H2020 Work Programme 2016-2017
H202-IoT-02-2016
Duration: 36 months

CROSS FERTILISATION THROUGH ALIGNMENT, SYNCHRONISATION AND EXCHANGES FOR IoT

H2020 – CREATE-IoT Project

Deliverable 05.04

IoT Data Value Chain Model Evaluation & Final IoT Data Value Chain Model

Revision: 1.00
Due date: 31-12-2019 (m36)
Actual submission date: 20-01-2020
Lead partner: SINTEF

Dissemination level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Summary					
No and name	**D05.04 IoT Data Value Chain Model Evaluation & Final IoT Data Value Chain Model**				
Status	Final	**Due**	m36	**Date**	31-12-2019
Authors	Ovidiu Vermesan (SINTEF), Roy Bahr (SINTEF), Bertrand Copigneaux (IDATE), Arthur van der Wees (AL), Dimitra Stefanatou (AL), Prakriti Pathania (AL)				
Editors	Ovidiu Vermesan (SINTEF)				
DoW	The present document falls under the scope of the "Task 05.02: Data in the context of IoT applications". It is the final report focusing on one of the IoT Data Value Chain Model Evaluation and the final IoT Data Value Chain Model.				
Comments					

Document history			
Rev.	**Date**	**Author**	**Description**
0.00	04-12-2019	SINTEF	Template and input
0.01	10-12-2019	SINTEF, AL	Further development of introduction and several sections building up on deliverable D05.03
0.02	16-12-2019	SINTEF	Input on several sections
0.03	17-12-2019	SINTEF	Input section 3
0.04	20-12-2019	SINTEF	Input on several sections
0.05	23-12-2019	IDATE, SINTEF	Input sections 2
0.06	27-12-2019	SINTEF	Input sections 6
0.07	29-12-2019	AL	Further integration of updated input under section 6
0.08	03-01-2020	SINTEF	Inputs and structure updates
0.09	08-01-2020	SINTEF	Input sections 3
0.10	10-01-2020	IDATE	Internal review and comments considered
1.00	20-01-2020	SINTEF	Report processing and final version released.

Table of contents

Executive summary

Internet of Things (IoT) data value chains place emphasis on the utilization potential of IoT data, rather than on the data per se. This may result in the addition of several layers of value on top of the original raw data, which can be both private and public. As IoT data value chains are non-linear, they allow for the continuous use and re-use of data, which creates a series of challenges of legal and strategic relevance. Regulatory developments at EU level, relevant for the IoT Data Value Chains are legislative changes such as the General Data Protection Regulation (GDPR) and the Directive on security of network and information systems (NIS Directive) and the Regulation for the free flow of non- personal data.

The IoT Value Chain and data monetization for different industrial vertical sectors are addressed. The enablers of data monetisation solutions are technologies such as business intelligence, data mining, smart data and deep analytics through artificial intelligence (AI). Smart data is the most innovative set of technologies among analytics technologies. IoT data monetization in automotive and healthcare are elaborated through market overview, data availability, drivers, barriers, cost savings and new revenue opportunities. These sectors are represented in CREATE-IoT and are two important examples of verticals that have potential benefits of utilizing IoT data.

In addition, the deliverable produces an IoT Data Model for reliable IoT Data Value Chains aiming to support and further unleash the potential of existing and future IoT value chains. Based on the earlier discussion, it focuses on the attributes of economic feasibility, liability and transparency.

Furthermore, free flow of non-personal data is a key element for a competitive data economy within the Digital Single Market (DSM) and there is a need to provide the mechanisms and the regulatory framework to ensure a free flow of data, allowing companies and public administrations to store and process non-personal data. The "EU Code of conduct on agricultural data sharing by contractual agreement" is an example of attempt to define such a framework.

Addressing the data value chain implies focusing as well on connected data infrastructure that deliver on the non-physical level that will connect with the IoT devices used in different industrial sectors. Today, the global public cloud computing market is dominated by non-European players, but GAIA-X is a highly ambitious cloud project which plan to launch a European cloud service platform in 2020, open for both European and foreign countries following the projects rule on data sovereignty.

This document is addressed to IoT European Large-Scale Pilots Programme partners as well as the broader community of the IoT stakeholders. This deliverable form the final report based on the initial work and report prepared in 2017-2018, and the further work carried out towards December 2019. Some of the initial work are processed in the annex sections, with respect to the whole picture. The work forms the basis for the topics to be discussed at the workshop on "Policies to Support Open Data Marketplaces - Data Sharing in IoT Ecosystems, Data-supported Services Concepts & Best Practices" to take place on 29 January 2020 at The Hague.

1. INTRODUCTION

1.1 Purpose and target group

This document forms the final report due under "Task 05.02: Data in the context of IoT applications"[1], and address the IoT European Large-Scale Pilots Programme partners as well as the broader community of the IoT stakeholders. Further in this section we provide an overview of the relevant scene for IoT data value chains introducing the set of the associated concepts surfacing the particularities of the IoT data value chains.

Considering the disruptive nature of the IoT, current approaches when developing IoT business models need to be adapted accordingly based on a dynamic flexible IoT business model framework. The important opportunity in this regard is convergence of value chains with value networks on the context of IoT ecosystems. This will also affect the IoT Value Chain Data Models developed and used by IoT architectures, IoT platforms, IoT applications and IoT ecosystems.

The value chain concept was introduced in the field of Business Management as a tool to model the chain of activities that an organisation performs to deliver a valuable product or service to the market [1]. The value chain categorises the generic value-adding activities of an organisation allowing them to be understood and optimised. A value chain is made up of a series of subsystems each with inputs, transformation processes, and outputs.

In the context of IoT we can identify physical, digital and virtual data value chains as part of the different IoT ecosystems and over the IoT architectural layers.

The EC defines the data value chain as the "centre of the future knowledge economy, bringing the opportunities of the digital developments to the more traditional sectors (e.g. transport, financial services, health, manufacturing, retail)" [11]. This brings at least three distinct aspects along which the development of European data standards ought to be fostered:

- Standardised entity identifiers (i.e. identifiers of legal and physical persons, artefacts and their components, as well as time and location).
- Standardised, compositional concept systems (thesauri, taxonomies, ontologies).
- Standardised formats.

These elements are the key for IoT as standardised identifiers allow entities of interests (i.e. legal and physical persons as well as artefacts and their parts) and "things" to be reliably traced across independently established processes, managed by different applications, stakeholders, across connected supply, logistics chains and value networks. Standardised systems of concepts allow for IoT semantic integration, while standards need to be established for the actual formats in which data will be recorded on a medium for substrate or communicated across IoT networks and platforms [12].

From a similar standpoint, one can think of IoT as of a highly complex supply chain which connects an unlimited number of various devices together making it possible for the devices to communicate and operate through different infrastructures across various supply chain layers. As the supply chains extend across borders and industry sectors and domains, this supply chain in the existing and rapidly developing hyperconnected world is no longer linear. As a result, relations between the developers, vendors, consumers and other stakeholders of the digital economy and society (including but not limited to IoT enabled devices, systems or services) are non-linear. Collectively, they create an extensive multi-dimensional system which can also be referred to as the *supply chain ecosystem*. Every participant within this multi-dimensional ecosystem is relevant

[1] D05.04 IoT Data Value Chain Model Evaluation & Final IoT Data Value Chain Model: The evaluation report as well as the updated IoT Data Value Chain Model (D05.03) and a recommendation report beyond this project is due in December 2019.

and plays an important role in the design, engineering, manufacturing, deploying and functioning of both a connected device, system and service, as well as hyperconnected (IoT) ones [1].

1.2 Contributions of partners

AL is the task leader of "Task 05.02: Data in the context of IoT applications". Based on the contribution to D05.03: IoT Data Value Chain Model (Initial Version), AL provided an update of the regulatory developments, to the extent relevant for the scope and aims of the present deliverable, that took place as of January 2018 until December 2019.

SINTEF is deliverable responsible and contributed to solutions related to key verticals, data value chain, data classification, data life cycle, identity and access, management, data access, digital rights management, security, data management and data protection with a cross-domain IoT approach as part of the Digital Single Market strategy. The work considers the policies of data management for IoT applications that are context-aware and situational, and thus more complex to identify and assess, while data protection and security risks depends on the context and the purpose of the objects that are considered.

IDATE provided an analysis of the chain of markets of IoT Data by looking at the economic value chain. The analysis looks into the general, horizontal, value chain of IoT data, as well as more vertical analysis focused on key verticals of the Large-Scale Pilot projects (automotive, health). The power structures of the value chain and key players are presented. IDATE also provided an analysis of the economic feasibility of data monetization in the IoT domain. This analysis concentrates on the main opportunities and barriers toward monetization and present the potential of new revenues streams.

1.3 Relations to other activities in the project

The discussion on IoT Data Value Chain Model is relevant for all activities falling under the scope of CREATE-IoT project. As data present one of the building blocks of IoT ecosystems, an IoT Data Value Chain Model stands at the core of the IoT debate. This delivery builds on delivery D05.03.

As has been mentioned earlier, IoT ecosystems are extensive and actions within can have far-reaching consequences. It is important to consider these from the perspective of each individual Large-Scale Pilots as well as from broader and cross-cutting perspective of CREATE-IoT. The analysis of the IoT Data Value Chain requires a holistic approach, that incorporates numerous relevant perspectives, including technological, economical, as well as legal.

The analysis presented within this deliverable can contribute to other deliverables of WP05 on "IoT Policy Framework - Trusted, Safe and Legal Environment for IoT". In addition, certain topics have also been identified of particular relevance for related Work Packages, namely WP04 on "European IoT Value Chain Integration Framework" and WP06 on "IoT Interoperability and Standardisation". It is, also , closely linked to the discussion on the IoT Policy Framework produced under "D05.01 IoT Policy Framework" given that certain elements of the policy framework (e.g. Data Life Cycle, Device Life Cycle etc.) constitute main traits of the IoT Data Value Chains to be discussed, and addressed in "D05.09 IoT Data Value Chain Model Common Event".

2. IoT VALUE CHAIN AND DATA MONETISATION OPTIONS FOR VERTICALS

This chapter addresses the IoT Value Chain and data monetisation for different industrial vertical sectors. The enablers of data monetisation solutions are technologies such as business intelligence, data mining, smart data and deep analytics (now AI). Smart data is the most innovative set of technologies among analytics technologies. It is a disruptive concept thanks to the use of new algorithms and new data (namely unstructured data), but it is mostly the continuation of existing analytics technologies.

2.1 Market description

The accumulations of data both in businesses and on the Web, along with the growing number of open data initiatives we are seeing, have enabled the concept of big data to emerge.

However, the actual concept is not yet fully defined but regarding the data itself, big data has at least three characteristics:

- Volume: The data generated are usually produced in large volumes.
- Velocity: They are generated at high speed, some continuously and some frequently.
- Variety: The data generated is usually in various formats or types. These can include text, video, picture, sound or Websites.

Figure 1: Variety of data sources (Source: Capgemini)

All the data are usually considered as being 'unstructured' and thus difficult to handle and analyse. That is why specific 'big data' techniques are required for use with unstructured data. Most of these techniques are an extension of traditional data mining solutions, although new technical bases have been developed recently solely for big data – *MapReduce* and *Hadoop* are two examples.

Indeed, big data techniques can provide solutions to companies which want to process all their unstructured data but also - and especially - data located on the Web.

Structured vs unstructured:

- 'Structured' data: All data can be classified automatically in a table.
  - Examples: Relational databases, spreadsheets and accounting data, telephone directories.
- Unstructured data: It is impossible to put these in a table without processing them.

  - Examples: Emails and text, pictures, video.
- Semi-structured data: Data are not rigorously structured and need processing before classification in a table.
  - Examples: Social network data, logs.

The other key concept around big data is the data lake, regrouping all data from a company. The data lake includes structured data from relational databases like Access or SQL-based databases (rows and columns), semi-structured data (as with csv, logs and xml), unstructured data (full text in documents or web pages) and even audio visual files (images, audio and video), thus creating a centralised single data repository. Data can be just raw data but also processed data.

2.1.1 Major opportunities of smart data for verticals

The data used in IoT solutions can come from various means:
- Internally (internal databases).
- From data aggregation (forms, documents, sensors, web browsing).
- From open data sources and APIs from third parties exchanging data (automatic data transfer). Open data and data provided through APIs often have limited value.

