ICT2018 VIENNA
NETWORKING SESSIONS HIGHLIGHTS

DATA DEMOCRATISATION: EMPOWERING THE CITIZENS IN THE DIGITAL TRANSFORMATION
PERSONAL AND MACHINE-GENERATED DATA: IMPACT ON PRIVACY AND SECURITY

Editors:

Jean-Christophe Pazzaglia and Daniel Alonso
BDVe Project
## TABLE OF CONTENTS

- **Acknowledgments** ........................................................................................................... 4
- **DISCLAIMERS** .................................................................................................................. 4
- **INTRODUCTION** .............................................................................................................. 5
- **Format of the sessions** ..................................................................................................... 6
- **Data Democratization Introduction** .................................................................................. 7
- **What DOES data democratization look like - what are potential benefits and what potential risks do you see?** .................................................................................................................... 9
  - BDVE question Moderator: Tjerk Timan, on behalf of BDVe ........................................... 9
    - Summary: data democratization in 4 clusters; concepts, risks, benefits and solutions ........ 9
    - Raw data .......................................................................................................................... 10
- **What do you see as the main economic, democratic, and societal benefit of data democratization?** ............................................................................................................................................ 13
  - MyData Question Moderator: Michael Shea, The Dingle Group (on behalf of MyData Global) ...... 13
    - Summary .......................................................................................................................... 13
    - Raw Data ....................................................................................................................... 14
    - About MyData – https://mydata.org.............................................................................. 16
- **What initiatives should take governments to explain what data democratization is about?** .......... 17
  - BDVA question Moderator: Daniel Alonso (BDVA TF4 & TF1.SG1 lead, BDVe project) .......... 17
    - Context: .......................................................................................................................... 17
    - Summary .......................................................................................................................... 18
    - Raw data .......................................................................................................................... 19
- **WHICH INFRASTRUCTURE, INTERMEDIARY, TOOLS OR SERVICES SHOULD WE PROVIDE TO BOOST DATA DRIVEN INNOVATORS?** ............................................................................................................ 21
  - European network of living labs (ENOLL) Question Moderator: ZSUZSANNA BODI .......... 21
    - Summary .......................................................................................................................... 21
    - Raw DATA ....................................................................................................................... 22
    - About ENoLL – https://enoll.org/ ............................................................................... 24
What Skills should be taught to enable Citizens to Understand, Decode and Appropriate themselves the data? .................................................................................................................................................................................. 25

EWSHOPP Question Moderator: Fernando Perales, JOT INTERNET MEDIA (On behalf of EWSHOPP project) .................................................................................................................................................................................. 25

Summary ............................................................................................................................................................................................................................................. 25
Raw data ........................................................................................................................................................................................................................................... 26
About EWSHOPP - https://www.ew-shopp.eu/ ........................................................................................................................................................................... 28

Which tools or services can be developed to help citizens to understand data and potentially detect biased or malignant usage? ........................................................................................................................................................................... 29

BigDataStack Question Moderator: Dimosthenis Kyriazis (BigDataStack project) ......................................................................................................................................................................... 29

Discussion summary .................................................................................................................................................................................................................. 29
raw data ............................................................................................................................................................................................................................................ 30
About BigDataStack – https://bigdatastack.eu/ ..................................................................................................................................................................... 32

Personal and Machine generated DATA: Introduction .................................................................................................................................................. 33

How AI may transform the game (of cybersecurity) in terms of new threats but also new forms of protection ........................................................................................................................................................................................................................................................................................................... 36

BDVe Question Moderator: Tjerk Timan, BDVe ........................................................................................................................................................................................................................................... 36

Summary ............................................................................................................................................................................................................................................. 36
Raw data ........................................................................................................................................................................................................................................... 37

How can we raise citizen awareness on potential risks about the data collected about you or your machines? ........................................................................................................................................................................................................................................................................................................... 40

ECSO Question Moderator: Geraud Canet ........................................................................................................................................................................................................................................... 40

Summary ............................................................................................................................................................................................................................................. 40
Raw Content .................................................................................................................................................................................................................................... 40
About ECSO – https://ecs-org.eu/ .................................................................................................................................................................................................................................... 42
ACKNOWLEDGMENTS

The organizers want to thank all moderators and participants that contribute to the fun and the success of the sessions.

DISCLAIMERS

This disclaimer informs readers that the views, thoughts, and opinions expressed in the text belong solely to the participants, and not necessarily to the moderators, moderators and authors employers, participants employers, organizations, committees or other groups or individuals.
INTRODUCTION

The BDVe project, in collaboration with BDVA, MyData and the European Network of LivingLabs, and PPP projects EW-Shopp and BigDataStack in one hand, and with BDVA and ECSO, and PPP projects MyHealth-MyData and Special in the other hand, organised two networking sessions:

- Data democratisation: empowering the citizens in the digital transformation
- Personal and machine-generated data: impact on privacy and security

This report provides some insights on the methodology and the result of the brainstorming.

If you liked the sessions and want to join the BDV PPP crowd, we are heading to north in 2019!


Book your agendas for Riga and Helsinki, and stay tuned ... more details coming soon!
FORMAT OF THE SESSIONS

The Carousel Brainstorm provides an opportunity to use the group’s collective prior knowledge to further individual understanding. It is an active, users-centred method for generating and sharing large amounts of data. To implement efficiently the method, 6 groups were pre-determined and a coloured sticker identified the group membership, each group answered 3 questions to cope with time constraints. Each station was equipped with a table, the usual set of coloured sticky notes and the topic of the station. After the first iteration, the coach explained the different clusters identified and the ideas developed during the previous iterations. Then a 3 minutes silent brainstorm was used to collect new ideas that complemented the initial ideas.
During ICT2018, BDVe project, with BDVA, MyData and the European Network of LivingLabs, and PPP projects EW-Shopp and BigDataStack, organized a networking session around the Data democratisation: empowering the citizens in the digital transformation. The brainstorm was organised around 6 questions and followed the carousel format.
Data democratisation: empowering the citizens in the digital transformation

Set #1
- What does data democratization mean for you and for the society? What are their benefits in your daily life? — BDVe
- What skills should be taught to enable the citizen to understand, decode and appropriate themselves the data? — EW-Shopp
- What do you see as the main economic, democratic and societal benefits of data democratization? — MyData

Set #2
- What initiatives should take governments to explain what data democratization is about? — BDVA
- Which tools or services can be developed to help citizen to understand data and potentially detect biased or malignant usage? — BigDataStack
- Which infrastructure, intermediary, tools or services should we provide to boost data driven innovators? — ENoLL

For all questions, success or failure stories are welcome.
WHAT DOES DATA DEMOCRATIZATION LOOK LIKE - WHAT ARE POTENTIAL BENEFITS AND WHAT POTENTIAL RISKS DO YOU SEE?

