity to resubmit a revised version of the application.
- **Application Rejected** = there is not enough evidence to show that the program adequately provides the required data science analytics training.

THE REVISION OF APPLICATIONS

If the reviewers of an application believe that the application would greatly benefit from the addition of missing information, they can offer the applicant the possibility of resubmitting a revised version of the application. If this is the case, the applicant will be offered two weeks to resubmit the application.

APPEALS

If after the receipt of the final decision of an application, the applicant believes that an error was committed in the review process, the applicant can submit an appeal. Appeals need to be submitted in the two weeks immediately following the receipt of a final decision.
CODE OF CONDUCT FOR BADGE ISSUING APPLICATION REVIEWERS*

* Document based on Annex 1 of the Model of Contract for EU Experts

A. ARTICLE 1 – PERFORMANCE OF THE REVIEW

1. The reviewer works independently, in a personal capacity and not on behalf of any organisation.
2. The reviewer must:
   ▪ evaluate each application in a confidential and fair way,
   ▪ assist the BDVe to the best of their abilities, professional skills, knowledge and apply the highest ethical and moral standards
   ▪ follow any instructions and time-schedules given by the BDVe and deliver consistently high-quality work.
3. The reviewer may not delegate another person to carry out the work or be replaced by any other person.
4. If an entity involved in an application approaches the reviewer during the evaluation of this application, s/he must immediately inform the BDVe.

B. ARTICLE 2 - OBLIGATIONS OF IMPARTIALITY

1. The reviewer must perform their work impartially. To this end, the reviewer is required to inform the BDVe of any conflicts of interest arising in the course of their work
2. Definition of the conflict of interest: for a given application, a **conflict of interest** exists if a reviewer:
   - was involved in the preparation of the application
   - stands to benefit directly or indirectly if the application is accepted or rejected
   - has a close family or personal relationship with any person representing an applicant
   - is a director, trustee or partner or is in any way involved in the management of an applicant
   - is employed or contracted by one of the applicants or any named subcontractors. However, the BDVe may decide to invite a reviewer who is employed or contracted by one of the applicants or any named subcontractors to take part in the review process, if the reviewer works in a different department/laboratory/institute from the one where the training is to be carried out, and if the constituent bodies operate with a high degree of autonomy
   - In the following situations, the BDVe will decide whether a conflict of interest exists, taking account of the objective circumstances, available information, and related risks.
     a. when a reviewer:
        i. was employed by the applicant in the last three years
        ii. is involved or in collaboration with the applicant, or had been so in the last three years
        iii. is in any other situation that could cast doubt on their ability to participate in the evaluation of the application impartially, or that could reasonably appear to do so in the eyes of an external third party.

3. Consequences of a situation of conflict of interest:
   - If a conflict of interest is reported by the reviewer or established by the BDVe, the reviewer must not evaluate the application concerned, or take part in any forum where the application is discussed.
   - If a conflict becomes apparent at any stage of the evaluation, the reviewer must immediately inform the BDVe. If a conflict is confirmed, the reviewer must stop evaluating the application concerned. Any comments and scores already given by the reviewer will be discounted. If necessary, the reviewer will be replaced. If it is revealed during an evaluation that a reviewer has knowingly concealed a conflict of interest, the reviewer will be immediately excluded. Any discussions in which s/he has participated will be declared null.
C. ARTICLE 3 - OBLIGATIONS OF CONFIDENTIALITY

1. The BDVe and the reviewer must treat confidentially any information and documents, in any form (i.e. paper or electronic), disclosed in writing or orally in relation to the performance of the review.

2. The reviewer undertakes to observe strict confidentiality in relation to their work. To this end, the reviewer:
   ▪ must not use information or documents provided by an applicant for any purpose other than fulfilling their obligations as part of the review process without the prior written approval of the BDVe
   ▪ must not disclose, directly or indirectly, information or documents relating to applications or applicants, without the prior written approval of the BDVe.

In particular, the reviewer:
   ▪ must not discuss any application with others, including other reviewers or BDVe staff not directly involved in evaluating the application, except during the formal discussion at the meetings moderated by or with the knowledge and approval of the BDVe
   ▪ must not disclose:
     a) any detail of the evaluation process and its outcomes or of any application submitted for evaluation for any purpose other than fulfilling their obligations to provide a review without the prior written approval of the BDVe
     b) their advice to the BDVe on any application to the applicants or to any other person (including colleagues, students, etc.)
     c) the names of other reviewers participating in the evaluation.
   ▪ must not communicate with applicants on any application:
     d) during the evaluation, except in panel hearings between reviewers and the applicants organised by the BDVe as part of the evaluation process;
     e) after the evaluation

3. If the applications are made available electronically to the reviewer who then works from their own or other suitable premises, s/he will be held personally responsible for maintaining the confidentiality of any documents or electronic files sent, and for returning, erasing or destroying all confidential documents or files upon completing the evaluation as instructed.

D. SIGNATURE

I, the undersigned, hereby agree to adhere to this code of conduct during the review of applications to issue BDV badges.
Name:

Institution:

Signature:

Date:
Appendix J: BDV Data Science Analytics Badge, Academic Level – Template for Reviewers

BDV

Data Science Analytics Badge
Academic Level

BADGE ISSUING APPLICATION
EVALUATION TEMPLATE
A. SUMMARY OF THE EVALUATION

Reviewer’s Name:

Evaluation of application number:

Recommendation: (please mark only one of the following)

- Application Accepted = there is no doubt that the program provides data science analytics skills. The program is granted permission to issue the badge for four years.
- Application Accepted with recommendations = the program provides the data science analytics skills but needs some improvements. The program is granted permission to issue the badge for one year and needs to resubmit its application within a year.
- Application Rejected with the possibility of revision = there is not enough evidence provided in the application to show that the program adequately provides the required data science analytics training, but this could possibly be due to missing or incomplete information. The program is given the opportunity to resubmit a revised version of the application.
- Application Rejected = there is not enough evidence to show that the program adequately provides the required data science analytics training.