Valuable data is generally blocked/controlled by their owners, as there is strong differentiation through data itself.

This is especially the case with major OTT (over-the-top) players like Facebook but also with major IoT players (data is generally only shared with user explicit consent on a case by case).

Beyond Internet giants, many vertical players have a wealth of information about millions of users. There are several opportunities for vertical stakeholders to use and share data:
- **Internal use**: Data are analysed internally and are not shared with third parties.
- **Intermediation for third parties**: Data are analysed internally, but the result of the analysis is shared with third parties, whether monetised or not – a way to monetise data without disclosing them.
- **Data sales to third parties**: Data are directly shared and monetised with third parties. Data will generally be anonymised and sold in aggregated versions.

Table 1: Main potential uses of big data by vertical players, by type of activity (Source: IDATE DigiWorld [23])

	Infrastructure (network, factories, inventory, logistics)	**Products and associated services**	**Customer management**
Internal use (no third party involved)	• Optimisation of infrastructure and logistics (potentially real time).	• Improvements of products and services based on customer feedback. • Development of new products. • Customer care.	• Increased sales of existing products (upselling, cross-selling). • Development of new services associated with products (servicisation). • Churn prevention. • Fraud detection.
Intermediation for third parties	• Insights on infrastructure performance to optimise their choices.	• APIs around added value data. • Billing.	• Ad networks. • Recommendations and sales of third-party products (reseller, trusted third party, trust-centric provider).
Sales to third parties	Insights and aggregated data sales.		

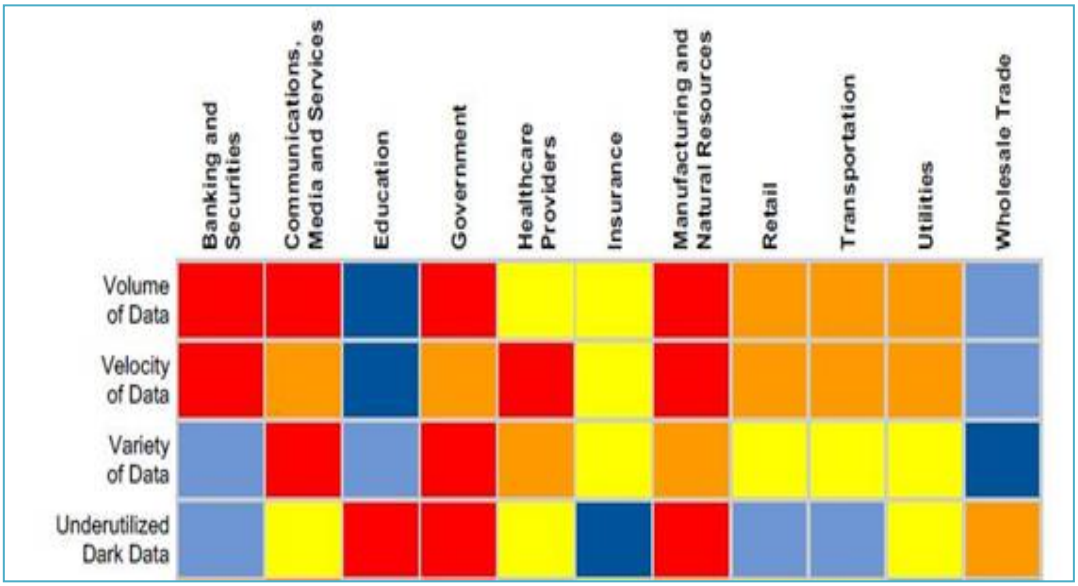
Data monetisation is only one of the potential options around data. Other models include customer care or improvement of existing infrastructure and products and services – these are not the focus in the remainder of this report. Similarly, free data-centric services provided to attract end users are not considered.

2.1.2 Vertical development in IoT

The IoT market is mainly developing vertical per vertical market, despite generally using the same types of technologies but often with different norms and sub technologies. The development has been faster for some verticals, providing more benefits than for other applications, by identifying more profitable business models. Industries involving moving objects have been more rapidly positively impacted by IoT/M2M, allowing remote monitoring (especially for the objects directly under control, i.e. machines that are located on a third-party footprint) and even remote control, reducing the need for staff.

The capacity to exploit data within vertical industries is also historically developing at different paces among the vertical markets, due to the data characteristics in terms of volume, value and velocity. Retail and finance (plus more recently telecom) industries are used to generate large amounts of data on a daily basis by following transactions and traffic. With the development of new channels (like mobile apps and social networks), B2B and B2B2C industries have now for instance access to more data related to their customers (with various types of formats).

Valuation of IoT data will therefore develop at least in the short/medium term by vertical markets. Indeed, the capacity to generate data through IoT sensors will depend on the vertical itself. The higher the number of machines, the bigger the perspectives thanks to larger amount of data. **But the value will come also from the capacity to generate new data** (and new dimensions of data) to cross them with other existing data. Thanks to IoT, some verticals will have access to new data, paving the way to new internal optimization and servicization business models.



Note: Red means high importance, blue means low

Figure 2: Data characteristics per vertical (Source: Gartner)

2.1.2.1 Selection of two vertical markets

To analyse the potential of IoT data, we selected two vertical markets to provide a first assessment of market potential, both in terms of cost savings and in terms of generation of new revenues. The selection has been based on:

- Number of connected machines in the short and long term, which favours logically consumer-facing verticals in B2B or B2B2C, especially automotive. Other markets are less important in terms of volumes.
- Capacity to create new types of data (or at least to communicate them), which provides more interest around verticals like healthcare (data was only produced in specific situations like within hospitals) and automotive (data was never used before, except for internal machine usage).

- Capacity to develop new usages thanks to IoT, using the connected object as a gateway to develop new services. This is the case of automotive (to develop new services like car or ride sharing or also services in interaction with the smart cities and the insurance companies). It may also be the case of healthcare with silver economy models.

Therefore, we decided to select the two verticals that are the most often quoted in the criteria above, i.e. automotive and healthcare, and that represented in CREATE-IoT. Some industries have more potential with data monetization (like finance, retail and telecom), but this will not come from IoT data.

2.2 IoT data monetization in automotive

2.2.1 Market overview

A 'connected car' is equipped with access to the Internet (the network of networks) whereby it can communicate with the outside world. The main goal of connectivity is to enable automakers to provide new services and extend existing ones. These services combine computer systems with remote communications technologies (such as GPS, wireless or cellular).

Core telematics services are:

- Emergency services, which include automatic collision notification, roadside assistance, good Samaritan assistance.
- Vehicle diagnostics, to receive regular emails detailed information (diagnostics) on the condition of the car.
- Navigation, with turn-by-turn instructions, using the GPS location.
- Information, such as traffic, gas price finder, local search and other (weather, sport, news)
- Entertainment services (streaming media for instance).
- Communication, including texts, emails, and calls.
- Remote information and access (unlocking the car, checking the status of batteries on an EV, finding the location of the car, etc.).

At the technical level, Internet connectivity is either embedded (with embedded SIM cards, using eUICC technology) or tethered to the driver's smartphone (using Bluetooth or a USB cable). Contrary to autonomous cars, connected cars are already widely available from all manufacturers, with many services existing. Figure 3 illustrates the whole automotive value chain including the main players.

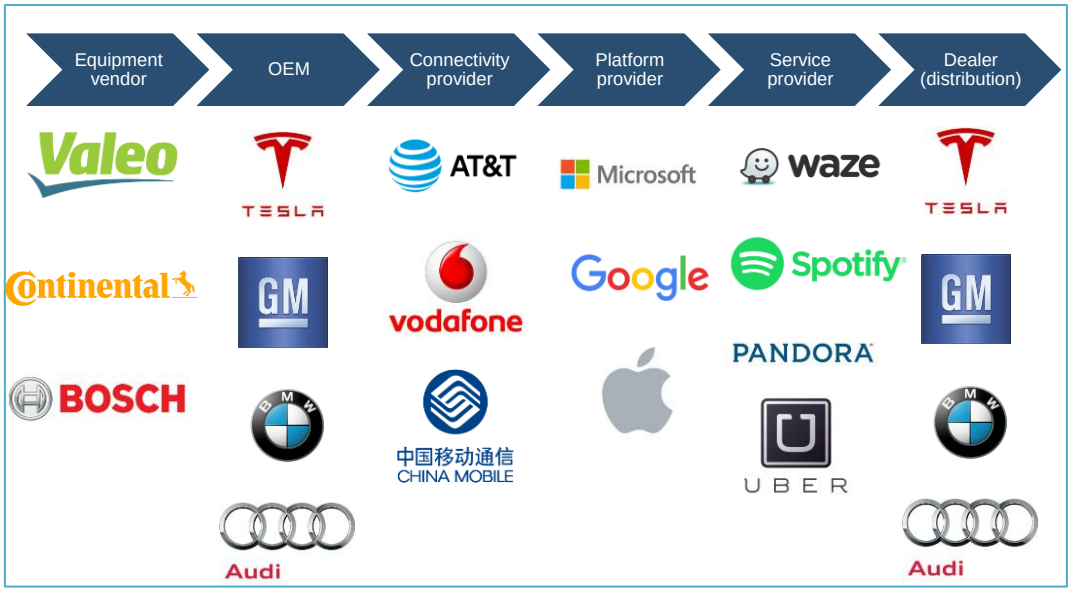


Figure 3: Main players of the automotive value chain (Source: IDATE DigiWorld [24])

2.2.2 The data available

2.2.2.1 Generating the data

The car has the potential to be a device carrying very valuable data. Currently some plug-in hybrid vehicles generate 25 GB of data in just one hour.

According to Intel, future cars are expected to generate around 4000 GB of data every day (for 1 hour of driving per day), mainly because of data from LIDAR and cameras.

Figure 4: Intel predictions for the data generated by future vehicles (Source: Intel)

This does not mean, however, that all the data generated by the car will be uploaded using an Internet connection; most of the data is expected to be used locally.

Tesla, for example, generated around 2GB of data for 3 months (a lot of which is due to video feeds being uploaded), mainly for internal purpose (improving their software).

The types of data that can be collected from automobiles are overall as follows:

- Vehicle, software, system health.
- Location and navigation.
- Speed, braking, acceleration and deceleration.
- In-cabin settings and activities including infotainment.
- Demographic information such as age range, revenue range and marital status.

General Motors (GM), for example, collects account information, vehicle-related information and driving information through its OnStar connected-car services, as detailed in its own words below:

- Account Information, like your contact and billing information and information about how you use certain OnStar services and our website(s).
- Vehicle-Related Information, like diagnostic information, odometer, oil life remaining, tire pressure, and information about collisions involving your vehicle.
- Driving Information, like location, GPS speed, safety belt usage, and other similar information about how the vehicle is used.

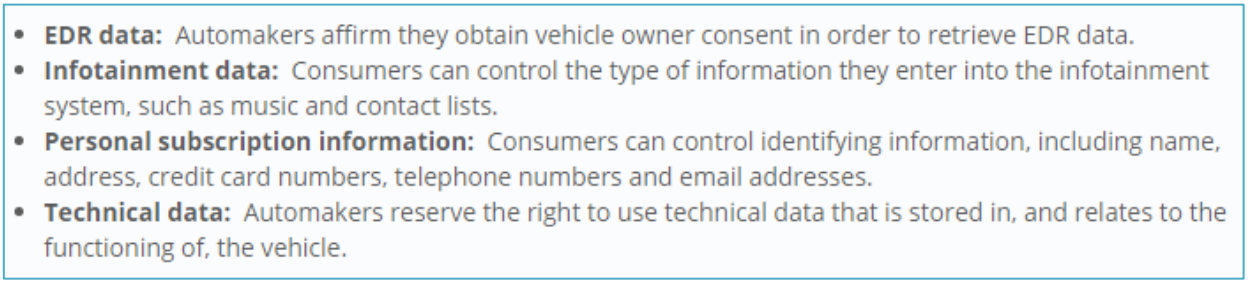
In addition, GM also collects information through a number of sources other than OnStar.

These include GM Websites, products and events, surveys, social media, incentive applications, GM credit card bank partners and other sources such as dealers that provide lists of potential vehicle purchasers and current owners – if, of course, such companies are permitted to share customers' information with GM pursuant to their privacy statements.