BDVE QUESTION
MODERATOR: TJERK TIMAN, ON BEHALF OF BDVE

SUMMARY: DATA DEMOCRATIZATION IN 4 CLUSTERS; CONCEPTS, RISKS, BENEFITS AND SOLUTIONS

Whereas data democratization sounds like logical and positive idea, in this session we wanted to explore what data democratization can mean according to the workshop participants. In a first round we explored concepts related to data democratization, and what it can mean. Here, terms such as data sovereignty, data ownership and data access came about. Furthermore, overarching questions were asked such as the question as to what would constitute data driven society and if people would actually recognize the premise of being in one. Other points raised evolved around the role and meaning of data, what openness of data actually means and who or what should be responsible for governance of data democracy.

In the second round, we focused on potential risks and benefits from data democratization. In terms of benefits, the two main categories that emerge were societal and commercial benefits. Societal benefits included an increase of understanding the role and meaning of data, more informed decision-making and more control, and sovereignty over one’s data. On the commercial side, benefits of a more democratic data-landscape were seen for startups, who can more easily explore and evidence new ideas or services by having better access to different types of data. Moreover, opening up data in general would lead to
new business ideas, better transparency and to a more efficient way of finding partners and to collaborate. It will also enable to better use and manage resources.

In terms of **risks**, the groups identified risks for individuals, and what effects radical openness of data can have on autonomy, decision making and privacy. On a societal level, risks were seen in data being used wrongly or falling into the wrong hands but also by the power given by potential of misuse of all that open data (e.g. social scoring) – how to build a governance of trust surrounding data?

Some directions for **solutions** were provided, such as more education and awareness and (the need for) better and easier to use tools for citizens to work with data, and to build confidence and trust in using their own data.

---

**RAW DATA**

**Concepts (related to data democratization)**

- data as a weapon (for good and for bad)
- transparency
- database-openness (no more paywalls)
- data commons
- ministry of data
- one-to-many – network effects for all
- open data
- data pool of anonymized and consented personal data
- equal terms
- (data) ownership
- data pluralism
- (citizen) empowerment
- understanding data value
- level of interest (do people want to know about this?)
- what is the difference with the now: what would it change compared to now (is the ‘now’ a data driven society?)
- (increase of) engagement
- data archives and governance
- private data provisioning
- data sovereignty

**Benefits (of data democratization and opening things up)**

- user sovereignty
- better filtering and value discovery
- data fairness and data education
- more insight into data streams
- legibility and accessibility
- more data inclusion
- credibility (data-based leads to better evidence-based)
- deciding on one’s data – what happens with it
- build knowledge and new ideas
- Collaboration and new ideas/ innovation through sharing of data
- better environment for startups – easier to start and test new ideas
- development of new services
- from producer and consumer to prosumer
- demand-focused services
- fact-based decision - making
- space for new types of platforms to work with data
- new age of data economy
- metadata search
- generate income from personal data sharing
- better resource management
- to generally increase understanding the environment and world around us

**Risks (of data democratization and opening things up)**

- loss of anonymity
- data secrets no longer kept – radical transparency
- inefficiency – too much data to handle
- increase of populism through data
- data laziness (loss of filtering and selection)
- data overload
- risk of losing control over one’s data – not knowing where what is
- loss of privacy
- data divide
- data-based biased decisions – hidden in the sheer amount of data (if n=large, truth++)
- wrong interpretations
- court evidence through data (data-based decisions in court)
- risk of exploitation – significant data in the wrong hands – misuse of data
- social scoring and ‘good behaviour’ measuring
- (data) vulnerability
- learning things you don’t want to know (about yourself or your neighbours) – risk of over-exposure

**Solutions**

- Connect different data sources
- Trust – to share data without regret
- Easier tools to work with data
- International data spaces (IDS)
The mission of BDVe is to support the Big Data Value PPP in realizing a vibrant data-driven EU economy or said in other words, BDVe will support the implementation of the PPP to be a SUCCESS. Behind that mission, there are multiple goals to achieve, which should be taken into full consideration when defining the directions of the PPP. Some of the most challenging ones are: (1) achieving a more competitive landscape of European Big Data providers, leading to bigger market share; (2) creating the context for a more competitive EU industry (transport, manufacturing, public sector, agrifood, media, energy...) in the advent of a data-driven revolution where many traditional players will have to transform their processes and re-think their business if they want to remain competitive—or in some cases, just to survive; (3) ensuring the sustainability of the investments and actions triggered by the PPP.

BDVe has broken down those high-level goals into 7 major priorities for the project:

- Being accurately informed about most important facts in Big Data so that we have a solid basis to support the decision-making process in the PPP
- Supporting the implementation of the Big Data PPP from an operational point of view
- Developing a vibrant community around the PPP
- Supporting the development of a European network of infrastructures and centers of excellence around Big Data
- Setting-up a professional Communications strategy
- Setting up a framework that supports the acceleration of data-driven businesses, and
- Ensuring the sustainability of the investments and actions triggered by the PPP.

The BDVe consortium includes a set of partners that have shown commitment and dedication to the success of the PPP for several years. They have already invested and they have committed to invest along the coming years. We believe that this CSA cannot be a neutral action that offers operational support without further commitment.
WHAT DO YOU SEE AS THE MAIN ECONOMIC, DEMOCRATIC, AND SOCIETAL BENEFIT OF DATA DEMOCRATIZATION?

MYDATA QUESTION
MODERATOR: MICHAEL SHEA, THE DINGLE GROUP (ON BEHALF OF MYDATA GLOBAL)

SUMMARY

As expected the results of the brainstorming sessions cluster around the three dimensions of the question: economic, governmental and societal (citizen). An interesting insight was the ‘leveling of the playing field’; this concept plays into all three areas and has the potential for strengthening all dimensions.
By shifting the 'center of data' back to the citizen it has the potential of modifying the power dynamic; elevating the power position of the citizen bringing it more in line with the power of the business or government dimensions.