Signature:

Date:

B. DETAILED EVALUATION

Please refer to the BDV Data Science Analytics Badge Academic Level GUIDELINES FOR REVIEWERS for general information regarding the review process and details regarding how to complete this section.

<table>
<thead>
<tr>
<th>DSA.1. Identify existing requirements to choose and execute the most appropriate data discovery techniques to solve a problem depending on the nature of the data and the goals to be achieved.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Score (0-5):</td>
<td></td>
</tr>
<tr>
<td>Comments to be given to the applicant justifying the score:</td>
<td></td>
</tr>
</tbody>
</table>
## DSA.2. Select the most appropriate techniques to understand and prepare data prior to modelling to deliver insights.

<table>
<thead>
<tr>
<th>Score (0-5):</th>
<th>Comments to be given to the applicant justifying the score:</th>
</tr>
</thead>
</table>

## DSA.3. Assess, adapt, and combine data sources to improve analytics.

<table>
<thead>
<tr>
<th>Score (0-5):</th>
<th>Comments to be given to the applicant justifying the score:</th>
</tr>
</thead>
</table>

## DSA.4. Use the most appropriate metrics to evaluate and validate results, proposing new metrics for new applications if required.

<p>| Score (0-5): | |
|-------------|</p>
<table>
<thead>
<tr>
<th>Comments to be given to the applicant justifying the score:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additional information not to be shared with the applicant:</td>
</tr>
</tbody>
</table>

**DSA.5. Design and evaluate analysis tools to discover new relations in order to improve decision-making.**

<table>
<thead>
<tr>
<th>Score (0-5):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments to be given to the applicant justifying the score:</td>
</tr>
<tr>
<td>Additional information not to be shared with the applicant:</td>
</tr>
</tbody>
</table>

**DSA.6. Use visualisation techniques to improve the presentation of the results of a data science project in any of its phases.**

<table>
<thead>
<tr>
<th>Score (0-5):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments to be given to the applicant justifying the score:</td>
</tr>
<tr>
<td>Additional information not to be shared with the applicant:</td>
</tr>
</tbody>
</table>
### Final Score of the badge

Sum of the scores for each of the individual skills:

<table>
<thead>
<tr>
<th>Overall Recommendations for Improvement (to be given to the applicant)</th>
</tr>
</thead>
</table>
Appendix K: BDV Data Science Analytics Badge, Academic Level – Guidelines for Reviewers

BDV

Data Science Analytics Badge
Academic Level

GUIDELINES FOR REVIEWERS
INTRODUCTION

Applicants to issue the Academic Level of the Data Science Analytics Badges there are required to submit applications according to the published guidelines (see the BDV Data Science Analytics Badge Academic Level INFORMATION AND GUIDELINES FOR APPLICANTS document available on the badge program’s web page: http://www.big-data-value.eu/skills/skills-recognition-program/). Each application will be reviewed, and a final decision will be made regarding whether each program can issue the badge. This document describes the role and activities of the reviewers of these applications. The skills required for the Data Analytics Badge are described in Table 42.

<table>
<thead>
<tr>
<th>Required skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSA.1. Identify existing requirements to choose and execute the most appropriate data discovery techniques to solve a problem depending on the nature of the data and the goals to be achieved.</td>
</tr>
<tr>
<td>DSA.2. Select the most appropriate techniques to understand and prepare data prior to modelling to deliver insights.</td>
</tr>
<tr>
<td>DSA.3. Assess, adapt, and combine data sources to improve analytics.</td>
</tr>
<tr>
<td>DSA.4. Use the most appropriate metrics to evaluate and validate results, proposing new metrics for new applications if required.</td>
</tr>
<tr>
<td>DSA.5. Design and evaluate analysis tools to discover new relations in order to improve decision-making.</td>
</tr>
<tr>
<td>DSA.6. Use visualisation techniques to improve the presentation of the results of a data science project in any of its phases.</td>
</tr>
</tbody>
</table>

Table 42: BDV Data Science Analytics Badge Skills

REQUIREMENTS FOR REVIEWERS

All reviewers must sign and adhere to the Code of Conduct.

EVALUATION GUIDELINES FOR THE ACADEMIC LEVEL OF THE DATA SCIENCE ANALYTICS BADGE

Only approved programmes can issue the Academic Level of the BDV Data Science Analytics Badge. To obtain this approval, a program must apply by providing evidence to show how they can ensure that their students acquire the corresponding skills.
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

To simplify the application process, the BDV Data Science Analytics Badge Academic Level Template for Issuing Badges is provided to applicants to help them collect specific details about how each of the skills required by the badge is acquired by the program’s students. The template also includes some general information about the program. Reviewers will be provided with the completed applications and the attached files containing the corresponding evidence.

Each application will be independently reviewed by two experts that will assess the submitted application as well as the evidence files. A final decision will be made if their recommendations coincide. If the two reviewers are not able to reach a consensus, a third reviewer will be asked to participate in the process.

A. General Guidelines

All reviews must be submitted in English.