Owners of automotive data and third-party sharing

Generally, the car manufacturer owns the data related to the car, and already uses it to evaluate or research the safety, quality and functionality of vehicles and services, to develop new services or to maintain a lasting customer relationship.

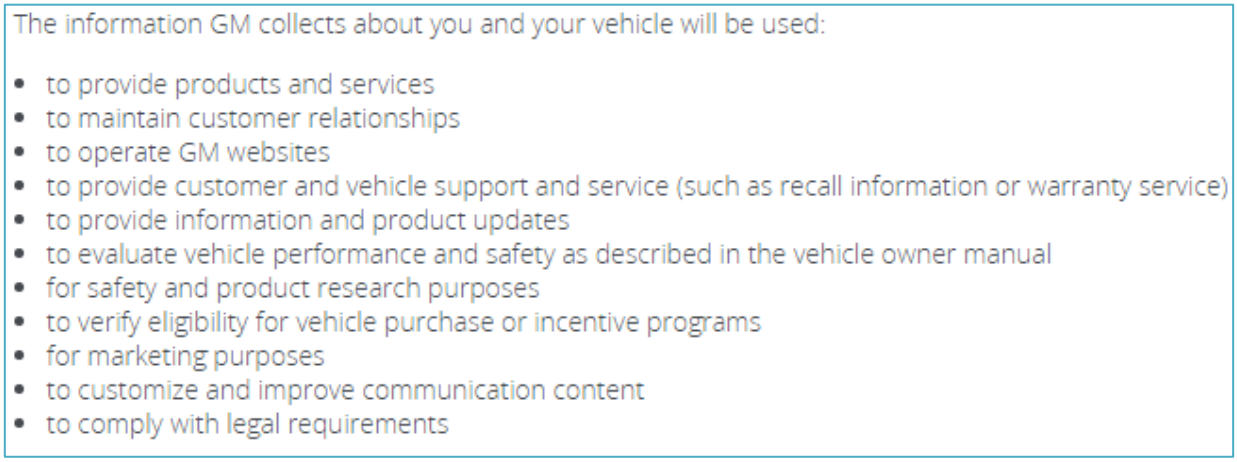
No laws or regulations have yet been made available to clarify such issues, whereas some common practices are taking shape within the industry itself, as cited below:


- **EDR data:** Automakers affirm they obtain vehicle owner consent in order to retrieve EDR data.
- **Infotainment data:** Consumers can control the type of information they enter into the infotainment system, such as music and contact lists.
- **Personal subscription information:** Consumers can control identifying information, including name, address, credit card numbers, telephone numbers and email addresses.
- **Technical data:** Automakers reserve the right to use technical data that is stored in, and relates to the functioning of, the vehicle.

Notes: EDR= Event data recorder

Figure 5: Commitments on data control (Source: Carvue)

Below is an example of how GM uses the collected data:


The information GM collects about you and your vehicle will be used:

- to provide products and services
- to maintain customer relationships
- to operate GM websites
- to provide customer and vehicle support and service (such as recall information or warranty service)
- to provide information and product updates
- to evaluate vehicle performance and safety as described in the vehicle owner manual
- for safety and product research purposes
- to verify eligibility for vehicle purchase or incentive programs
- for marketing purposes
- to customize and improve communication content
- to comply with legal requirements

Figure 6: How General Motors uses the automobile data (Source: General Motors)

According to a survey conducted by Capgemini, car owners prefer sharing car-related data to vehicle manufacturers and dealers. Not surprisingly, this willingness comes with a price, expressed as “anonymization” and “incentive programmes” by customers.


- as required by law, such as in conjunction with a subpoena, government inquiry, litigation, dispute resolution or similar legal process
- when we believe in good faith that disclosure is necessary to protect our rights, protect your safety or the safety of others, investigate fraud or respond to a law enforcement request
- with our services providers who work on our behalf and who do not have an independent right to use the information to which they have access or that we disclose to them
- with our business partners for GM marketing activities, business partner marketing activities or both
- with third parties for research and development purposes (such as university research institutes for improving highway safety)
- in connection with the sale, transfer or financing of a significant part of a GM business or its assets, including any such activities associated with a bankruptcy proceeding
- within GM, with our GM controlled subsidiaries and affiliates, with GM dealers and with GM licensees; however, transaction information regarding your GM Card will not be shared with GM dealers.

Figure 7: How automobile data is shared (Source: General Motors)

GM provides opt-out options to customers for certain data usage. For instance, if a customer opts out of receiving marketing communications from GM, his/her personal information will not be used for marketing. However, like in other industries, the collected data can be processed and shared, often in an anonymous and aggregated form, to third parties, for different applications.

As a result, GM indicates in the Privacy Statement that "GM will not share information about you or your vehicle with other third parties for their independent use without your permission".

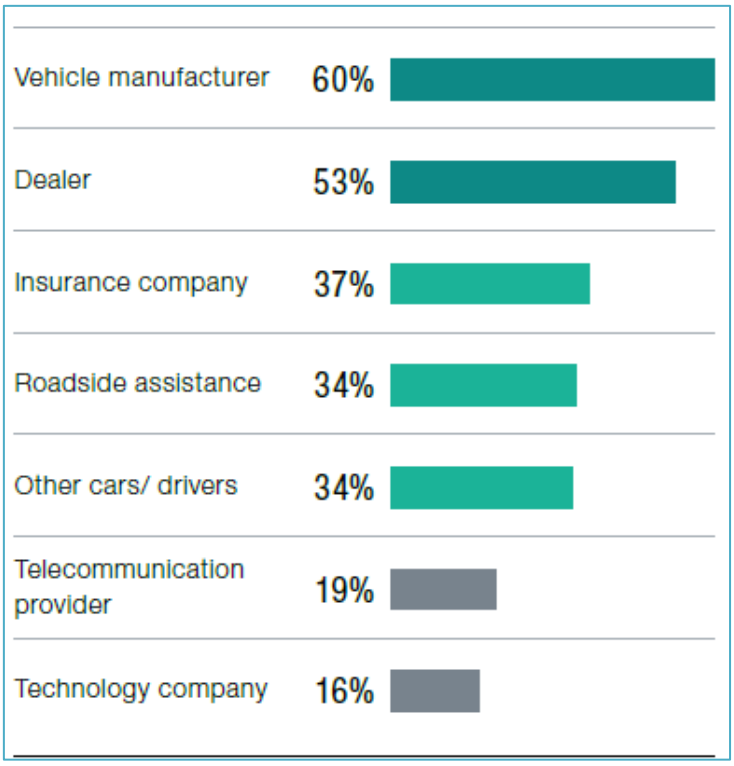


Figure 8: Preferred parties for connected-car data sharing (Source: Capgemini)

2.2.3 Drivers and barriers

The main drivers and barriers are illustrated in the following table and then explicitly developed in the sections below.

Table 2: Summary of key elements for the connected car data market (Source: IDATE DigiWorld)

Drivers	Barriers
As the number of connected cars grow, the pool of data grows as well	Sharing data with third parties is a sensible topic
Most consumers seem ready to share their data, if they know how it is used	User willingness to pay remains uncertain, which could lead to a lack of generated data

2.2.3.1 Drivers

As the number of connected cars grow, the pool of data, and services, grows as well

Newer cars tend to have the possibility to be connected to Internet, contrary to older models.

As a result, the number of connected cars is logically growing, and this trend is expected to continue.

With an ever-increasing number of connected cars, the total amount of generated data is growing as well, as is the relevance of services/products based on such data.

Consumers are looking for more services based on connected car data

Together with the growing adoption of connected cars, consumers expect to get more services based on the data from their car.

The majority of consumers (94%) [25] expressed their interest in apps and services based on this data.

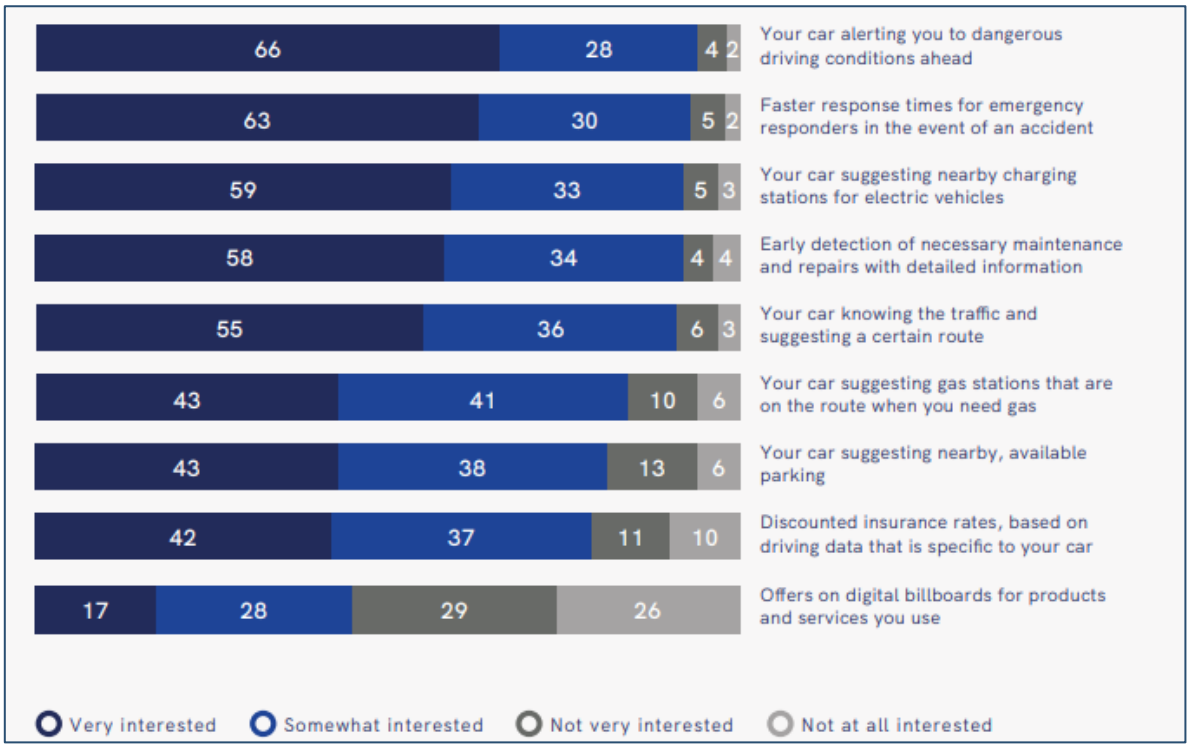


Figure 9: Drivers' interest in new apps and services (Source: Otonomo-Edison [25])

Most consumers seem ready to share their data if they know how it is used

Most consumers today are interested in connected car apps and services and the vast majority of those who expressed this interest are willing to share their personal data in order to get a high-quality service. However, they want to know how exactly the data collected by car manufacturers is used and want to be sure it is only used internally.

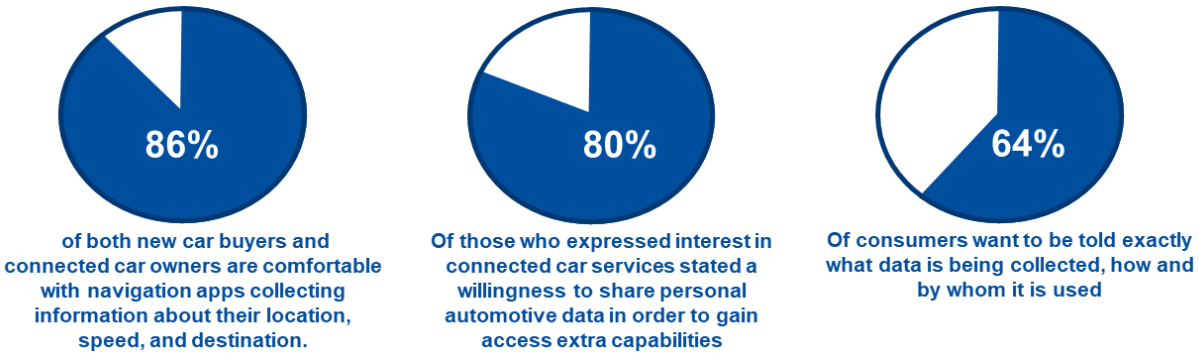


Figure 10: Customers' willingness to share connected car data with carmakers (Source: Otonomo-Edison [26])

2.2.3.2 Barriers

Sharing data with third parties is a sensible topic

While most consumers seem ready to allow their carmakers to use the data internally, selling or sharing this data with third parties is significantly less accepted. This could be especially the case if we consider unwarranted in-car advertising applications, since consumers appear weary of additional ads. If we consider personal computers, 47% of internet users globally use ad-blocker in 2019[2], generally blaming too frequent as well as annoying/intrusive ads. More generally, privacy fears (especially related to localization) are stronger if the data is shared/sold with third parties.