Another key insight is how democratization of data can be opportunity to strengthen democratic institutions and processes. As citizens understand how and where their information is being used, and the outcomes that are possible, the result is increased engagement with governmental institutions. Increased engagement can unlock empowerment and increase the feeling that governmental bodies are truly responsive to the needs of their citizens.

Finally, there was a single concept presented around data stewards that opens a very interesting set of questions on who do citizens feel comfortable and confident in to act on their behalf as a Data Steward? There is not a clear answer on this at this time, but the comment made positions/implies public sector institutions as data stewards, and not the private sector.

One objective of question framed by MyData Global, was to assess the perception of individuals who were/are outside the current MyData community on the value of data democratization. As with any passionate organization, getting consistent, objective outside comments is very important to ensure efforts are focused correctly. The thoughts expressed by the three groups, validate the MyData belief that data democratization is something of interest and is desired. The exact road to get there is still a ‘work in progress’, and further validation points will be required.

---

**Inclusion**
- Better knowledge and awareness
- Collectivity becomes a *unicum*
- Starting confrontation

**Citizen Engagement**
- User POV: personalization of services
- Problem solving
- More collaborative innovation for the end user
- Engagement
- Cost saving
- Grassroot solutions and creativity
- Citizen participation in policy making
- More individual power/control for each individual
- Citizens are informed
- Empowered Citizens - I check, I control, I analyze, I understand
- More active citizens
- Able to Identify “Fake News” and Mis/Disinformation
- Public participation
Data Control Inversion
- Flip the pyramid of power to the “citizen consumer”

Privacy
- More time, less burden around privacy

Information Leveled Playing Field
- Improve data quality (once only principle, data exists only in one location)
- Information
- Everyone has the same information
- Knowledge of the Complex social phenomena

Better Business Support Environment
- Economic: creating a link between clusters of industry that only in appearance are far from each other.
- Businesses save time collecting data - now focus time on using the data versus collecting it.
- Build value on top of data
- Gaining money/benefits for sharing personal data $/Euros
- Access to different types of resources
- Win - win strategy

Business Information
- Support new business innovation
- Grow sharing economy

Government Services / Democracy
- Responsibilities of Companies and Governments
- Better quality of public services
- Trust in democracy and public administration in general
- Reinforce trust in public institutions as data stewards
- Accountability of government
- Empowerment and Education
- Data driven (fact based?) decision making
- Accountability of companies, governments, etc. (political, environmental, societal)

Improved Services
- Cross societal inventions, services, etc.
- Boost to openness of science
- Improved medical treatment (open access to health data)
ABOUT MYDATA – HTTPS://MYDATA.ORG

MyData Global's mission is to empower individuals by improving their right to self-determination regarding their personal data. The human-centric paradigm is aimed at a fair, sustainable, and prosperous digital society, where the sharing of personal data is based on trust as well as balanced and fair relationship between individuals and organisations.

MYDATA DECLARATION

We are entrepreneurs, activists, academics, listed corporations, public agencies, and developers. For years, we’ve been using different words for what we do – MyData, Self Data, VRM (Vendor Relationship Management), Internet of Me, PIMS (Personal Information Management Services) etc, while sharing a common goal: **to empower individuals with their personal data, thus helping them and their communities develop knowledge, make informed decisions, and interact more consciously and efficiently with each other as well as with organisations.**

Together, in recent years, we have formed a network whose participants share experience, develop common projects, meet at the MyData conference, and take part in collective endeavours towards a human-centric approach to personal data.

It is now time to take this work out in the world and prove its potential impact on individuals, society, and the economy. Today, we believe it is time to publicly assert the values that drive us – and call on those who share those values to act upon them. Join us in reversing the paradigm of personal data. Join us in creating the MyData movement.
WHAT INITIATIVES SHOULD TAKE GOVERNMENTS TO EXPLAIN WHAT DATA DEMOCRATIZATION IS ABOUT?

BDVA QUESTION
MODERATOR: DANIEL ALONSO (BDVA TF4 & TF1.SG1 LEAD, BDVE PROJECT)

CONTEXT:
The future of Europe’s economic prosperity is dependent on its citizens and organisations to successfully establish and maintain a leadership position in the global data economy. In this way, the Big Data Value Association considers crucial to have a society fully informed about the value they can obtain from their data, not only in an economic way but also to improve services and, in general, to increase their welfare, so they can actively collaborate in the development of the data economy.

BDVA has included the citizen as the main beneficiary of the different pillars that the association has identified crucial in its position paper towards the Framework Programme 9 (http://www.bdva.eu/sites/default/files/BDVA%20position%20to%20Fp9_v1.pdf). Therefore, the first step for the society to get value from the data economy is to raise awareness about what this value is, how the citizen can benefit from it and pros and cons of this data democratization.
SUMMARY

The different ideas and answers provided touched several aspects, but they were around three main pillars: education and training, communication and dissemination, and direct and proactive involvement of the people in different activities. Education and training should cover from the very beginning (primary school) until courses aimed at professionals in different sectors. Communication and dissemination should use the traditional channels, but also more imaginative ways to reach all ranges of society (not digital, young people, etc ...). Both education and dissemination messages should include use case and success stories coming from different sources, as well as make use of demonstrators that clearly show the value of data. All activities are aimed to foster the engagement of the citizen in the whole process (real democratization), from the co-design of tools, the creation of apps and platforms, the exchange via marketplaces, the use, etc ..., looking for proactive citizens that demand more and more information from the relevant stakeholders in the process of their empowerment in the digital transformation.
RAW DATA

**Government and other stakeholders direct actions**
- Showing flow of public data
- Training for civil servants
- Incorporate data democratization in the role of governors
- Work with industry associations
- Trust brokerage
- Create environment for project creators
- Scale up local initiatives (smart cities)
- Good examples from beneficiaries
- Incorporate data democratization in the SRIAs
- Data producers inform users
- Open innovation platforms for new offerings

**Use case stories**
- Examples of benefits from data
- Demonstrations
- Citizen’s use cases demonstrating added value
- Simple use case stories for general public

**Dissemination (channels, mass media, others ....) and information campaigns**
- Create a website with all the necessary information
- Reality / fiction TV-soap for young people
- Create a story with an identifiable character: TV / radio / newspaper, in a real physical place
- Radio spots to take away fear
- Social network around data democratization
- Creating small groups of people with f2f meetings
- TV programs
- Information campaigns about the role of data and right to have access
- Public debates with relevant people and influencers
- Demonstrating how data are at the basis of all the services people use
- Campaigns with very concrete topics
- Data models and metadata must be understood
- Involve smart communities