If an application contains evidence files which cannot be understood by the reviewer for language reasons, then the reviewer can request that the BDV contact the applicant to provide a translation of the document.

Reviewers should be anonymous and should not provide any information in their reviews which could allow them to be identified.

Reviewers are given a deadline to complete their reviews, failure to submit the reviews in the given time period can result in their exclusion as reviewers.

A reviewer must notify the BDV and not submit a review if there could exist a conflict of interest as described by the Code of Conduct.

B. Criteria for Evaluation

Reviewers will complete the corresponding evaluation form, the Data Science Analytics Evaluation Template.

Based on the information presented, the acquisition of each of the skills of the Data Science Analytics Badge in an application will be scored as follows by each reviewer:

- 0 — Fail: The application fails to address the skill or the acquisition of the skill cannot be assessed due to missing or incomplete information. (‘Obvious clerical errors’ are not considered to result in incomplete information.)
- 1 — Poor: the skill is inadequately addressed, or there are serious inherent weaknesses.
- 2 — Fair: the application broadly addresses the skill, but there are significant weaknesses.
- 3 — Good: the application addresses the skill well but with a number of shortcomings.
- 4 — Very good: the application addresses the skill very well but with a small number of shortcomings.
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

- 5 — Excellent: the application successfully addresses all relevant aspects of the skill; any shortcomings are minor.

Such scores will be assigned based upon the following criteria:

- Is the average total number of hours that a student dedicates to the skill sufficient?
- Are the activities chosen to evaluate the skill appropriate for determining whether a student has acquired the skill?
- Is the level of achievement expected of students in line with the industry’s expectations for new graduates?
- In each evaluation activity, is there a clear description of the tasks that will be performed, for students to obtain the skill?

When scoring the previous criteria, the following must be taken into consideration:

- The application will be evaluated as submitted – evaluations are not based upon the possibility of the result of certain changes.
- Re-submissions will be evaluated as if it were the first time they are submitted, i.e., avoid comparison with previous versions of the application.
- Evaluators will not give a 5 if they identify any shortcomings or even worst, in the case of weaknesses.
- If no shortcomings are mentioned for a specific skill, then awarding 3 or less would not be appropriate.
- If evaluators use the term “weakness” in describing their opinion about a skill, then the score for that skill cannot be above 2.

Reviewers will also provide comments justifying or discussing the scores for each skill. Two kinds of comments can be provided:

a) Comments are given to the applicant. Those comments represent the rationale behind the score of each skill and will be shared with the applicant.
b) Additional information not shared with the applicants. Non-public comments that the reviewer might want to provide to the BDV regarding his/her score for each of the skills.

Each reviewer will provide a Final Score for the application by summing up the scores of each of the skills. The reviewer recommends that the applicant program be able to issue the badge if the program is awarded a total of at least 20 points and no skill is scored as less than 3. If the final score of an application is between 20 and 24, the reviewer has the option to recommend that the badge issuing period of the program be limited and that the program be required to resubmit another application to issue badges in the following year.

Finally, the reviewer will provide an Overall Recommendation for Improvement with a summary of his/her evaluation. This overall recommendation should be consistent with the final score provided and will be given to the applicant.
C. The Final Decision

Once reviewers have finished their reviews, they will submit their recommendations to the BDV. The BDV will process these recommendations and make a final decision regarding the application. There are four possible Final Decisions for each application:

- **Application Accepted**: there is no doubt that the program provides the data science analytics skills. The program is granted permission to issue the badge for four years.
- **Application Accepted with Recommendations**: the program provides the data science analytics skills but needs some improvements. The program is granted permission to issue the badge for one year and needs to resubmit its application within a year.
- **Application Rejected with the possibility of revision**: there is not enough evidence provided in the application to show that the program adequately provides the required data science analytics training, but this could possibly be due to missing or incomplete information. The program is given the opportunity to resubmit a revised version of the application.
- **Application Rejected**: there is not enough evidence to show that the program adequately provides the required data science analytics training.
Appendix L: FAQ

Who can apply to issue BDV badges?
Any EU based university-level educational institution can apply to issue BDV badges. Applicant programmes will most typically be at the MSc’s level, but undergraduate programmes can also apply. Approval to issue a badge is given to a particular degree program. Different programmes at the same institution must apply separately.

Can non-EU institutions apply to issue badges?
At the moment, only institutions within the EU can issue BDV badges.

Do separate programmes in multiple degrees (e.g., EIT Digital) need to apply separately?
Yes, as each individual program can evaluate skills in different ways, each program needs to apply separately.

Must all students of an approved badge issuing program be eligible to apply for a badge?
Badges are issued on a student by student basis; thus, the required skills for a badge can be acquired in elective courses or optional activities within a program. However, in this case only students who successfully acquire the required skills in those courses/activities can be issued a badge. Also, students do not need to be graduates of the programmes to earn a badge; they only need to demonstrate mastery of the required skills.

If optional courses/activities are included in the application, they should be listed in the application form according to the instructions provided in the next section.

Can a program charge additional fees to issue badges?
Programmes must provide badges to students at no additional cost as part of their standard student fees.

What language must be used in the applications?
Application templates must be submitted in English. When possible, evidence files should also be submitted in English, but when not possible, they can be submitted in the program’s native language.

Who can sign the application form?
The legal representative of each program must sign the application form.

What is an evidence?
Any evaluated activity which is included in a degree program can be considered to be an evidence. Typical examples include exam questions, laboratory exercises, sections of oral exams, and parts of final projects. Typical evidence files in these cases would include a copy of the exam’s questions, the instructions and description of the laboratory’s submitted work, the syllabus of an oral exam and the description of the requirements for the final projects. If a file includes evidence for multiple skills, please clearly specify which evidence corresponds with each skill by marking it in the
evidence file. Please do not submit examples of student work or any other information considered to be confidential by your institution.