[2] GlobalWebIndex, 2019

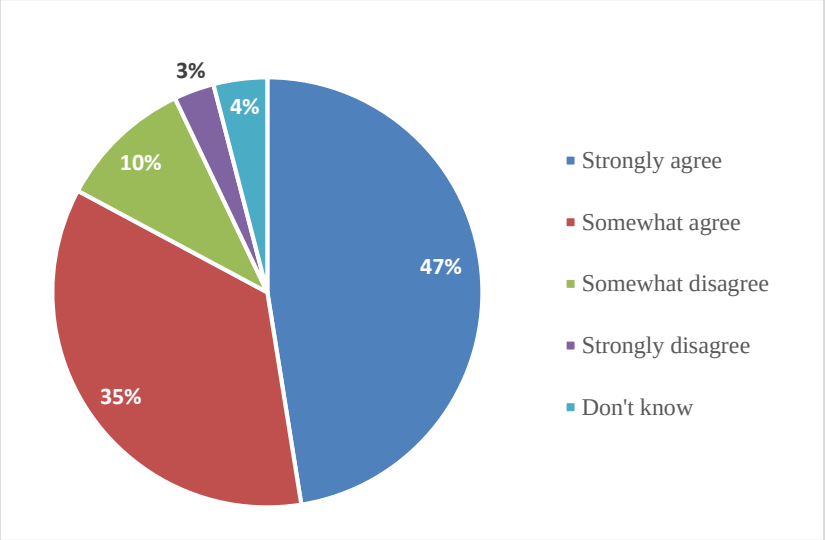


Figure 11: "I am concerned about companies selling my data to advertisers and other companies" (Source: Marketing Land [27])

As a result, automakers may be wary of monetizing such services externally, as new revenues (likely not gigantic) could be negligible compared to consumer backlash.

User willingness to pay remains uncertain, which could lead to a lack of generated data

For paying, data-generating services, the willingness of consumers to actually pay is, and will remain the major issue. Again, some questions may arise around the probability that ultimately, services will be mainly consumed individually, using a smartphone connected to mobile Internet, even in the car.

The cost of car connectivity plans remains non negligible, while the penetration rate of smartphone adoption is very high.

In the case that such services are not actively used, despite a high number of "connectable" cars, it will likely result in a lack of generated data (especially related to infotainment) and will thus hamper the development of the car data market.

This means an average of potential up to 45 EUR per month from declared intentions, which can be extrapolated in terms of real adoption to 15 to 30 EUR per month.

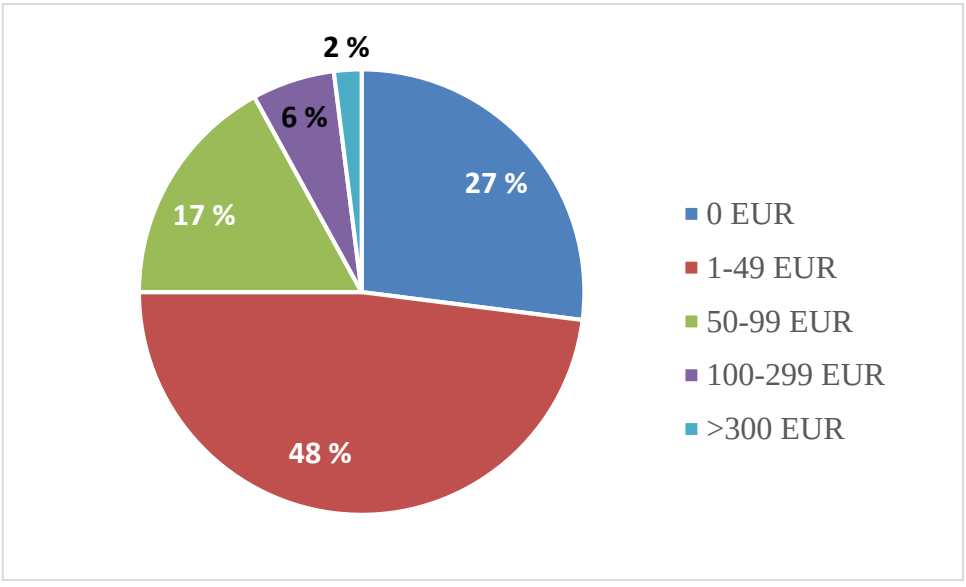


Figure 12: How much are customers willing to pay for connected car service per month? (Source: Deloitte [28])

2.2.4 Cost savings and new revenue opportunities for 2025

This section will look into how the use of data in connected cars could potentially help save costs for the industry, together with new revenue generation potential.

To start with the conclusion, the table below provides IDATE's final calculation results.

Table 3: Total cost savings and new revenues through data for 2025 in automobile

Cost savings	New revenues
36 to 39 billion EUR	17 to 24 billion EUR
0.9% of total automobile industry	0.45 to 0.6% of total automobile industry

2.2.4.1 Cost saving opportunities for 2025

The costs involved in the automotive industry

According to Automotive Engineering Partners, the cost breakdown within the automotive industry is as follows.

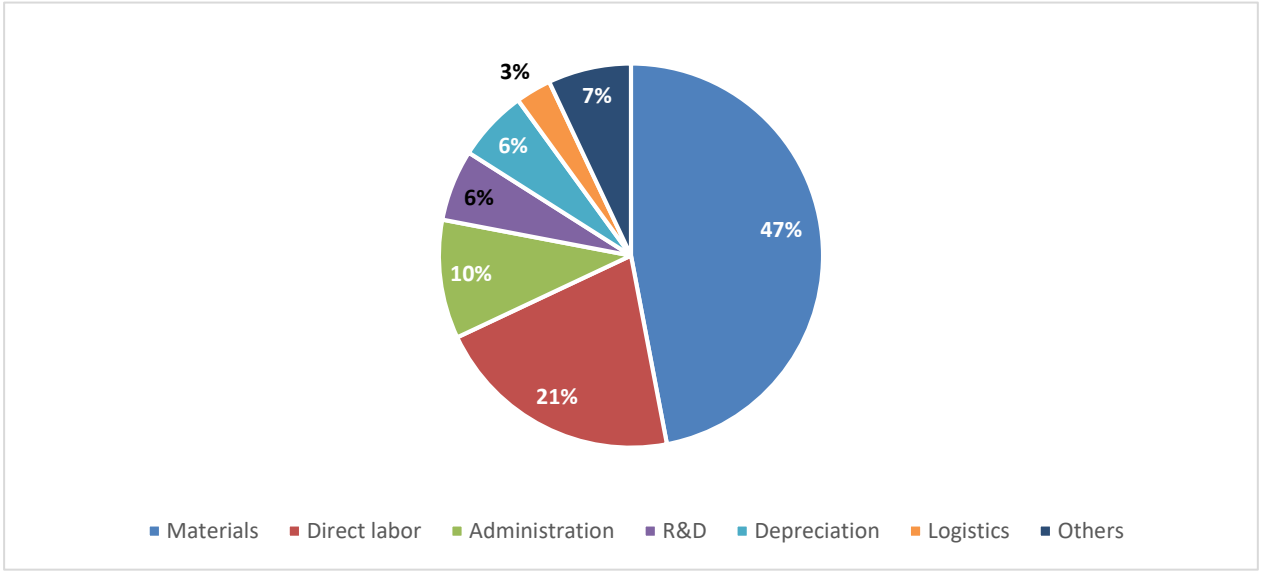


Figure 13: Cost breakdown within the automotive industry (Source: Automobile Engineering Partners)

Cost savings in logistics

The cost of raw materials account for the largest costs, at close to 50%, meaning that cost savings here would be the most ideal. However, the use of data cannot affect the pricing of the materials, whether it be steel, plastic, aluminium or any other material. While a car is made 47% of steel with a trend of shift towards aluminium (which is more costly but lighter, leading to better fuel consumption), the bottom line in terms of how data can provide cost savings here are negligible.

That said, the efficiency of their logistics, accounting for 3% of the total automotive industry costs, can be improved using data. The routes can be optimized (using for example real time traffic information) and the delivery trucks can be located and tracked. Warehouses can be made more autonomous by using such data, and shipping can be adapted to meet real time needs. This also extends to supply and inventory efficiency. It should be noted, however, that the use of data here is not necessarily directly linked to connected car data, but rather data can be used to improve the production of connected cars.

Embedding such data analytics can potentially increase efficiency leading to savings of around 10%. For example, Accenture estimates a 10% greater savings through improved supply chain efficiency, and Llamasoft estimates savings of between 5 to 15% if optimization analyses have not

been done in the last 3 to 5 years. DPDgroup has announced in 2019 to be able to reduce costs by 7+% for fleet utilization with a software-based solution (Transmetrics).

Cost savings in direct labour costs

While data may not be able to reduce material costs, the second largest cost factor within the automobile industry, the direct labour costs, can be reduced by using data.

The concept of Labour Cost Optimization for example, based on the idea of reducing variances and controls within the various processes and technologies within the organization, to focus on measurable and quantifiable financial return, could be applied here.

Critical to such concepts is the use of verifiable data to provide statistical analyses to aid the optimization decision making process, with the effect of anything between 0.5 to 2.5% savings on the labour costs according to Deloitte.

Cost savings in R&D

Third in the costs ranking for automotive industries is R&D spending. It is worth noting that the automotive industry is the world's third largest industry in terms of R&D spending.

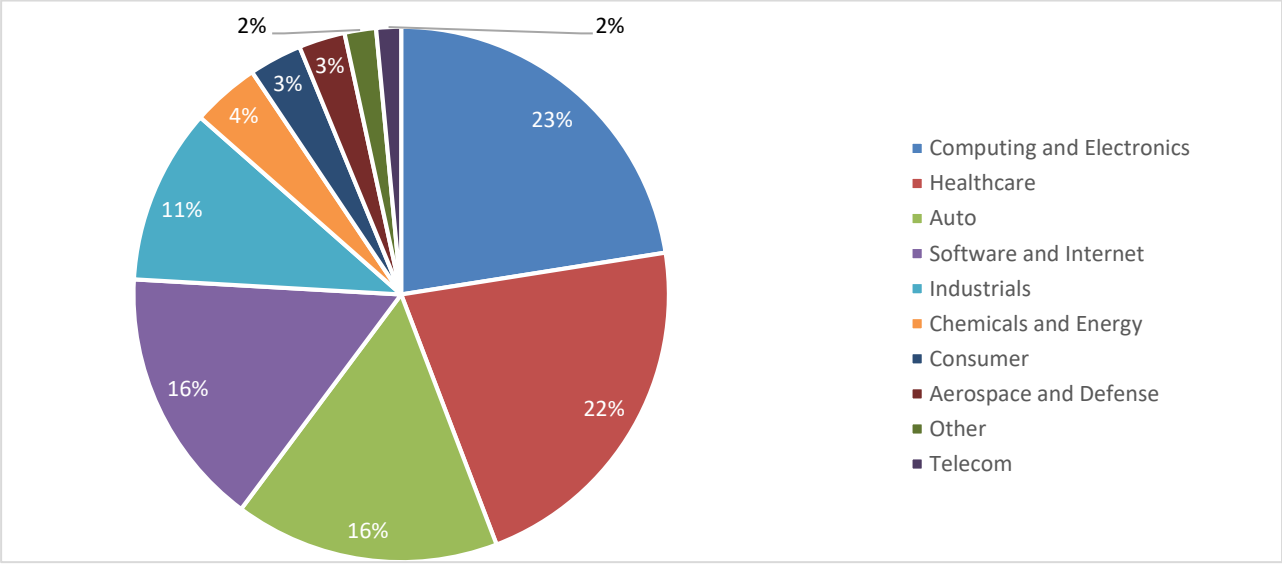


Figure 14: Percentage of total R&D spend by industry sectors in 2018 (Source: Bloomberg; Capital IQ; Thomson Reuters)

The idea of cost savings within R&D is of course not to cut the investment itself, but rather to make it more efficient.