**Education**
- Communicating the idea to young people (schools, universities, ...)
- Different learning activities from very simple to more complicated
- Be part of education programs at schools
- Open courses
• Data demo in school curricula
• Improve their presence in public spaces (new ways of interaction)
• Different projects with schools
• Education programs to reduce citizens fair

**Foster people active engagement**

• Civil society engagement
• Citizens should ask to be informed (proactive)
• Co-design / engagement sessions
• Empowering people in the co-design of digital tools
• Involving people in creating services / solutions which use data
• Apps for easy access (i.e.: smart cities)
• Creation of e-platforms and apps
• Marketplaces and exchanges

**ABOUT BDVA - HTTP://WWW.BDVA.EU/**

The **Big Data Value Association** (BDVA) is an industry-driven international not–for-profit organisation with 200 members all over Europe and a well-balanced composition of large, small, and medium-sized industries as well as research and user organizations. BDVA is the private counterpart to the EU Commission to implement the **Big Data Value PPP program**. BDVA and the Big Data Value PPP pursue a common shared vision of positioning Europe as the world leader in the creation of Big Data Value.

The **mission of the BDVA** is to **develop the Innovation Ecosystem** that will enable the **data and AI-driven digital transformation in Europe** delivering maximum economic and societal benefit, and, achieving and sustaining Europe’s leadership on **Big Data Value creation** and **Artificial Intelligence**.

BDVA is **open to new members** to further enrich the data value ecosystem and play an active role. These include Data Users, Data Providers, Data Technology Providers and Researchers. BDVA enables existing regional multi-partner cooperation, to collaborate at European level through the provision of tools and knowhow to support the co-creation, development and experimentation of pan-European data-driven applications and services, and know-how exchange.
WHICH INFRASTRUCTURE, INTERMEDIARY, TOOLS OR SERVICES SHOULD WE PROVIDE TO BOOST DATA DRIVEN INNOVATORS?

EUROPEAN NETWORK OF LIVING LABS (ENOLL) QUESTION
MODERATOR: ZSUZSANNA BODI
SUMMARY

The session provided fruitful insight for both the question hosting organisation (ENoLL) and other participants and co-hosts of the workshop.

The core discussion points on how the data users’ work could be enhanced were around:

- the interconnectedness of data;
- online platforms/marketplaces to host the data;
- quality measures to understand the value of data;
- discovering potential usage of data;
- commercial opportunities the access of data represents; and importantly
- the user perspective of making the best use of data and co-create innovations.

As a starting point there was a consensus between all session participants that the interoperability and standardisation of data is essential in the context of the European and the global digital market. Going beyond the ability of data communication the quality and value of data was under the microscope, where we concluded that further mechanisms are needed to engage a bigger audience making use of the availability and presence of data. Besides access, the understanding was found equally important, where experts’ help is needed to be translated into openly and publicly available toolkits and educational guidelines, trainings.

While we are aware lots of progress have been made the last years on the above said recommendations, participants articulated that there are still a lot of concerns and in general poor public understanding on data IPR related issues.

Once and in parallel while this key issue is being handled and answered, it was suggested that more experimentation and marketplace alike tools should be setup in order to forge new collaborations in PP(P)P -private-public-(people)-partnership. This type of large scale experimentations seems to bring lots of value added on the quality of life of citizens in the urban context and from a commercial point of view as well.

Some shortcomings and further potentials were identified in the context of using new technologies for data visualization and analysis, for which some concrete ideas were shared (please see below).

Last but not least based, on these exchanges, participants thought tools and methods are the least available to prioritize user perspective and user interface for data driven innovations. The discussion included the lack of involvement of final users in co-creation activities where there is data reused.

RAW DATA

Open and linked data

- Providing data enablers at different levels
- IoT as a data service
- Include open and private data (IPRs)
Data Markets
- Potential vertical or domain specific ones
- Setting up data quality standards
- Create a technical sandbox for data management lifecycle
- Common space, common digital single market

Ontologies
- Easy APIs/data sharing mechanism for pouring data from various sources
- Focus on data semantics and interoperability
- Common language used (multidisciplinary) for data coming from different sources

Data Processing Visual framework
- Focus on user experience and usability
- Problem based approach on the platforms (individual)
- DIY – what can I use / what is in it for me?
- Data Apps for playing with the available data around
- Data infographics

Data analysis/drill
- Services to analyse data without having to become a Data Analyst expert
- Create a regulatory sandbox that hosts trials of solutions
- IPR management and regulations

Advisors/Trainings/Methods and tools
- Create tools for facilitating cross-disciplinary workshops
- Train people in methods
- A smart way to find “mind buddies”
- Educational tools: guidelines, guidebooks
- Design thinking tools
- Data visualization tools -> making data more accessible
- Co-design tools and platform with end users of data

Monetization
- Intermediary that certifies the quality of data
- Data exchange between private and public sector
- Possibility for the final user to create/modify data set (e.g.: the information is not correct) -> crowd
- Build facilities where business meets government agencies
ABOUT ENOLL – HTTPS://ENOLL.ORG/

The European Network of Living Labs (ENoLL) is the international federation of benchmarked Living Labs in Europe and worldwide. Founded in November 2006 under the auspices of the Finnish European Presidency, the network has grown in ‘waves’ up to this day.

ENoLL counts today over 150+ active Living Labs members worldwide (440+ historically recognised over 12 years), including active members in 20 of the 28 EU Member States, 2 of the candidates and it is present in 5 continents in addition to Europe. Directly, as well as through its active members, ENoLL provides co-creation, user engagement, test and experimentation facilities targeting innovation in many different domains such as energy, media, mobility, healthcare, agrifood, etc. As such, ENoLL is well placed to act as a platform for best practice exchange, learning and support, and Living Lab international project development.

Living Labs (LLs) are defined as user-centred, open innovation ecosystems based on systematic user co-creation approach, integrating research and innovation processes in real life communities and settings.

- LLs are both practice-driven organisations that facilitate and foster open, collaborative innovation, as well as real-life environments or arenas where both open innovation and user innovation processes can be studied and subject to experiments and where new solutions are developed.
- LLs operate as intermediaries among citizens, research organisations, companies, cities and regions for joint value co-creation, rapid prototyping or validation to scale up innovation and businesses. LLs have common elements but multiple different implementations.
WHAT SKILLS SHOULD BE TAUGHT TO ENABLE CITIZENS TO UNDERSTAND, DECODE AND APPROPRIATE THEMSELVES THE DATA?