**How many hours are considered to be the minimum required to acquire a skill?**
There is no simple answer to this question. The decision will be made by each reviewer based not just upon the number of hours which a student dedicates to a skill but also upon the skill itself and the level and quality of the work expected to acquire that skill.

**Can the same evidence file be used as evidence for more than one skill?**
Yes, different sections of the same evidence file can provide evidence for different skills. In this case, please clearly specify which parts of each evidence file corresponds with each skill.

**What information should a program send to evaluate its program?**
The application should include the completed Template for Badge Issuing Applications along with the different files containing the evidences for each skill.

**If an application is not accepted, is there a minimum wait period before reapplying?**
If an application is not accepted, the program can reapply in the following call for applications to issue badges.

**Can a new program apply to issue badges?**
To be eligible to issue badges, a program must have students who have already completed the program’s degree.

**Once approved, can a program issue badges to students whose skills were evaluated prior to the approval?**
Once approved, a program can issue badges to any student actively studying in their program regardless of when they demonstrated the acquisition of the required skills. Badges cannot be issued to students who completed or abandoned the program prior to the approval.

**Is there a limit to the number of badges that a program can issue each year?**
There is no limit; however, additional fees may be charged by the badge issuing platform when exceeding the platform’s limit.

**How does this badge program address concerns related to the GDPR?**
The BDV will not have access to any of the personal data of the students who earn badges. Only summary data, in which no student can be identified, will be available to the BDV. A third-party badge issuing, and validation service will be used as part of the program. Only this platform, which is compliant with the GDPR and will require a data usage agreement from each student, and the student’s institution will have access to the personal data contained in the issued badges.
Appendix M: Instructions for Issuing Badges

BDV
Data Science Analytics Badge
Academic Level

INSTRUCTIONS FOR ISSUING BADGES
A. INTRODUCTION

Congratulations! Your application to issue the academic level of the BDV Data Science Analytics Badge was accepted.

This document provides instructions for getting started with the badge issuing platform and for issuing badges.

After evaluating all available and Open Badge 2.0 certified badge platforms (https://site.imsglobal.org/certifications?refinementList%5Bstandards_lvlx%5D%5D=Open%20Badges) the BDV badge program selected the Open Badge Factory (https://openbadgefactory.com) as the most appropriate badge platform for issuing BDV Badges. It should be noted that the Open Badge Factory is hosted in Finland by a “professional operator providing services compliant with GDPR”, for more information regarding the GDPR compliance of the Open Badge Platform, please see https://openbadgefactory.com/?open=modal-faq.

B. What is an Open Badge?

To help you understand the badge issuing process, we would like to provide a brief introduction to Open Badges. Open Badges (https://openbadges.org/) are simply an image (the visual representation of the badge) which contains metadata. In the case of the BDV Data Science Badges, they are issued in the form of a PNG file containing JSON data compliant with version 2.0 of the Open Badge Standard (https://www.imsglobal.org/sites/default/files/Badges/OBv2p0Final/index.html).

When Open Badges are displayed in a platform designed to store Open Badges, the platform offers viewers the possibility to:

- Verify the authenticity of the badge (confirm that it was issued to the user’s email address by the issuing institution)
- View the metadata contained in the badge including:
  - A name, description and a narrative for the badge (see Appendix N for this information in the case of the Academic Level of the BDV Data Science Analytics Badge)
  - A link to an online evidence file.

This metadata provides details which potential employers and others can use to help them evaluate the skills acquired by the badge’s owner in your program.

C. Registering with the Open Badge Factory

You first need to open a free trial account with the Open Badge Factory (https://openbadgefactory.com/c/signup).
After logging in to your account, please ensure that your organisation’s name is “Searchable” (in the Admin Tools page, Organisation Details section, click in the box next to “Searchable”). Once complete you will see the following:

Send the name that you registered for your organisation and the country you used to bdv.badges@big-data-value.eu so that the BDV badge program can register you as a partner of the BDV badge program in the Open Badge Factory Platform. You will receive a reply to your email from the BDV badge program once your organisation is registered as a partner.

This same email will also contain a subscription code which will allow your organisation access to the Open Badge Factory for the time you have been approved to issue BDV badges.

To apply this subscription code to your account, click on the “Use subscription code” link in the Admin tools page, Subscription section as shown below.

You can now set up your account to issue the badge!
D. Setting up the Badge for Issue

Log into your Open Badge Factory account.

In the Network page, Badges shared with our section, click on the BDV Data Science Analytics – Academic Level badge and accept the offer to share the badge. Once complete you will see the following:

Now you need to define an application form for your badges. This form will be used by your students to apply for a badge. Please note that the following examples are given in English, but many other languages can be used.

In the Creator tools page, Badge applications section choose the badge and select “After review”. Now follow the steps:

1. The Instructions for issuers section is for the internal use of users in your organisation.
2. In the Form section, you can provide instructions and define fields for your badge’s application form. Note that by default the “Name” and “Email” fields are automatically included at the top of the form. We encourage including all the data provided in the application form as evidence inside the badge. We also recommend that you request evidence files (one or more) from applicants (for example code or PDF’s from particularly relevant assignments) as these will be made available to viewers of the badge and can provide useful and very detailed information to employers regarding the applicant’s educational experiences. However, remember that if you do this, all the information contained in the application will be publicly provided to anyone who has access to the badge so choose your fields carefully. This example of application fields:
Produces the following application form:
3. In the Messages section, you can modify the default acceptance and rejection email templates.