Using data, R&D can be improved all the way from the initial planning phase through to the final output. R&D starts from the process to deciding what projects to go for and what to abandon, and the connected car data can offer insights to help such decisions.

All the data can be stored, managed and tagged as required to aid the entire research process. As the amount of data obtained from the connected car increases, so will the accuracy of the data and the variety of data, allowing for various analyses to improve the products and also plan future innovations.

Within the automotive industry, Tesla was the standout leader in terms of R&D intensity with close to 18% in 2016, the ratio between R&D investment and revenues, often used as a measure of how innovative a company is. But other big names in the industry are closing the gaps, while Tesla stands now at 12% in 2018.

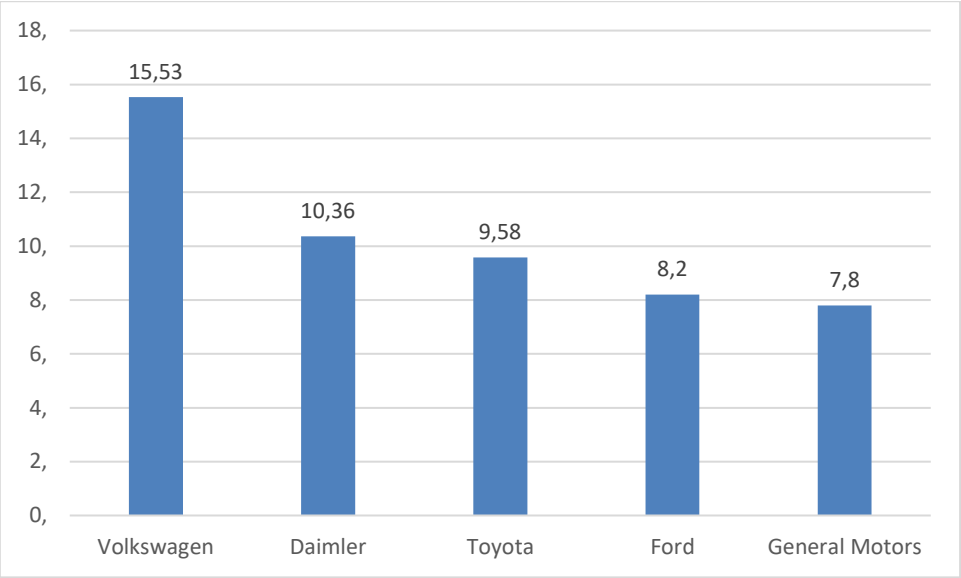


Figure 15: Global R&D spending: key companies in the automotive sector 2018, billion USD (Source: Statista, March 2019)

Cost savings through predictive maintenance

Although not provided as a cost component in the initial graph, data has the potential to provide cost savings through predictive maintenance. In essence, the car will be able to send data on its current conditions, and analysis of such data can lead to identifying and predicting problems which can then be tackled at an earlier stage, i.e. before that problem materializes as opposed to after. For example, predictive maintenance may deduce that a certain car part needs replacing, but without this prediction the car would continue until it breaks down creating a more serious and costly situation.

According to McKinsey, such IoT data based predictive maintenance could bring "economic value" of between 0.2 and 0.7 trillion USD by 2025, but it should be noted that it assumes certain preconditions to be in place such as interoperability between IoT systems and much more of the current available data be used. For example, as of today, on an oil rig that has 30,000 sensors, only 1% of the data are examined. It also includes elements such as insurance which is out of scope of this report (see also the section of revenue generation through new insurance).

The predictive maintenance market is expected to represent a 12.4 billion EUR market by 2024 (Source: IDATE Digiworld, Edge Computing, 2019) and close to 15 billion EUR, when addressing all industries (and not just automotive). It should be a key enabler of short-term and long-term savings, with a target of 12% of cost reduction[3].

With this in mind, IDATE predicts a realistic cost saving of around 6.7 billion in 2025, based on current maintenance costs accounting for 2% of the entire industry cost, with data allowing for an 10% reduction.

Conclusion: Cost savings around 0.9%, or around 37 billion EUR, can be obtained in automotive industry through data by 2025

To summarise, while both R&D and logistics can be expected to see cost savings through the use of data, their proportion to the overall automobile industry costs are small (6 to 9% and 3% respectively.

Labour costs account for 21% of overall automobile industry costs, and IDATE estimates that overall a cost saving of 0.3% will be made on the automotive industry.

[3] https://www.pwc.be/en/documents/20180926-pdm40-beyond-the-hype-report.pdf

Thus, in total IDATE forecasts a 0.8% to 0.9% cost saving in 2025 for the automotive industry through the use of data. The automobile industry as a whole stood at 4 trillion EUR as of 2019[4] and show grow by something close to 1% per year, and thus in terms of value cost savings of 34.4 to 37 billion EUR can be obtained through data use.

Table 4: Breakdown of cost saving calculation through IoT data for automotive industry

Costs	Share in automotive industry	Estimated cost saving through data	Estimated cost savings on automobile industry	Value (billion EUR)
R&D	6-9%	2%	0.12% - 0.18%	5-7.5
Logistics	3%	7.5%	0.22%	9.45
Labour	21%	1.5%	0.315%	13.23
Maintenance	2%	10%	0.2%	8,4

2.2.4.2 New revenues for 2025

New revenues through data enabled telematics and infotainment to generate 14 to 18 billion EUR of net value for the automotive industry, as part of 70 to 90 billion EUR industry.

The main applications of connectivity in the automotive field are:

- **Telematics**: This combines the power of computers and computer systems with remote communications technologies (such as GPS, wireless or cellular) to obtain information about remote automotive vehicles. Users can unlock their cars, check the status of batteries on electric cars, find the location of the car, or remotely activate the climate control system. It mainly relates to driver assistance services and also includes emergency call services. According to Market Study Report, in 2019, the market size of Automotive OEM Telematics is 4830 million US$ and it will reach 26000 million US$ in 2025, growing at a CAGR of 23.4% from 2019[5]. But telematics services may be offered by third parties too (through the smartphone for instance).
- **Infotainment**: By definition, this information-based media content includes entertainment content. It also includes telematics and more advanced tools. It is increasingly related to the mobile Internet through mobile applications. According to Market Study Report, in-vehicle Infotainment market is expected to grow to US$ 33.16 billion by 2025 from US$ 19.66 billion in 2016[6]. As for telematics, infotainment may be obtained through other systems (including smartphones).

Through the analytics of data coming from the car, it should be possible to provide advice to the driver based on his or her habits, and personalized services based on such data can be expected on the market for a subscription fee.

For example, should the data find that the driver goes from point A to B and back every Monday to Friday this is most likely for work, and then the personalized service could provide the driver with advice such as

- Cheapest and/or available parking spaces at destination.
- Cheapest petrol station on route.
- Real time congestion information together with alternative (faster) routes.
- Information on any non-daily events (police speed checks, accidents, roadworks, etc.).

[4] IBISWorld, Global Car & Automobile Sales Industry - Market Research Report, November 2019

[5] MarketStudyReport, Global (United States, European Union and China) Automotive OEM Telematics Market Research Report 2019-2025, July 2019

[6] MarketStudyReport, In-vehicle Infotainment Market 2025, February 2018

Through such personalized services, drivers can then make savings on their fuel consumption and refilling costs, parking costs and so on. Advances in OTA (over-the-air) software updates will also allow for further personalization.

Such connected car services are in fact already in use today. In general, the pioneer vehicles in this space are luxury cars as their clientele is more willing to pay for such services as driver assistance, emergency calls and concierge. Today, all major car manufacturers have developed a connected-car strategy.

General Motors is, among others, the pioneer manufacturer of the connected car, introducing service since 1995 in collaboration with Electronic Data Systems and the Hughes Electronics Corp. In the fiscal year of 2018, OnStar had around 14.5 million users.

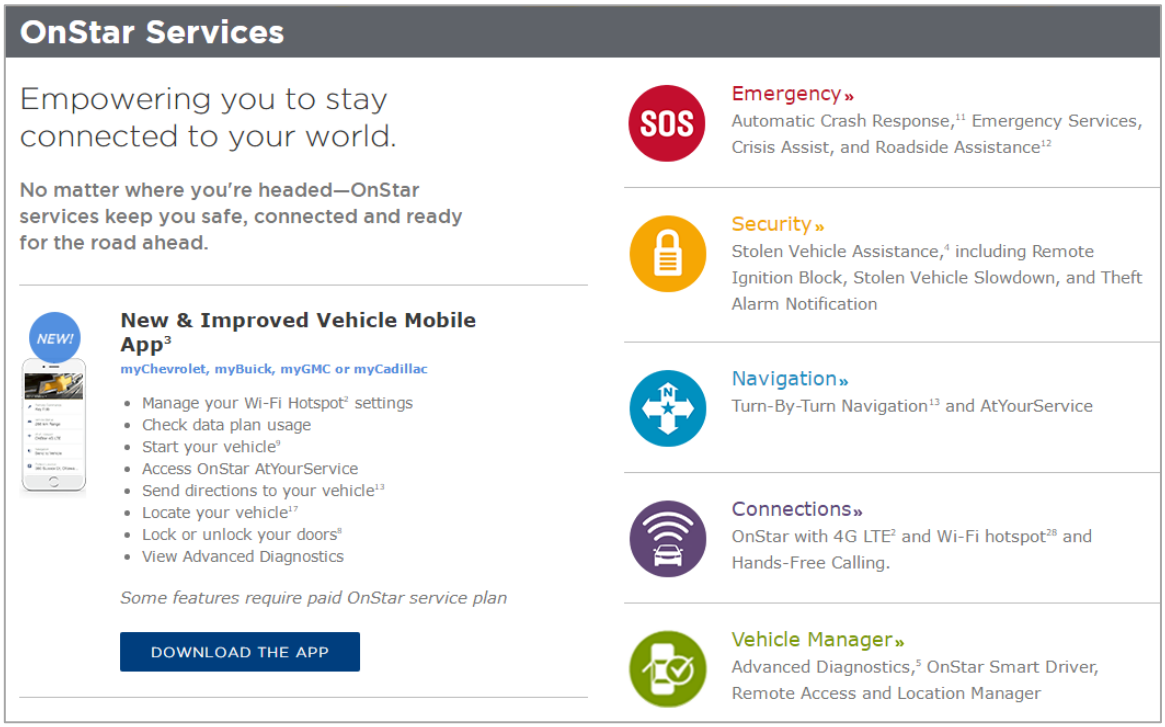


Figure 16: General Motors' OnStar connected car service portfolio (Source: OnStar)

Concerning the pricing, OnStar provides three plans with differing levels of services. The price range are relatively similar across all car manufacturers, at around 200 to 300 USD per year for a mid-range service.

Figure 17: OnStar plans and pricing (Source: OnStar)

Looking ahead to 2025, we can reasonably assume that the prices will fall as the services become more widespread, especially for the lower priced packages as the connected car services become more standard for the non-luxury cars. That being said, the price is not likely to come down that quickly either, with many subscribers opting to pay monthly as opposed to yearly (or longer) since the bulk sum is high. Taking all of this into consideration, IDATE estimates that in 2025 the yearly ARPU (average revenue per user) of connected car services will be between 150 to 200 EUR (see 2.3.2), including non-users. Counterpoint estimates the highest annual ARPU for connected car services at 180 USD in the US[7] for today's services.

Concerning the number of subscribers to such connected car services, GM's OnStar service has approximately 15 million users as of today, and it is one of the pioneers of this service. Considering that now virtually all car manufacturers offer similar services and taking into account global market share (see below), IDATE estimates that by 2019 the global number of subscribers to connected car services could reach 150 to 165 million (i.e. 75 to 80% of connected cars in 2020). According to IDATE, the installed base of connected cars (with embedded module) should reach 612 million units on the road by 2025, therefore we estimate (using the 75-80% ratio) that they will be around 450 million subscribers for connected cars services.