EWSHOPP QUESTION
MODERATOR: FERNANDO PERALES, JOT INTERNET MEDIA (ON BEHALF OF EWSHOPP PROJECT)

SUMMARY
Discussion focused on three main clusters: What to teach (top right), how (left) and societal benefits (down right). It is clearly stated that the collection and processing of data are key in the industrial domain, however in the society there is a clear lack of awareness about the availability and potential benefits of using the existing data sources and tools. This is mainly due to the limited access to data related subjects and lessons during the scholar period, covering from primary school till University and post grade courses.

In terms of WHAT topics should be taught, there were highlighted both theoretical concepts related to maths like logics, statistics and algorithms; and critical/logical thinking and security awareness about data protection.

In relation with HOW all these topics should be offered there was a great brainstorming, covering from the traditional learning by doing (which is not that common at school level) to gaming, promotion of social discussion (including social media), dedicated lessons for visualization and application/usage of the tools to real application.

Finally, the groups remarked which can be the potential benefits of a high data skilled society: identification of fake news, awareness of benefits for their business (even for small companies) and the opportunity offered to create new revenue models.
This discussion is closely related to the core EWSHOPP project, where is being implemented a toolkit covering all the data value chain (collection, integration, aggregation, processing, visualization and reporting) and enabling the integration of business data with external data source to implement new and enriched business services. During the project execution it has been clearly identified that companies are requiring more technical knowledge about how to manage and process the data, and we are trying to reduce this knowledge gap thanks to the collaboration with research centres and universities.

WHAT KNOWLEDGE HAS TO BE PROMOTED?

- Data literacy
- Logics/statistics/algorithms
- Logical and critical thinking
- Learn to evaluate the truth of the data
- Data bases usage and data analysis
- Understanding decision-making process made with data
- Understanding BIAS
- Understand which data you should protect
- How protect data and information
- Deconstructing systems, understanding components

HOW SHOULD THEY BE TAUGHT?

- Basic digital skills as important part in lessons is school / Included in school curricula
- Teach in primary school / civil society and professional
- Control mechanism, traceability
- Mechanisms to protect data
- Share, social discussion
- Overcoming fear for statistics
- Learning by doing
- More exposure to data everywhere
- Link to every data needs
- Code as: programming, material, medium (for communication)
- Semantic annotations / tools for semantics and data linking
- Visualization literacy
- Free on line courses
- Games about data
WHY ARE THEY BENEFITS FOR THE SOCIETY?

- Data is good for business
- More understanding of the services provided, more participation
- Avoid that other exploit your data
- Fake news/data uncover
- Awareness campaign “Lie detector”
ABOUT EWSHOPP - HTTPS://WWW.EW-SHO PPP.EU/

Objectives from https://cordis.europa.eu/project/rcn/207028/factsheet/en

In this project we aim at supporting companies operating in the fragmented European ecosystem of the eCommerce, Retail and Marketing industries to increase their efficiency and competitiveness by leveraging deep customer insights that are too challenging for them to obtain today. Improved insights will result from the analysis of large amount of data, acquired from different sources and sectors, and in multiple languages. The integration of consumer and market data collected by different business partners will ensure to cover customer interactions and activities across different channels, providing insights on rich customer journeys. These integrated data will be further enriched with information about weather and events, two crucial factors impacting on consumer choices. By disruptively increasing the analytical power coming from the integration of cross-sectorial and cross-language data sources and new data sources companies will deploy real-time responsive services for digital marketing, reporting-style services for market research, and advanced data and resource management services for Retail & eCommerce and their technology providers.

As of today, developing these services is too costly or nearly impossible for a large number of European companies. Even when these companies have developed excellent skills in analyzing data in their sector, they lack knowledge, technology and resources that are needed to integrate and analyze large and divers data in a timely manner. Enriching business data with weather data is difficult and costly. Using event data to obtain precise customer and market insights is even more challenging because of the difficulty of collecting and accessing data about events at a large scale. Language barriers, lack of agreed models and shared systems of identifiers to interlink data make these data integration tasks even only more challenging.
WHICH TOOLS OR SERVICES CAN BE DEVELOPED TO HELP CITIZENS TO UNDERSTAND DATA AND POTENTIALLY DETECT BIASED OR MALIGNANT USAGE?

BIGDATASTACK QUESTION
MODERATOR: DIMOSTHENIS KYRIAZIS (BIGDATASTACK PROJECT)

As the variety of (non) digital data sources increases, the generated data are inherently of questionable quality. The latter has a direct effect and impact on the decisions taken based on these data, as well as on the provision of services utilizing the aforementioned datasets. In this context, the aim of the session was to trigger discussions in the scope of tools and services that will help citizens both to identify and to exploit such tools to ensure that data usage and data-driven decisions are based on data of quality, and data that are not biased.

DISCUSSION SUMMARY

Considering the different dimensions and potential interpretations of the question, from the beginning of the session the contributors started structuring their inputs and ideas around key pillars of the data path. Thus, clusters of inputs have been created and evolved / further enhanced following the inputs from participants in the different sub-sessions. Based on the data path, the discussion captured aspects relevant to data understanding reflected in the need for data usage explanations (that can be achieved through concrete definitions), as well as through activities that would educate citizens. A key aspect raised in this context also relates to regulatory and legal frameworks both for usage and for contributions of datasets from different providers and sources. Moving towards the data sources themselves, participants highlighted the needs for the association of information for the data sources regarding who provided the data as well as reporting tools and services for the sources. The session participants concluded that such information could also be utilized if associated with the datasets, thus if datasets are annotated with the source-related information. Further to that, annotations could get into the space of
how data were actually interpreted and used. In the services space / cluster, participants contributed with ideas regarding data stream inspection / checking services (i.e. providing information on where the data resides, who and why provided the data, how long can the data be stored and used), as well as services for comparing data (in an attempt to identify potential biases). A quite interesting view refers to the provision of choice to citizens and data providers on how the data will be used. Thus, services need to be provided allowing a wide set of options to data contributors. Moreover, visualization has been identified as an important means both for understanding data per se and for understanding their quality. A variety of dashboards, visualizations, and infographics are required. Contributors also discussed the concept of human touch, enabling a social check of data through a kind of “facebook for data”. Finally, the importance of data quality has been made clear by several participants, raising the need for tangible outcomes relevant to quality in terms of labels or trust marks.