4. In the Reviewers section, you can define external reviewers of applications if you wish.

5. Nothing needs to be done in the Permissions section (unless you have other users in your organisation).

6. In the Settings section, you need to activate the application form. There you will find the URL of your application form and an embed code to include a link to include the form in web pages. You can also set the language of the application form. Lastly, you can decide whether to include a link to the application form as evidence in the badge (something which we do encourage).

Once complete and activated, your application form will appear as follows:
Now you can send the URL provided in the settings page to your students or include the embed code in web pages.

Please note that when students apply for a badge, they are voluntarily providing their personal information for the purpose of issuing the badge, and as they must agree with the terms and conditions of the Open Badge Factory, they are complying with all of the requirements of the GDPR.

E. Issuing Badges

As students apply for badges, their applications need to be reviewed. You can accept, reject or request a revision of an application.

In the Issuer tools page, Review Applications section, you can click on the badge to see your applications.

If you wish you can configure the platform to send daily emails when applications to issue badges are pending:
You can also use the Issue badge section of the Issuer tools page to issue badges in bulk, but this is discouraged as you must previously have permission to publish the user’s personal information in the Open Badge Platform and in this case, the metadata of the badges will not include an evidence file.

F. Receiving a Badge

After accepting a badge application, the applicant will receive an email from the Open Badge Factory with a link to redeem the badge. Upon following this link, the applicant can choose to: reject the badge, download the badge, or make a free account in the Open Badge Passport to display the badge online as shown below:
If the applicant chooses to download the badge, it can then be uploaded to any other badge hosting platform.

**G. Final Considerations**

By default, the Academic Level of the BDV Data Science Analytics Badge expires after 4 years.

At the basic level, an organisation can issue up to 5,000 badges per year using the Open Badge Platform. You can see your badge issuing quota in the Admin tools page, subscription section. The Open Badge Platform agrees to continue hosting all badges (and evidence files) issued for at least 2 years after the completion of an organisation’s contract and will do so indefinitely if the organisation continues to pay the annual fees of the Open Badge Platform.
Appendix N: Details of the BDV Data Science Analytics – Academic Level Badge

BDV DATA SCIENCE ANALYTICS – ACADEMIC LEVEL

<program name> (<program URL>)

<Program contact email>

A. Creator

Big Data Value eCosystem Project (http://www.big-data-value.eu)
bdv.badges@big-data-value.eu

Data Science Analytics is the use of techniques to gather, organise and process a collection of raw data to discover new and useful information and support the making of conclusions and decisions based upon that data. Earners of this badge have demonstrated their ability to apply Data Science Analytics techniques in an academic setting.

BDV Data Science Analytics Badge v1.0
- Academic Level

B. General Description

This badge is part of the BDV Data Science Skills Recognition Program (http://www.big-data-value.eu/skills/skills-recognition-program/) organised in a collaboration between the BDVe Project (http://www.big-data-value.eu/) and the BDVA Skills Task Force (http://www.bdva.eu/task-force-9).

The BDV Data Science Analytics badge serves to recognise an academic level of understanding in Data Analytics. Earners of this badge have shown that they can use appropriate techniques to discover new relations in a collection of data and thereby give insight into that data set to support decision making in an academic setting. This badge
includes skills related to Machine Learning, Data Mining, Statistical Methods and Algorithms.

This badge was issued by an educational programme which meets the educational standards established by the BDV Data Science Skills Recognition Programme.

C. Required Skills

To earn this badge, a student must acquire and demonstrate their competence in the following data science skills in an academic setting:

- DSA.1 Identify existing requirements to choose and execute the most appropriate data discovery techniques to solve a problem depending on the nature of the data and the goals to be achieved.
- DSA.2 Select the most appropriate techniques to understand and prepare data prior to modelling to deliver insights.
- DSA.3 Assess, adapt and combine data sources to improve analytics.
- DSA.4 Use the most appropriate metrics to evaluate and validate results, proposing new metrics for new applications if required.
- DSA.5 Design and evaluate analysis tools to discover new relations in order to improve decision-making.
- DSA.6 Use visualisation techniques to improve the presentation of the results of a data science project in any of its phases.

D. Credits


The description, requirements and types of badges used in the BDVe Data Science Skills Recognition Programme (http://www.big-data-value.eu/skills/skills-recognition-program/) were based upon the Edison EDSF (http://edison-project.eu/edison/edison-data-science-framework-edsf) and include changes suggested by the data science community.
Appendix O: Mobility Framework

A. Mobility Framework Survey Template

The Big Data Value project (BDVe), an H2020 Coordination and Support Action of the BDV PPP, is exploring the opportunity of developing a mobility framework to stimulate and promote mobility and to develop skills and competencies.

The purpose of the mobility program is to help increase the chance of solving a common technical challenge while also improving the mutual learning and transfer of knowledge among the participants.