Table 5: Global car market share of the world's largest automobile OEMs in 2018 and 2019 (Source: Global Data Auto)

Company	Rank 2019	Rank 2018	Sales YTD 2019	Sales YTD 2018	+/- YTD 2019	Share YTD 2019	Share YTD 2018
Volkswagen Group	1	1	8.878.607	9.056.191	- 2,0 %	12,1 %	11,8 %
Toyota Group	2	3	8.517.681	8.354.921	+ 1,9 %	11,6 %	10,9 %
Renault Nissan Group	3	2	8.174.177	8.706.162	- 6,1 %	11,2 %	11,3 %
General Motors	4	4	6.317.737	7.110.766	- 11,2 %	8,6 %	9,2 %
Hyundai-Kia	5	5	5.933.296	6.108.094	- 2,9 %	8,1 %	7,9 %
Ford Group	6	6	4.362.540	4.753.555	- 8,2 %	6,0 %	6,2 %
Honda Motor	7	7	4.283.632	4.280.101	+ 0,1 %	5,8 %	5,6 %
F.C.A.	8	8	3.888.216	4.064.328	- 4,3 %	5,3 %	5,3 %
P.S.A.	9	9	3.168.460	3.436.419	- 7,8 %	4,3 %	4,5 %
Mercedes Daimler	10	11	2.339.839	2.233.870	+ 4,7 %	3,2 %	2,9 %

Thus, in conclusion in 2025, with 450 million subscribers with an average yearly ARPU of 150 to 200 EUR, IDATE expects new revenue generation of 70 to 90 billion EUR (numbers globally in line with figures identified earlier. Strictly speaking, assuming the car manufacturers share their revenues with connectivity providers (typically telcos) but also content/maps providers not all of the ARPU will belong to the automotive industry. Assuming such third-party providers take an 80% cut on average, **this leaves 14 to 18 billion EUR for the automobile industry.**

A note on theft protection type services

In the OnStar example given above, the service also includes security features such as theft alarm notification and remote ignition block. While there are also dedicated players such as LoJack who are dedicated to theft protection and recovery, the bulk of new revenue generation created in this regard is expected to come from car manufacturer services.

[7] Counterpoint, Connected Car Revenues to Grow Five-Fold by 2025, June 2019

New revenues through new insurance models; 2 to 4 billion EUR for the industry

Data could also be shared with insurers, with the goal of providing **pay-as-you-drive (PAYD) insurance**, whereby the costs of motor insurance are dependent upon type of vehicle used, measured against time, distance travelled and place. With the customer's prior consent, insurers are allowed to collect detailed granular driving data, and combine it with external information such as traffic patterns and weather data, in order to enhance the accuracy and value of the vehicle score based on the data analysis. Drivers are therefore charged a range of flexible premium fees on a usage basis and have optional value-added services such as the tracking of stolen vehicles. As illustrated in the figure below, consumers still view insurance discounts as the primary benefit of using PAYD.

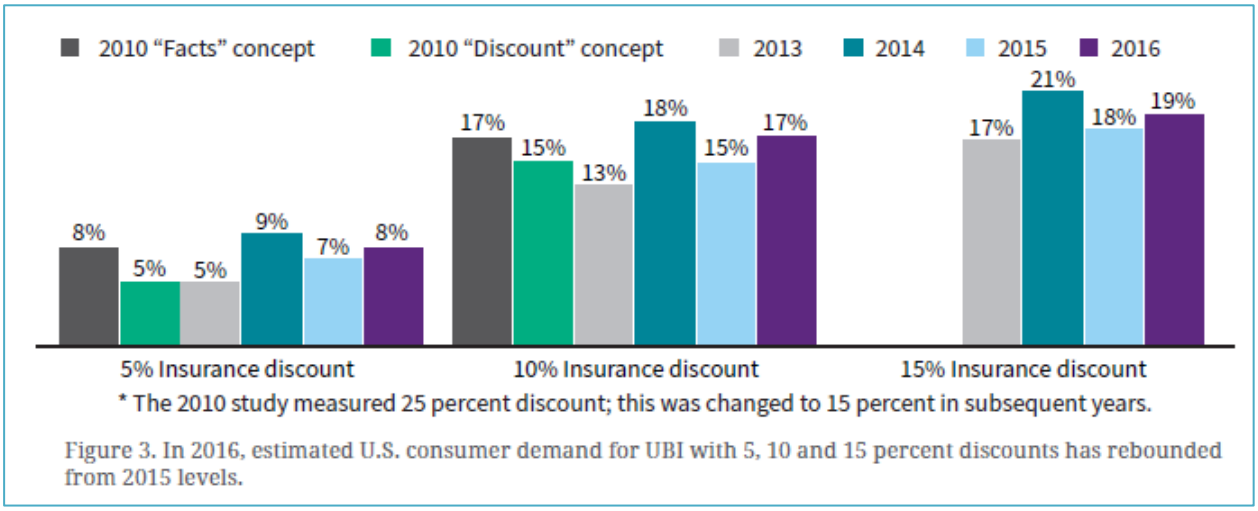


Figure 18: UBI interest growing with insurance discounts (Source: LexisNexis)

Some large insurers, along with independent car insurance start-ups, are joining the PAYD trend. In the US, National General Insurance is one of the first and largest insurance companies to have offered PAYD plans to OnStar subscribers. In Europe, Allianz launched PAYD programmes by partnering with BMW in the UK in 2014 on the 'FlexiMile' telematics insurance scheme, and then in France in 2015 with TomTom Telematics. The insurers claim that customers can generally save between 20% and 50% on premium fees through such programmes.

According to MarketsandMarkets, the usage-based insurance market is expected to garner 95.8 billion USD by 2025, with a CAGR of 18.95% from 2018. The real interest here is how much of this revenue will go back to the automotive industry. The PAYD insurance model requires data gained from the car, and thus it is only logical that the automotive industry who provides this data also gets a share of the generated revenue. This is not likely to be large, as ultimately it is still an insurance market driven by insurance specific players, but assuming automobile manufacturers negotiate a 2 to 4% share **then this would generate a 2 to 4 billion EUR market**.

New revenues through advertising and aggregated data sales expected to be negligible

The principle of use of personal data in the Internet service world today is advertising, with the likes of Google and Facebook generating 90% of their revenues from advertising. The data generated from connected cars are no different, and potentially provide invaluable data for advertisers.

In particular, targeted advertising specific to the car and/or driver could be delivered, with advertising made specifically matching the driving route. For example, restaurants could target drivers who will pass in front of their shop, and in fact target further by other factors such as gender, time of day, whether it is near the destination, and so on.

For example, in September 2017, Adobe announced new automotive focused analytics, personalization and advertising capabilities in Adobe Experience Cloud "that give brands the ability to deliver unique consumer experiences". In addition to offering more personalized services such as personalized playlists and on-route recommendations for auto makers and in-car app developers, the new analytics allows for personalized audio advertising.

Automobile industry expected to gain a billion EUR through advertising in 2025

The overall market of advertising is set to become increasingly digital. IDATE predicts that the digital advertising market worldwide will reach 581.5 billion EUR by 2025 [29]. While this is a big figure, the reality is that over half of this market is already taken by Google and Facebook, and this power balance is not expected to change any time soon.

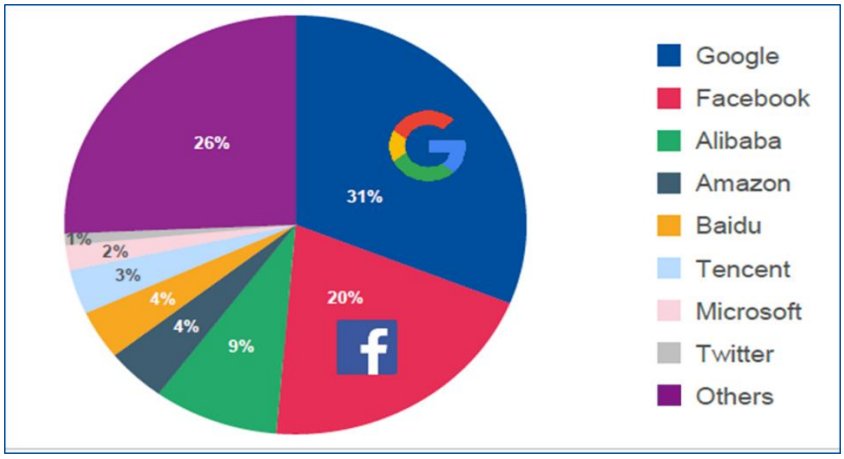


Figure 19: Player shares of online advertising revenue, 2018 (Source: IDATE DigiWorld [29])

As the figure shows, 26% of the digital advertising market belongs to "others"; in short, there is a myriad of players who own a very small percentage (less than 1%) of the market. This makes sense when considering that any website with some form of advertising receives a share of the advertising revenue pie. Automobile industry players are no different and are currently just another piece within the "others" category. Even with the potential of being able to offer targeted in-car advertising, the reach will remain very small in comparison to the Internet giants, and moreover there will almost certainly be intermediaries involved (such as Adobe, see earlier) who will also take their share of the revenue. Assuming the automobile industry can gain a 0.2% share of the entire advertising revenues share in 2025, this will amount to 1 billion EUR.

Aggregated data / insight sales still in its infancy with negligible revenues

Some car manufacturers are also looking at the potential of new revenue generation through the sale of aggregated data. The concept exists and players such as PSA and BMW have started to provide these services, for instance to some cities in order to improve urban planning (identify hot spots for accidents, zones with higher than expected speeds, etc.). For PSA, the sale of car data is done in partnership with IBM (data service partner).

Carmakers are in competition with telecoms operators, which can provide similar data (and even pedestrian data) thanks to smartphones. However, carmakers have access to more precise data, for instance related to braking. As a result, some telecom operators are also interested in buying missing data from carmakers.

At best, revenues expected should be lower than advertising, therefore between 0 and 1 billion EUR by 2025.

Conclusion: New revenues of 17 to 24 billion EUR, can be obtained in automotive industry through data

To summarise, the main source of new revenue generation for the automotive industry will come from more personalised telematics and infotainment services, around 16 billion EUR. Furthermore, new forms of insurance, headed by PAYD, is expected to generate around 3 billion EUR of revenues in 2025.

Table 6: Breakdown new revenue calculation through IoT data for automotive industry

Revenue	Value (billion EUR)
Personalized services (including infotainment and telematics)	14 to 18
Advertisements	1
Aggregated data sales	0 to 1
Insurance	2 to 4

2.3 IoT data monetization in healthcare (remote patient monitoring)

2.3.1 Market overview

In the healthcare industries, IoT solutions make self-managed and home-centred care available and hold potential to lower down overall healthcare expenditures. Through connected devices and care delivery platforms, care givers, patients and funders are connected within health systems in a continuous manner. Connected medical devices are not necessarily wearable but are incorporated with connectivity module so that they can communicate with the health-hub (a gateway device), which then transmits personal health data to caregivers at preconfigured interval.

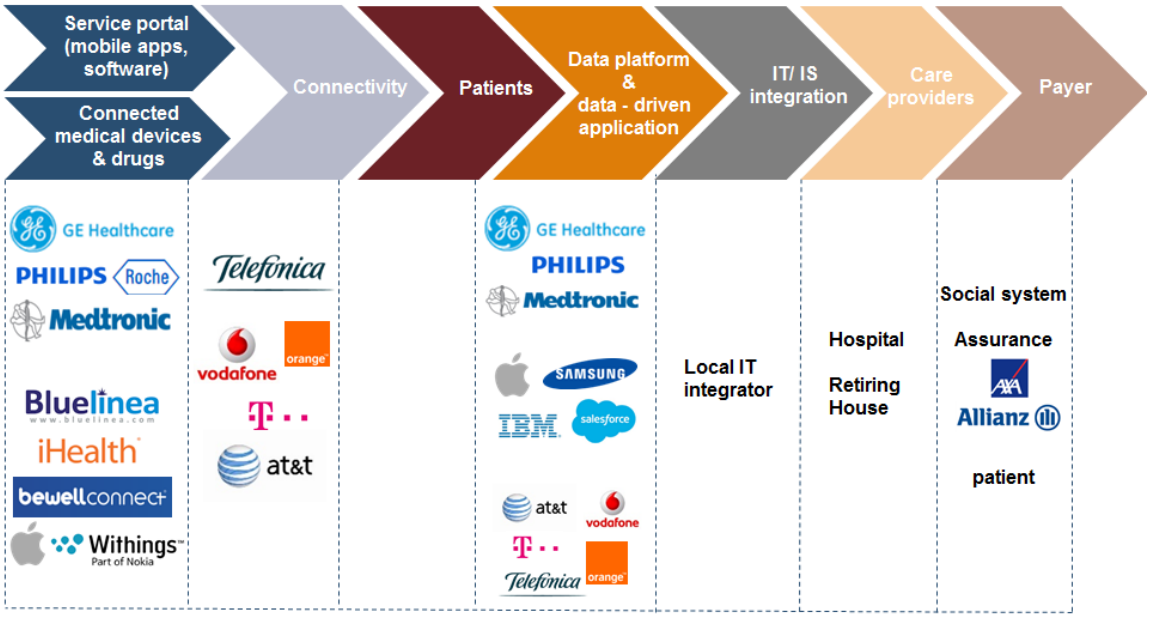


Figure 20: eHealth value chain (Source: IDATE DigiWorld [30])

The principle application of IoT in healthcare is **remote patient monitoring (RPM)**, particularly for those living with chronic conditions, senior and fragile population. An effective RPM portfolio encompasses:

- **RPM:** 1) Self-monitoring multiple vital signs via connected sensors, normally including weight, pulse, temperature, blood sugar, blood oxygen, and ECG or respiratory conditions. 2) Medical apps allowing share self-measured health data with designated persons (doctors or families), accessible in a patient-controlled manner. 3) Access to health data and analytics.