**RAW DATA**

**Understanding the data provisioning and usage**

- Data explanations (definitions)
- Data walks: “walkshops” towards j
- Joint training
- Data education in schools (primary)

**Regulatory frameworks for data provisioning**

- Liability measures in place (legal frameworks)

**Data sources**

- Information on who provided the data
- Citizen data monthly report
- Privacy app to inspect sensors

**Annotation of datasets**

- Semantic data
- Data annotations (metadata)
- Learn how it was interpreted (through annotations of data insights)

**Services**

- Data stream inspectors (where, who, why, how long?)
- Atomic service for data source quality
- Data comparisons
- Fact checker
- My data management app / tool
- Alternative choices can be made only when you have several options
Visualization

- Data dashboards
- Different visualizations
- Visuals (infographics)
- “My data” services (I see, I report mistakes)
- I can use data for my work, business, studies, etc (to produce result)
- The human touch: social check, FB for data

Outcome / validation of data quality

- Trusted mark / label (like ecommerce label)
- Traffic light for data quality & “level of trust”

Quality data added value

- Data improves services (quality)
- Higher identification with the product
ABOUT BIGDATASTACK – HTTPS://BIGDATASTACK.EU/

Objectives from https://cordis.europa.eu/project/rcn/213081/factsheet/en

The new data-driven industrial revolution highlights the need for big data technologies to unlock the potential in various application domains. To this end, BigDataStack delivers a complete high-performant stack of technologies addressing the emerging needs of data operations and applications. The stack is based on a frontrunner infrastructure management system that drives decisions according to data aspects thus being fully scalable, runtime adaptable and performant for big data operations and data-intensive Applications.

BigDataStack promotes automation and quality and ensures that the provided data are meaningful, of value and fit-for-purpose through its Data as a Service offering that addresses the complete data path with approaches for data cleaning, modelling, semantic interoperability, and distributed storage. BigDataStack introduces a pioneering technique for seamless analytics which analyses data in a holistic fashion across multiple data stores and locations, handling analytics on both data in flight and at rest. Complemented with an innovative CEP running in federated environments for real-time cross-stream processing, predictive algorithms and process mining, BigDataStack offers a complete suite for big data analytics.

BigDataStack holistic solution incorporates approaches for data-focused application analysis and dimensioning, and process modelling towards increased performance, agility and efficiency. A toolkit allowing the specification of analytics tasks in a declarative way, their integration in the data path, as well as an adaptive visualization environment, realize BigDataStack’s vision of openness and extensibility. With an emphasis on standardisation and open source contributions targeting high impact, BigDataStack will enable data operations and data-intensive applications to take full advantage of the developed technologies, exhibiting their applicability through three commercial use cases from the maritime, market and financing domains.
During ICT2019, BDVe project, with BDVA AND ECSO, and PPP projects Special and MyHealth-MyData, organized a networking session around the impact on privacy and security related to personal and machine-generated data. The brainstorm was organised around 6 questions and followed the carousel format.
Question Set #1
A. How AI may transform the game in term of threats but also protection? ---- BDVe (Tjerk Timan)
B. What are the main threats that you perceive due to exposure of personal and machine generated data? ---- ECSO (Fabio Martinelli)
C. How to provide incentives to data subjects/owners for engaging in the management of their data? ---- MHMD (Edwin Morley-Fletcher)

Question Set #2
a) How does data protection for "big data" differ from data protection for "small data"? ---- BDVA (Zoltan Mann)
b) How can we raise citizen awareness on potential risks about the data collected about you or your machines? ---- ECSO (Geraud Canet)
c) What additional privacy, transparency and trust mechanisms are needed to support emerging and future applications, such as Cyber Physical Social Systems and Intelligent Software Agent Systems? ---- Special (Sabrina Kirrane)

For all questions do or don’t, success or failure stories are welcome.

Coordination: Daniel Alonso, Jean-Christophe Pazzaglia (BDV PPP)
Roberto Cascella (ECSO)
BDVA SRIA

- Strategic Research and Innovation Agenda on Big Data Value
- See http://www.bdva.eu/SRIA
- Dedicated chapter on data protection challenges and solutions: Chapter 3.5

Your inputs for

Upcoming BDVA position paper on Privacy Enhancing Technologies (PET)

- What are the main developments on PETs?
- How can we accelerate uptake?
- What are the main challenges that lie ahead?

Website: www.bdva.eu
E-mail: secretarygeneral@core.bdva.eu
Follow us: Twitter: @BDVA_PPP

ECSO SRIA

Identify research priorities for 2018-2020

- A strategic vision is needed to demonstrate how industrial priorities contribute to the implementation of the strategy
- 7 thrusts organised in 4 different areas have been identified

Your inputs for

ECSO Technical paper on AI planned publication in 2019

- Strategic relevance of AI for EU
- Technical roadmap: cyber security challenges, their relevance, current status and future directions
- Impact on vertical sectors and their needs
- Relevant regulations and legislations, and their implications

Website: www.ecs.org.eu
E-mail: secretariat@ecs.org.eu
Follow us: Twitter: @ecso_eu
HOW AI MAY TRANSFORM THE GAME (OF CYBERSECURITY) IN TERMS OF NEW THREATS BUT ALSO NEW FORMS OF PROTECTION

BDVE QUESTION

MODERATOR: TJERK TIMAN, BDVE

SUMMARY

In this session we explored the role of developments and innovations that are generally grouped under “AI” on cybercrime-and cybersecurity. Looking at both drivers and barriers, or opportunities and risks of AI for cybersecurity, one of the main threads/lines of inquiry throughout the session was if and how AI triggers a new arms race and if the adversary can still be known (due to the informational, but also infrastructural asymmetry when it comes to both cyber-attacks and cyber security). On the side of drivers and opportunities, rather different categories were mentioned; a first being AI for data protection and better forms of data obfuscation, via for example tokenizing personal data via Neural Networks. Another strand of opportunities mentioned is that of AI for anomaly detection in networks and AI for pattern recognition of attacks: If, in the EU, we would share across sectors the types of attacks in terms of patterns and behaviours and which types of weaknesses are (attempted to be) exploited, we could build more resilience. AI could also be used to combined heterogeneous sources of data to get a better

1 Although, as one participant noted, in many cases inappropriately, or too imprecisely

36
understanding of an anomaly or an attack and classify/make better, more informed decisions (situational awareness).