This survey is coordinated by EIT Digital. The data collected will not be made public and will be used only for statistical purposes. For more information please contact Angelo Giuliana at Angelo.Giuliana@eitdigital.eu

*The name of the person filling the survey

*Name of the Company

*E-mail of the person filling in the survey

*In which BDV PPP project(s) is your company part of? at least 1 choice(s)

<table>
<thead>
<tr>
<th>BDV Project</th>
<th>Company</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEGIS</td>
<td>DataPitch</td>
<td>MyHealthMyData</td>
</tr>
<tr>
<td>BigDataGrapes</td>
<td>e-Sides</td>
<td>QROWD</td>
</tr>
<tr>
<td>BigDataOcean</td>
<td>E2Data</td>
<td>SLIPO</td>
</tr>
<tr>
<td>BigDataStack</td>
<td>eBusinessGraph</td>
<td>SODA</td>
</tr>
<tr>
<td>BigMedilytics</td>
<td>EDI</td>
<td>Special</td>
</tr>
<tr>
<td>BodyPass</td>
<td>EW-Shopp</td>
<td>TheyBuyForYou</td>
</tr>
<tr>
<td>Boost4.0</td>
<td>FANDANGO</td>
<td>Track&amp;Know</td>
</tr>
<tr>
<td>CLASS</td>
<td>FashionBrain</td>
<td>TT</td>
</tr>
<tr>
<td>Cross-CPP</td>
<td>i-BiDaas</td>
<td>TYPHON</td>
</tr>
<tr>
<td>DataBench</td>
<td>Icarus</td>
<td>other</td>
</tr>
</tbody>
</table>
1. Would you be interested in taking part in a pilot setting up a co-located team between projects representing partners from two or more PPPs to work on challenge to be solved?

☐ No

☐ Yes under which conditions if any:

2. Would you be interested in joining in EU Big Data Mobility Program if available?

☐ No

☐ Yes under which conditions if any:
**B. Mobility Framework Projects candidate**

<table>
<thead>
<tr>
<th>Project name</th>
<th>SINTEF</th>
<th>DemoKritos</th>
<th>UBI TECH Software AG Germany</th>
<th>ATB</th>
<th>University of Southampton</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeepHealth</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>EVOLVE</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>CYSELE</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>CloudButton</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>ELASTIC</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>EXA MODE</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>ExtremeEarth</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>INFORE</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>SmartDataLake</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>Data Market Services</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>MUSACROWN</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>Safe-DEED</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>BigDataOcean</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>SLIPo</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>AREGIS</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>euBusinessGraph</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>QROAD</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>FashionBrain</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>EW-Shopp</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>Data Pitch</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>EDI: European Data Incubator</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>Track and Know</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>Lynx</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>TheBuyerForYou</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>FANDANGO</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>DataBio</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>TT</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>BOOST4.0</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>BigDataStack</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>E2DATA</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>i-BiDataS</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>CLASS</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>Cross-CPP</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>Typhon</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>BDve</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>DataBench</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
<td>2020</td>
</tr>
<tr>
<td>SODA</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>SPECIAL</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>e-Sides</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>AI4EU</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>IASS</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
<td>2021</td>
</tr>
<tr>
<td>K-PLEX</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
<tr>
<td>Presto Cloud</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
<td>2019</td>
</tr>
</tbody>
</table>
T4.4: Functional Specifications for the Data Scientist Mobility Program – Market Portal for Internship

<table>
<thead>
<tr>
<th>Work Package</th>
<th>WP4 - SKILLS: Skills, Education, and Centers of Excellence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task</td>
<td>T4.4 - Data Scientist Mobility Programme</td>
</tr>
<tr>
<td>Editor(s):</td>
<td>Angelo Giuliana</td>
</tr>
<tr>
<td>Responsible Partner:</td>
<td>EIT Digital</td>
</tr>
<tr>
<td>Contributors</td>
<td>ATOS, INSIGHT, UPM, BDVA/ITI</td>
</tr>
<tr>
<td>Status-Version:</td>
<td>Draft – v1.0</td>
</tr>
<tr>
<td>Abstract:</td>
<td></td>
</tr>
</tbody>
</table>

Appendix P: Internship Portal

A. Functional Specification for the Internship Portal
1. Introduction

This document defines the minimum functional specifications for the Mobility Portal generated by Task 4.4 Mobility Program for Data Scientist to be published in the BDVe project web page (i.e. www.big-data-value.eu). One of the actions of Task 4.4 is to provide a simple portal to the BDV PPP partners (large companies and SME) and to the BDVA members for posting internship positions and mobility opportunities at industries targeting data scientists.

This section of the web should be also reached by other BDVe web pages (such as Big Data Landscape) to complement the information about the Big Data companies with additional information about their internship programs, if they have them.

2. Proposed Site Structure

We propose the inclusion for the Mobility Portal in the section “Ecosystem” of the BDVe web page (About us) as shown below:
Finally, BDVA is putting its marketing capabilities at the service of the whole Community to produce the European Big Data Value Forum, the Big Data flagship event of this and the coming years.

Figure 48: Home Page from pop-up and from People&Talent

There are two options reaching the Internship portal:

1.) Clicking from the Pop-Up menu under Ecosystem you will be redirected to the Internship portal,

2.) when clicking on People and Talent on the main site you will arrive at the ecosystem homepage as shown in the mock-up below where a new “button” should show the Internship opportunities:
From both options as described above it is possible to access the Internship portal with a minimum set of functionalities:

- Search for an internship position
- Post an internship position

3. Find an internship position

In this section the technical specifications for the set of features required for the Internship Mobility Portal section are defined.

3.1 Find an internship position
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

This web page has a search section (see Figure 50) where the entered keyword comes from choosing in a drop down menu, where the each interested person can search for a specific country or for a company name listed.

**Figure 50: Internship main page**

If you enter a keyword with the search criteria

As result a list of the companies that match the search criteria appears in the page as show in Figure 51 together with further filter on the left side.