- **Telemedicine**: Virtual consultation with doctors that enlarges care access. In the meantime, it allows an interpretation of self-measured data without physical visits to doctors.
- **Tele-assistance:** 1) Immediate alert to monitored patients, their families and hospitals/doctors. 2) Redesigned home-care delivery by particular caring teams, for regular follow-up or emergencies.

Data is at core of connected healthcare since data silo problems are historically a severe challenge to health industries. The proliferation of personal health data and genomic data poses high requirement for data management and securing. For those reasons, both medtech giants and ICT players come to fierce competition in this segment.

The position of the medtech companies in health industries is firm, and not at all likely to be shaken in the near future. Medtech companies have inherent advantages over others in medical device development and in their knowledge of regulatory affairs, care delivery paths and care providers' working modalities.

Telcos are also eyeing this promising market with initiatives mainly sitting on "connected hospital", covering system interoperability, cloud-based data platform and data securing. Beyond that, the rising application is in remote patient monitoring, through a wholesale approach, partnering with health insurers in most cases.

2.3.2 The data available

2.3.2.1 Who creates the data?

Sensors are the main source for data generation, which can take place at home rather than having to perform at hospitals or clinics. The major components could measure different parameters like:

- Health parameters: Blood pressure, heart rate, glucose levels in blood, body temperature, etc.
- Activities: Steps, physical activity, distance run, etc.
- Wellness: weight, sleep quality, etc.
- Environmental parameters: fall up detection, temperature, pressure, wind, etc.
- And all other features which integrated sensors can already measure.

Through the sensors, anything in the human body and surroundings can be measured, potentially. This is often called the **'**Quantified Self**'** concept. It is not a recent notion, but connected objects make it easier and more efficient than before, when it was mainly through the paper-based dashboard.

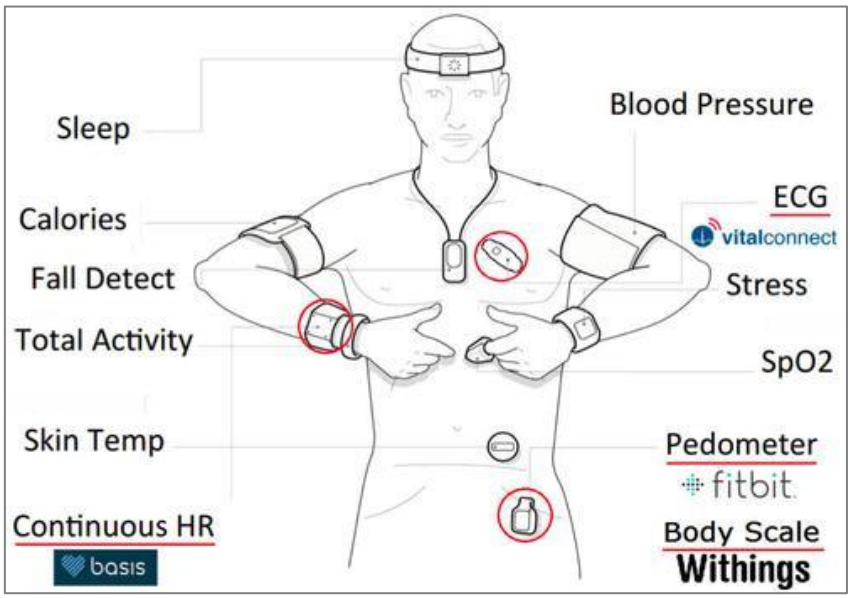


Figure 21: Different sensors on the human body (Source: IDATE DigiWorld [31])

2.3.2.2 Who owns the data?

Actually, there is to date, no consensus on who owns health and medical records at regulatory level, neither at practicing level. In the US, the Health Insurance Portability and Accountability Act (HIPAA) does not specify ownership and state laws are inconsistent. In April 2017, European Commission closed a public consultation on "Building a European Data Economy", but addressing the barriers impeding the free flow of data mainly, while not clarifying on data ownership yet.

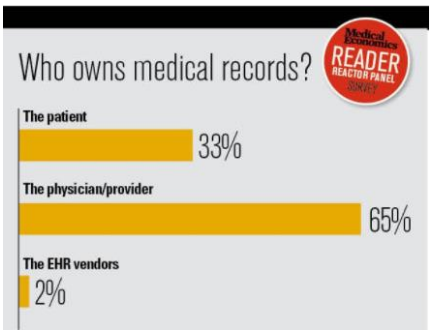


Figure 22: Who owns medical records? (Source: Medical Economics Reader Reactor Panel)

A survey conducted by US Medical Economics Reader Reactor Panel showed though healthcare providers believe patients ultimately own their health records, the default setting is that providers have most control over it. Nevertheless, the authorities do implement a series of rules to regulate the health data access, exchange and data processing.

Regulation (EU) 2016/679 (GDPR: General Data Protection Regulation)

The European Unions has adopted this new regulation with regard to the processing of personal data and on the free movement of such data. It will come into effect from 28 May 2018, repealing Directive 95/46/EC (General Data Protection Regulation).

With particular interest into personal health data, the new regulation will continue to providing EU citizens with fundamental privacy rights, notably with respect to the collection, recording, storage, distribution, consultation, modification and use of their health data.

On this basis, the Regulation 2016/679 provides **clearer rights to patients** compare to the Directive from 1995, and **advocates for a balanced approach** to protect patients' privacy while ensuring patient's data can be shared smoothly for healthcare and research purposes. For instance, citizens have the rights to access their own personal health data, to transfer the data from one data controller[8] to another with consent, and to object to the processing of their health data.

Health Insurance Portability and Accountability Act (HIPAA)

At the federal level, the US continues to take a sector-specific approach to data protection legislation. The aim of the US HIPAA is to guarantee the protection of individually identifiable health data. In particular, HIPAA defines who can have access to health-related information. In most cases, they can only be used by healthcare professionals as part of a medical treatment or for coordinating care.

Beyond that, under the new **HIPAA omnibus rule, health IT vendors become directly responsible for patient security** in their system. Meanwhile, patients have increased control over their personal health information (PHI). Before marketing a third-party service based on PHI or sell/provide access to this data at a fee, this business entity must obtain permission from patients whose PHI will be used.

[8] Data controller: The persons or entities which collect and process personal data. They determine the purpose(s) and means for processing the data. For instance, medical practitioners are usually controllers of their patients' data.

2.3.2.3 What is the data flow?

IoT-based personal healthcare requires a smooth data exchange from remote personal health devices, to health gateway device and finally to health systems (as with EHRs). However, the sensor-generated data at distance is scattered, and not of the same data format and structure which allows the direct integration to Hospital Information Systems (HIS).

To tackle such interoperability problems, Personal Connected Health Alliance (PCHA) publishes the Continua Design Guidelines that clearly define end-to-end interoperable interfaces that enable the secure flow of medical data among sensors, gateways, end services and HIS.

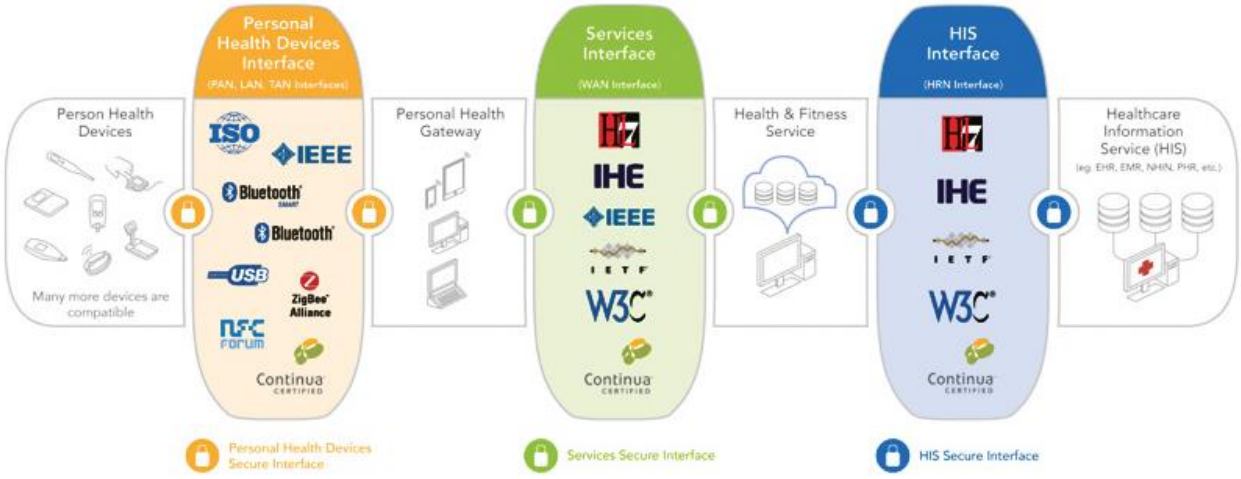


Figure 23: End-to-end interoperability solutions advised by Continua Design Guidelines (Source: The Personal Connected Health Alliance (PCHA))

The Services Interface allows the uploading of the data gathered from personal health device over a wide area network. The data relay must obtain users' consent. On this basis, security is achieved through confidentiality and service authentication, consent management and enforcement, auditing, and entity authentication.

The Health Information Services Interface provides for the electronic exchange of health records employing an HL7-based[9] PHMR (Personal Healthcare Monitoring Report). HL7-based PHMR is a specification defined and developed by Continua and HL7 to aggregate and deliver personal healthcare monitoring information to EMRs (electronic medical records), including the measurements captured by personal health devices.

2.3.3 Drivers and barriers

2.3.3.1 Drivers

Potentials empowered by health data

The health data sharing and democratization will not only extend care access and help in affordable healthcare coordination, the sensitive nature of health data also nurtures new business opportunities. Personalised care (or precise medicine) is expected to be the next step of eHealth leveraging big data, AI and genetic technologies. In the long run, precise medicine may also provide high economic potential. It could assist in foretell of chronic diseases and prevent deterioration in the future, thus lowering the overall social and care cost down.

[9] Health Level Seven International (HL7) is a not-for-profit, ANSI-accredited standards developing organization dedicated to providing a comprehensive framework and related standards for the exchange, integration, sharing and retrieval of electronic health information that supports clinical practice and the management, delivery and evaluation of health services (http://www.hl7.org/).

Public authorities' support

Healthcare, as a public good associated with many sensitive issues for patients and practitioners alike, is always being charged by public authorities to benefit the citizen. Health data, in this sense, are not seen as a goal in itself, but as a tool to reach certain purposes that benefit the public.

A number of government initiatives are taking place to unleash the health data potential. For instance, National Health Service (NHS) England has worked on care.data initiatives, aiming to upload all GP (general practitioner)-held data to a central repository, while give patients the opt-out right to reject the upload. On EU-level, a series of implementation focus on giving guidance on the data access control and data use, data-sharing without compromising patients' rights and improved interoperability.