In terms of **threats**, one line of reasoning was in line of the risks that come with adversarial algorithms (an AI-arms race): there can be (attacking) AIs for fooling an (protecting) AI that does anomaly detection, or that learns from failed attacks about the types of protection mechanisms present- another forms of exploiting weaknesses. Many threats were seen in this area – from feeding a protection system false data to automating attacks and dynamically learning from the defence while trying new forms of attacks. Flowing from this, a more philosophical (yet potentially very real) risks is that AI’s start developing patterns of risks-and accompanying interventions based on a system or logic of reasoning that we humans no longer grasp: the human-out-of-the-loop scenario is seen as a real threat in cybersecurity: since on the once hand we will have to make use of AI and ML to help deal with ever-increasing complexities and variations of cyber-attacks, yet on the other we want to still maintain a form of human-based risk analysis (e.g. severity and appropriate countermeasures)

**RAW DATA**

**Questions**

- Can we develop better authentication methods through AI?

**Drivers/ opportunities**

- Better data obfuscation
- Tokenizing personal data via NN (neural networks) – no one can reverse it – making use of the black box
- Network monitoring and protection through AI – mapping behavioural patterns of threats and ‘bad guys’
- ML techniques applied over anonymous cyber incident data exchanged over sectorial levels EU wide
- Better anomaly detection
- AI combined with Identity Analytics to reliably detect suspicious behaviour
- Anonymization of ML training data
- Classifying code commits for security relevance
- Use of edge/for computing for smart yet secure networks
- AI and heterogeneous sources for better situational awareness and classification of risks/ more contextual knowledge
- Better ways of learning from attacks
- Automatically checking of formal policies / sticky policies / increased accountability and transparency
- New field of expertise, new specialisations in AI4SEC
- Explain-ability and knowledge graphs in AI for security
- Real-time support users of possible risks while interacting online
Barriers/ risks

- Fake news (automated feeding of)
- Adversarial AI
- Incorrect automated conclusions from following online behaviour
- Dependency and outsourcing of security risk assessment to an AI
- AI asymmetry
- Fake data
- AI imitating behaviour of the person or thing being attacked (link to adversarial AI)
- Automation of attacks
- AI to find weaknesses in networks
- Attacks by/on physical assets via AI-instructed autonomous vehicles (drones for example)
- Emergence of novel types of risks outside our understanding (AI-epistemology beyond our reach)
- Combination with bio – organic computing
- Making incorrect predictions about future actions
- Skills-gaps and geographical boundaries of action – from cyber to real (e.g. intervention space)
- Human-out-of-the-loop interventions in cybersecurity
- Adversarial AI learning to be a ‘normal’ user by learning from anomalies: thereby also learning from types of defence mechanisms.
- AI to find patterns for potential re-identification in de-identified data
- Deep fakes (in video) making face recognition foolable (for example)
The mission of BDVe is to support the Big Data Value PPP in realizing a vibrant data-driven EU economy or said in other words, BDVe will support the implementation of the PPP to be a SUCCESS. Behind that mission, there are multiple goals to achieve, which should be taken into full consideration when defining the directions of the PPP. Some of the most challenging ones are: (1) achieving a more competitive landscape of European Big Data providers, leading to bigger market share; (2) creating the context for a more competitive EU industry (transport, manufacturing, public sector, agrifood, media, energy…) in the advent of a data-driven revolution where many traditional players will have to transform their processes and re-think their business if they want to remain competitive—or in some cases, just to survive; (3) ensuring the sustainability of the investments and actions triggered by the PPP.

BDVe has broken down those high-level goals into 7 major priorities for the project:

- Being accurately informed about most important facts in Big Data so that we have a solid basis to support the decision-making process in the PPP
- Supporting the implementation of the Big Data PPP from an operational point of view
- Developing a vibrant community around the PPP
- Supporting the development of a European network of infrastructures and centers of excellence around Big Data
- Setting-up a professional Communications strategy
- Setting up a framework that supports the acceleration of data-driven businesses, and
- Ensuring the sustainability of the investments and actions triggered by the PPP.

The BDVe consortium includes a set of partners that have shown commitment and dedication to the success of the PPP for several years. They have already invested and they have committed to invest along the coming years. We believe that this CSA cannot be a neutral action that offers operational support without further commitment.
HOW CAN WE RAISE CITIZEN AWARENESS ON POTENTIAL RISKS ABOUT THE DATA COLLECTED ABOUT YOU OR YOUR MACHINES?

ECSO QUESTION
MODERATOR: GERAUD CANET

SUMMARY

In this session, a lot of (more or less strong) educational measures have been proposed, spanning from children-oriented educational operations to shock campaigns, in the same spirit as tobacco, safe-driving alcohol campaigns. Measures helping the customer evaluate the impact, or actions towards the design of human interfaces, have also been proposed.

The simplest educational measures concern young consumers: organise workshops at school, inform with ads in social networks as Instagram, Snapchat etc. New technologies as immersive training or gamification could be used. Operations toward the parents could also be organised.

A more efficient educational measure consists in exposing the risk as it is, which means making the risk explicit: highlight the actual value of the data, explain consequences of data loss in terms of actual value or public reputation. AI can be used for the purpose of user empowerment: AI systems can be run to analyse what can be inferred from the data willingly stored on line by the users.

Finally, a lot could be done on enhancing human-machine interfaces of common data-consuming apps: explain data policy with an easier language, explore different visual means about the use of data, adapted to different kinds of users; add a what-if feature (what happens if I refuse to give up some data; in the opposite, what happens if my data leak); help visualise, with cartoons and narrative, what happens after a data breach.

In a more general view, many actions could be taken at many levels: with regards to the providers and the consumers, at educational level...

A greater awareness of the users is the best tool to pressure the providers.