The filter criteria are based on the following categories:

- **DEGREE** (Bachelor, Master, PhD)
- Big Data Categories (see Appendix)
- Duration (in months max 12 months)
- Compensation (Financial Compensation – No Financial Compensation)
- Commitment (Full-Time, Part-Time)
- Languages (all)
The shown criteria depend on the match between the actual date of the search and the deadline date, if the deadline is passed the opportunity is not shown. Clicking on the CEFRIEL “Read More” button, a detailed web page with generic information about the position and the location and a direct link to the Company Web page appears – see Figure 52.|Error! No se encuentra el origen de la referencia.|

A generic description of the internship position appears under the tab “ABOUT”, while under “Skills and Competencies” there is what is expected from the candidate in terms of expertise – see Figure 52.|Error! No se encuentra el origen de la referencia.|
Figure 52: Internship position details
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

When clicking on the below button “Internship Information” the student/researcher/professional will be redirected to the Company Web Page (for ex. https://www.cefriel.com/category/internships/).

If at the main page (Figure 50) no search keyword is entered, but just clicked view complete list, a similar page as shown in the Figure 51 appears with all the Companies in alphabetic order independent from the Country.

3.1.1 Post an Internship Position

This entry in the Mobility Portal is dedicated to the Industries and SMEs that want to offer internship opportunities to students/researcher/professionals that are interested in a working experience.

The section is dedicated only to the internship opportunities related the Big Data topics, which is one of the areas with the major lack of skills.

This section will provide all the needed functionalities to permit the employers (Industries/SME’s) to post their job request in the Data Scientist landscape.

The upload will be a batch process, offline, done by an administrator of the portal that will fill the Internship Portal Datasets with the updated information.

The HR department of BDVA/BDV PPP members/partners that wants to upload their Big Data internship positions will have to fill in a template form (could be an excel with some contents as data list fields) with the minimum required information as below described:

<table>
<thead>
<tr>
<th>Content</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company Name</td>
<td></td>
</tr>
<tr>
<td>Company Address</td>
<td></td>
</tr>
<tr>
<td>Company Contact mail</td>
<td></td>
</tr>
<tr>
<td>Company Contact number</td>
<td></td>
</tr>
<tr>
<td>Company Market Category</td>
<td>From a service Table</td>
</tr>
<tr>
<td>Internship Position</td>
<td>Name of the open position</td>
</tr>
<tr>
<td>Duration</td>
<td>From a service table (1-12 months)</td>
</tr>
<tr>
<td>Post Date</td>
<td>dd/mm/yy</td>
</tr>
<tr>
<td>Deadline</td>
<td>dd/mm/yy</td>
</tr>
<tr>
<td>Degree required</td>
<td>From a service table</td>
</tr>
<tr>
<td>Big Data category *</td>
<td>From a predefined list as in Appendix</td>
</tr>
<tr>
<td>Compensation</td>
<td>Yes/No</td>
</tr>
<tr>
<td>Commitment</td>
<td>Full-Time; Part-time</td>
</tr>
</tbody>
</table>
### D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

<table>
<thead>
<tr>
<th>Content</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>From a service table</td>
</tr>
<tr>
<td>Company country</td>
<td>Link to corresponding Country table</td>
</tr>
<tr>
<td>Internship position short description</td>
<td>Free text (max 5 rows)</td>
</tr>
<tr>
<td>Skill &amp; Competencies required</td>
<td>Free text (max 5 rows)</td>
</tr>
<tr>
<td>Link to internship company web page</td>
<td>URL</td>
</tr>
</tbody>
</table>

For Example, for CEFRIEL the table will have the following information:

<table>
<thead>
<tr>
<th>Content</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company Name</td>
<td>CEFRIEL Politecnico di Milano</td>
</tr>
<tr>
<td>Company Address</td>
<td>Viale Sarca 226 20126 Milano</td>
</tr>
<tr>
<td>Company Contact mail</td>
<td><a href="mailto:info@cefriel.com">info@cefriel.com</a></td>
</tr>
<tr>
<td>Company Contact number</td>
<td>(+39) 02 239541</td>
</tr>
<tr>
<td>Company Market Category</td>
<td></td>
</tr>
<tr>
<td>Internship Position</td>
<td>Big Data and Analytics Cloud Consultant</td>
</tr>
<tr>
<td>Duration</td>
<td>3 months</td>
</tr>
<tr>
<td>Post Date</td>
<td>25/05/18</td>
</tr>
<tr>
<td>Deadline</td>
<td>07/09/18</td>
</tr>
<tr>
<td>Degree required</td>
<td>PhD</td>
</tr>
<tr>
<td>Big Data category of the position</td>
<td>Analytics, Cloud Computing</td>
</tr>
<tr>
<td>Compensation</td>
<td>Yes</td>
</tr>
<tr>
<td>Commitment</td>
<td>Full-Time</td>
</tr>
<tr>
<td>Language</td>
<td>Italian</td>
</tr>
<tr>
<td>Owner country</td>
<td>Italy</td>
</tr>
<tr>
<td>Internship position short description</td>
<td>Analytic consultants work with project managers, businesses owners,</td>
</tr>
<tr>
<td></td>
<td>analyst teams and clients, building ETL processes, data warehouses,</td>
</tr>
<tr>
<td></td>
<td>database prototypes, reports and implement predictive algorithms to</td>
</tr>
<tr>
<td></td>
<td>deliver the customer requirements. They also document code, provide</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
progress reports, and perform code review and peer feedback.

Analytic consultants job duties include developing reports, analytic applications, SQL databases and integrating various business applications.

The internship position will focus on developing the soft and hard skills to become an analytics consultant as well as give an opportunity to work in live projects together with other more experienced team members.

Skill & Competencies required

Strong communication, problem solving and presentation skills
Fluent English
Learning capacity and the will to acquire new skills
Ready to travel for assignments across EMEA or worldwide up to 25% of working time
Experience / Knowledge in software projects
Knowledge of programming and/or database administration
Operating systems and Network Administration is desirable

Link to the internship company web page

https://www.cefriel.com/category/internships/

Some of the supporting table are pre-loaded or pointed to tables already used for other functionalities (for ex. Language, country, degree used in EduHub or Vertical Market in Marketplace), the same value are provided to the companies when filling the template form (data list fields); other contents will be taken from the template form provided to the companies and will upload the main dataset “Internship Position”.

3.1.2 Database

The main datasets that derive from the concept of the Internship Portal is the following:

- Internship Position
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

Supporting information:
- Language table
- Degree
- Big Data Skills
- Country table
- Vertical Market Category

The main dataset is the catalogue of internship positions, with a few supporting tables like the ones mentioned above.

**Internship Position Table**

<table>
<thead>
<tr>
<th>Content</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company Name</td>
<td>Text</td>
</tr>
<tr>
<td>Company Address</td>
<td>Text</td>
</tr>
<tr>
<td>Company Contact mail</td>
<td>Text</td>
</tr>
<tr>
<td>Company Contact number</td>
<td>Text</td>
</tr>
<tr>
<td>Company Market Category</td>
<td>Link to Vertical Market Category as defined in Marketplace</td>
</tr>
<tr>
<td>Company Picture</td>
<td>Company Logo Picture attached the general Info</td>
</tr>
<tr>
<td>Internship Position</td>
<td>Text</td>
</tr>
<tr>
<td>Duration</td>
<td>Number</td>
</tr>
<tr>
<td>Post Date</td>
<td>Date</td>
</tr>
<tr>
<td>Deadline</td>
<td>Date</td>
</tr>
<tr>
<td>Degree required</td>
<td>Link to Degree table</td>
</tr>
<tr>
<td>Big Data category *</td>
<td>Link to Big Data Skills Table</td>
</tr>
<tr>
<td>Compensation</td>
<td>Select one of the values:</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Commitment</td>
<td>Select one of the Value:</td>
</tr>
<tr>
<td></td>
<td>Full-Time</td>
</tr>
<tr>
<td></td>
<td>Part-time</td>
</tr>
<tr>
<td>Language</td>
<td>Link to language table see EduHub</td>
</tr>
<tr>
<td>Company country</td>
<td>Link to Country table see EduHub</td>
</tr>
<tr>
<td>Internship position short description</td>
<td>Text</td>
</tr>
</tbody>
</table>
4. Non-functional Requirements

The BDVe Mobility Portal has the following non-functional requirements:
Usability of this section across multiple devices is the most important non-functional requirement. The user interface for providing previous described functions should look and feel consistent across desktops, laptop, tablets, and smartphones.

Performance is also critical in terms of transitions between web pages, loading of search results and their associated items.

Availability of this section should be guaranteed in similar terms as is the availability of the web page itself.

5. Appendix

5.1 Appendix 1 Big Data Skills List

In this paragraph there is a list of the most common required skills related to Big Data. Some other skills will increment and update this list, coming from the result of the work done in BDVA TF9 and BDVe T4.2 T4.3.

Big Data Skills List

A - C

- Adaptability
- Advanced Microsoft Excel
- Advanced SQL
- Amazon Web Services (AWS)
- Analytical
- Analytical Solutions
- Analytics
- Applying Knowledge of the Software Development Lifecycle
- Assessing the Data Needs of Internal Stakeholders or Clients
- Attention to Detail
- Big Data Solutions
- Big Data Strategy
- Big Data Technologies
- Coaching Executives Regarding the Impact of Big Data on Strategic Plans
- Cloud Computing
- Collaboration
- Communication
- Conducting Statistical Analyses
- Continual Learning
- Conveying Technical Information to Non-Technical Audiences
- Creating Visualisations for Data Systems
- Creative Thinking
- Critical Thinking

D - L
- Data Access Systems
- Data Architecture
- Data Flow
- Data Management
- Data Mining
- Data Modelling
- Data Profiling
- Data Sets
- Data Wrangling
- Decision Making Regarding Complex Technical Designs
- Designing Data Warehouses
- Drafting Proposals
- Drawing Consensus
- Estimating Costs for Projects
- Facilitating Group Discussion
- Hadoop
- Handling Criticism Non-Defensively
- Implementing Data Warehouse Systems
- Interpreting Data Output
- Large Data Sets
- Leadership
- Leading Cross-Functional Groups

M - R
- Manipulating Relational Databases
- Matlab
- Multitasking
- NoSQL
- Organizational
- Persuading Colleagues to Adopt Preferred Big Data Systems and Strategies
- PowerPoint
- Presentation to Groups
- Problem Solving
- Programming with Java
- Project Management
- Python
- Quantitative
- R
- Research

S - Z
- SAS
- Shell Scripting
- Spark
- SPSS
- Stress Management
D4.3: Skills, Education, and Centers of Excellence Period 2 Report M36

- Structuring Cloud Storage Solutions for Big Data
- Taking Initiative
- Teamwork
- Time Management
- Tracking Trends and Emerging Developments in Big Data
- Translating Data Analysis into Key Insights
- Verbal Communication
- Visualisations
- Working Independently
- Writing Reports with Data Findings