RAW CONTENT

Education

- Ads in young social networks: Instagram, Music.ly, Snapchat etc.
- training for parents: → Educate the parents to educate the children
- immersive VR training
- gamification
- Involve users in workshops to participate in concrete/real experiences
- Through living labs examples
- workshops with kids at schools/parent schools
Attention-grabbing information

- pretend attacks
- 15-minute-glory consequences
- through a collection of worst-practice examples promoted to users by their favorite app
- publish the consequences of known data breaches to individuals (cyber-bullying, loss of money, etc.)
- advertising campaigns like those for alcohol and smoking cigarettes
- find the way to the data subject
- examples of bad things happening with data
- catalogue good and bad stories about data loss (what people did right or wrong)

Explicit Impact

- give figures on money made by other companies with your data
- ID numbers can count as much as access to other data
- extroversion initiatives
- workshops
- city campaigns
- school education for youngsters
- info kiosks on city services
- highlight the value of their data (a password can cost as much as your bank savings)
- run AI on data to show some abuses
- show how apparently unrelated data can be deduced with algorithms

Customer policy

- Legislation/regulation for data collectors: info day about potential risks
- children limitations or information on data collection
- do you really need to say yes when you are asked to provide personal data?
- identify the number of times you are pooled with the information you don't need (just because you gave data)
- easy and plain language
- different visual means about the use of data adapted to different kinds of users / different levels of knowledge
- very simple messages: no legal language with users
- kill social networks such as ...
- every portal requesting data has a what if feature
- visualisation/cartoon/narrative of what happens after a data breach
The European Cyber Security Organisation (ECSO) ASBL is a fully self-financed non-for-profit organisation under the Belgian law, established in June 2016.

ECSO represents the industry-led contractual counterpart to the European Commission for the implementation of the Cyber Security contractual Public-Private Partnership (cPPP). ECSO members include a wide variety of stakeholders such as large companies, SMEs and Start-ups, research centres, universities, end-users, operators, clusters and association as well as European Member State’s local, regional and national administrations, countries part of the European Economic Area (EEA) and the European Free Trade Association (EFTA) and H2020 associated countries.

The main objective of ECSO is to support all types of initiatives or projects that aim to develop, promote, encourage European cybersecurity, and in particular to:

- Foster and protect from cyber threats the growth of the European Digital Single Market;
- Develop the cybersecurity market in Europe and the growth of a competitive cybersecurity and ICT industry, with an increased market position;
- Develop and implement cybersecurity solutions for the critical steps of trusted supply chains, in sectoral applications where Europe is a leader.
“WHAT ADDITIONAL PRIVACY, TRANSPARENCY AND TRUST MECHANISMS ARE NEEDED TO SUPPORT EMERGING AND FUTURE APPLICATIONS, SUCH AS CYBER PHYSICAL SOCIAL SYSTEMS AND INTELLIGENT SOFTWARE AGENT SYSTEMS?”

SPECIAL QUESTION MODERATOR: SABRINA KIRRANE (VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS)

SUMMARY

The input provided by participants can broadly be classified under three core headings: user centric requirements, technical requirements, and process, education and policy requirements.

In terms of user centric requirements the discussion focused on the need for control and transparency mechanisms that are both usable and understandable. Participants were particularly concerned about the need to protect vulnerable users and children, and the need to safeguard particularly sensitive data such as genomics and health data. Other key considerations included the need to consider diversity in terms of both culture and multilingualism and to engage younger people in discussions with respect to privacy trust and transparency.
The discussion on technical requirements included the enhancement of steadfast security and privacy mechanisms, such as, encryption, anonymisation, authentication, and authorisation. However, there was also a major focus on explainability, in the context of how data is processed and with whom it is shared, with a particular concern on providing transparency with respect to latent and inferred data. In terms of potential solutions, participants where especially interested in automated control, transparency, and deletion mechanisms. In this context, self-describing objects and systems, and transparency with respect to data sharing in a chain of custody type fashion, were seen as potential starting points.

Although most of the discussion focused on user and technical requirements, participants also highlighted the need to educate citizens with respect to privacy preservation mechanisms, the limits of technical solutions, and gaps in terms of current legislation. Other key discussion points included the need for risk assessment and management techniques, and the need for additional legislative protection mechanisms.

**RAW DATA**

**User centric requirements**

- Transparency and privacy in genomics and health data (owned by citizens)
- End user centric privacy mechanisms
- User control frameworks
- Parent control versus technology children
- Protect vulnerable users and data subjects
- Multilingualism
- Clear and understandable information to the data subject
- Automatic cancellation of private data when action is completed with information to the owner
- Include cultural diversity
- Give control to young generation to define privacy, trust and transparency

**Technical requirements**

- We need efficient encrypted domain processing to allow safe data sharing and pooling
- We need anonymity mechanism that protect us from profiling
- One single identifier per person with a trust party system
- Privacy preserving collaboration/co-operation
- Trust – “level” of sharing organisation/agency
- Trusted automated non-attention and forgetting
- Explainability of decisions and actions
- Latent personal data transparency
- Physical systems authentication: are they trustworthy?
- Certified/trusted third party services
- Privacy vs transparency in future applications and blockchain
- Tractful algorithms
- Inference
- Chain of custody (unlimited copies challenge)
- Automated privacy aware data sharing mechanisms
• Self-describing smart policies for every connected object/system
• Need to manage bias

**Process, education and policy requirements**

• Knowledge on privacy bounds
• We need better laws: GDPR is full of gaps
• We need continuous risk level assessment and training in risk management for data subjects
• Increase people’s knowledge in basic technology so they can understand the possibilities
ABOUT SPECIAL - HTTPS://WWW.SPECIALPRIVACY.EU/

The SPECIAL project will address the contradiction between Big Data innovation and privacy-aware data protection by proposing a technical solution that makes both of these goals realistic. We will develop technology that: (i) supports the acquisition of user consent at collection time and the recording of both data and metadata (consent, policies, event data, context) according to legislative and user-specified policies; (ii) caters for privacy-aware, secure workflows that include usage/access control, transparency and compliance verification; (iii) demonstrates robustness in terms of performance, scalability and security all of which are necessary to support privacy preserving innovation in Big Data environments; and (iv) provides a dashboard with feedback and control features that make privacy in Big Data comprehensible and manageable for data subjects, controllers, and processors. SPECIAL shall allow citizens and organisations to share more data, while guaranteeing data protection compliance, thus enabling both trust and the creation of valuable new insights from shared data. Our vision will be realised and validated via real world use cases that - in order to be viable - need to overcome current challenges concerning the processing and sharing of data in a privacy preserving manner. In order to realise this vision, we will combine and significantly extend big data architectures to handle Linked Data, harness them with sticky policies as well as scalable queryable encryption, and develop advanced user interaction and control features: SPECIAL will build on top of the Big Data Europe and PrimeLife Projects, exploit their results, and further advance the state of the art of privacy enhancing technologies.