2.3.3.2 Barriers

Strict data regulation and security concern

Healthcare industries operate in a highly regulated space, which differentiates it from other IoT markets. Although consumers are relatively more open than before to sharing health data, the health data access and processing is still constrained by to-be-solved data silo problems, and related regulation and policies (as discussed before, HIPAA, GDPR and likewise). Due to those constraints, the data-driven power (for example, in personalized care) is yet to unleash.

Figure 24: Healthcare organizations experiencing a lot of cyberattacks (Source: Armis, 2019; Irdetto, 2019)

Beyond that, cyber security also arises a lot of concerns. The most recent and severe attack was caused by WannaCry virus - around 50 NHS Trusts in the UK were badly affected. The increasing use of connected medical devices at home and processing of personal health data put more demanding requirements to increase cyber security level, and to standardize the personal health data use.

To-be-proven capex/opex efficiency for carers

The incentives for other stakeholders than patients (caregivers, clinician and payers) to participate in data-centric health programs, so far, insufficient, given the capex/opex issues coming with new system investment and care pathway alignment.

Some countries are hence prioritizing investment in eHealth domain by doing medical-social evaluation. For example, quite recently, since 2017, NHS has started to redesign healthcare services and delivery model at scale, with objective to accelerating the deployment of eHealth services in a cost-effective manner.

2.3.4 Cost savings and new revenue opportunities for 2025

This section will look into how the use of IoT data in the health industry, specifically remote health monitoring could potentially help save costs for the industry, together with new revenue generation potential.

To start with the conclusion, the table below provides IDATE's final calculation results.

Table 7: Total cost savings and new revenues through data for 2025 in health vertical

Cost savings	New revenues
271.6 billion EUR	32.5 billion EUR
3.36% of total health industry	0.35% of total health industry

2.3.4.1 Cost saving opportunities for 2025

Healthcare expenditures/costs structure

The costs involved in healthcare can be broken down into three main categories, as follows:

1. **Curative/rehabilitative care (inpatient):** Hospital operation & hospitalization, GP consultation, emergency care.
2. **Long-term care + outpatient:** Home-delivered care and nursing centre and ancillary services
3. **Medical goods:** medical devices & equipment, consumable supplies, etc.

Out of these three categories, the cost of medical goods will not be affected directly by the use of IoT data. One could potentially argue that through the use of data, optimization and cost savings could be made and reduce prices, but this will not be considered as it does not rely directly on the health data made available through patient monitoring.

In contrast, items 1 and 2 of the above will be directly impacted by the use of patient IoT data. In simple terms, remote patient monitoring allows patients to stay more at home and less at hospitals or clinics. Thus, expenditure on the curative and the rehabilitative care will go down (cost savings), while long term care and outpatient related new revenue generation can be expected.

Breakdown of health expenditure

According to the OECD and various other sources, long term and outpatient care accounts for 28% of total health expenditure. Inpatient care accounts for 45%. It should be noted that there are quite large differences depending on the country.

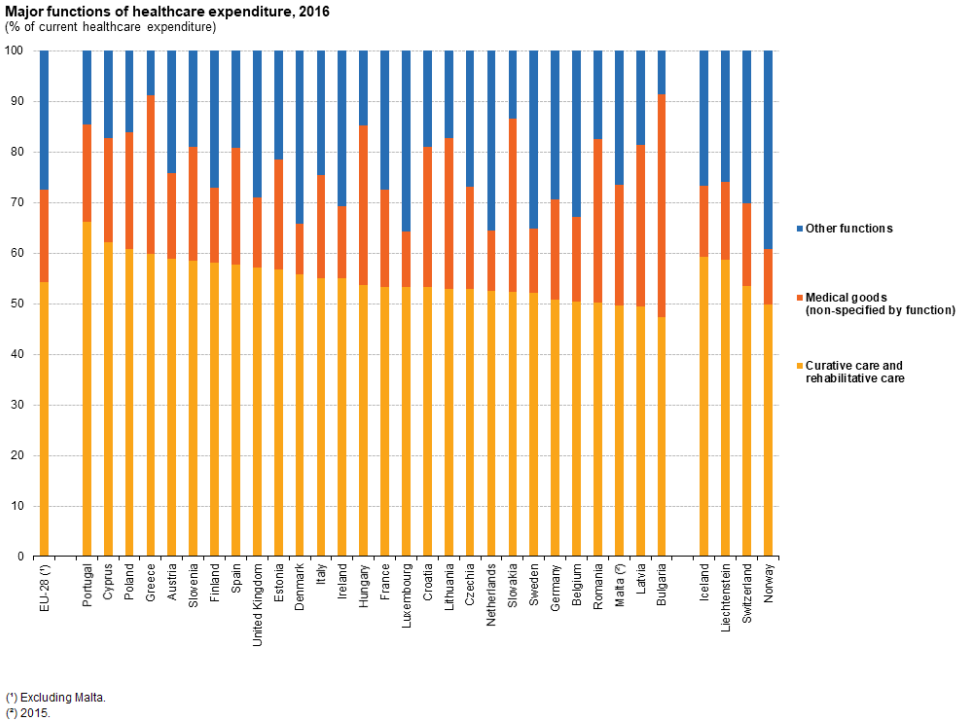


Figure 25: Major functions of healthcare expenditure in EU countries, 2016, % of current healthcare expenditure (Source: Eurostat, 2016)

Meanwhile, the world health expenditure stands at roughly 10% of global GDP. The 2018 global GDP was 85 trillion USD; thus, health expenditure is roughly 8.5 trillion USD.

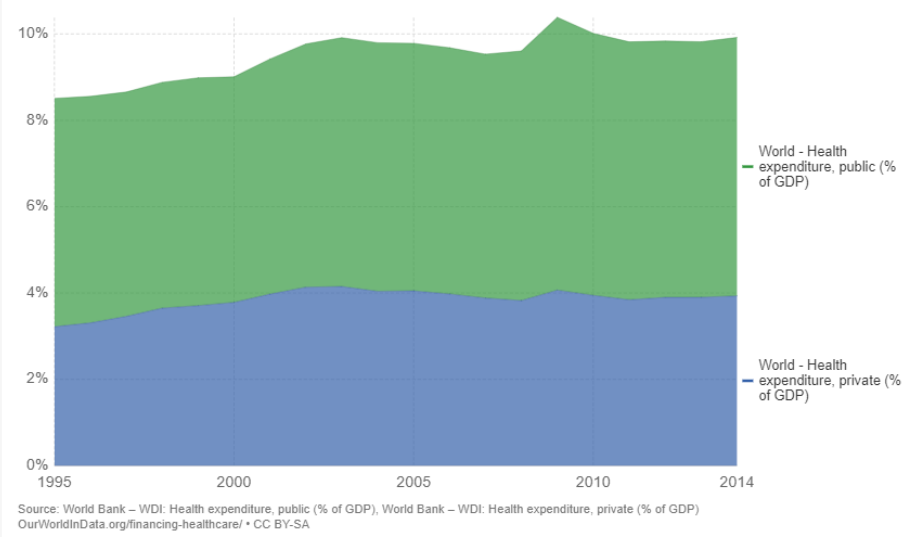


Figure 26: World health expenditure as share of the global GDP by source of funds (Source: World Bank)

Savings per patient

There are various reports that try to calculate the cost savings, which can made on health expenditure, specifically on inpatient expenditure. However, these studies vary wildly on their results. For example, Biotricity's study involving 90,000 patients found that remote patient monitoring could bring 3,700 USD savings per patient per year. Genenia claims this figure is actually more than 8,000 USD, while a pilot at Christus St. Michael Health System showed savings of more than USD 10,000 per patient (although over how long is not clear, and concentrates on elderly with chronic diseases so is likely to skew the results in favour of cost savings). A Deloitte report also claims a savings of between roughly 1,000 and 2,000 USD per year per patient.

All of these examples are in the US, and it is worth noting that the health expenditure in the US is disproportionately high compared to other nations, as the data from OECD below shows. Remote patient monitoring services are most developed in the USA but even with this in mind, when considering the saving globally the figure is likely to be lower.

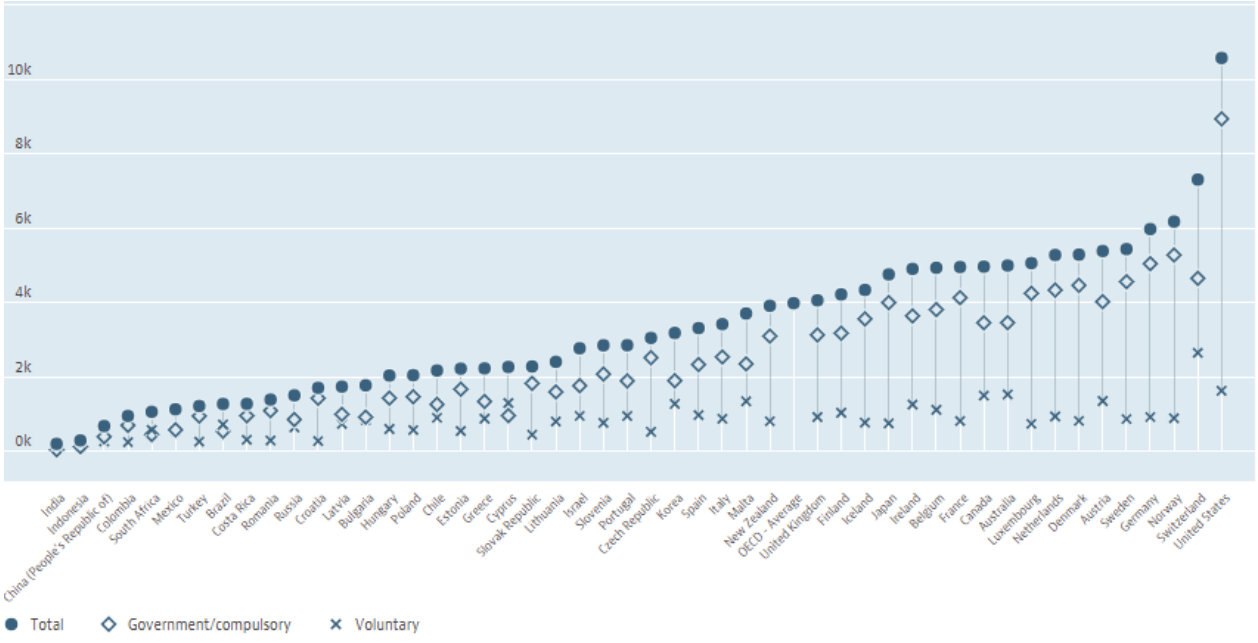


Figure 27: Financing of health expenditure by country, latest data possible (Source: OECD Data)

Thus, taking a cautious approach of an estimation of 1,000 USD savings per year per patient seems reasonable and this will be used in the calculation of this report.

Most of it is related to data-based process (patients still need sensors/devices) and can therefore be considered as data benefits from IoT.

Conclusion: Cost savings of 3.36%, or 271.8 billion EUR, can be obtained in health industry through data

Through IDATE DigiWorld's IoT reports, the forecast for 2025 is that there will be more than 302 million connected remote patient monitoring devices installed.

In particular, the highest savings generated by telecare solutions in 2025 will be in the following countries: in the UK (6.5%), in the France (5.7%) and in Spain (5.3%).

The average savings in top-10 countries by healthcare expenditure will be 3,6% in 2025. (Thus, multiplying this by the 1,000 USD savings per patient per year, this gives **a final estimate of 303 billion USD (271.8 billion EUR) savings in 2025**.

If we estimate the total health expenditure to be 9 trillion USD in 2025 (10% of 90 trillion global GDP), this gives a final result of **IoT data through remote patient monitoring to provide 3.36% savings of the total health expenditure in 2025**.

Table 8: Breakdown of cost saving calculation through IoT data for health industry (Source: IDATE DigiWorld)

Service	Number of connections 2025	Cost savings per patient	Total savings 2025	As % of total health market
Remote patient monitoring	302.6 million	850 EUR	271.8 billion EUR	3.36%

This overall calculation is in line with analysis of health care national plans leveraging digital technologies.

IDATE Digiworld conducted a benchmark of 10 advanced countries in the world in the Smart Ageing report in 2019 [33]. The analysis showed that 1,5% reduction of hospitalization fees in 2018 thanks to telecare, expected to reach 3% by 2023.

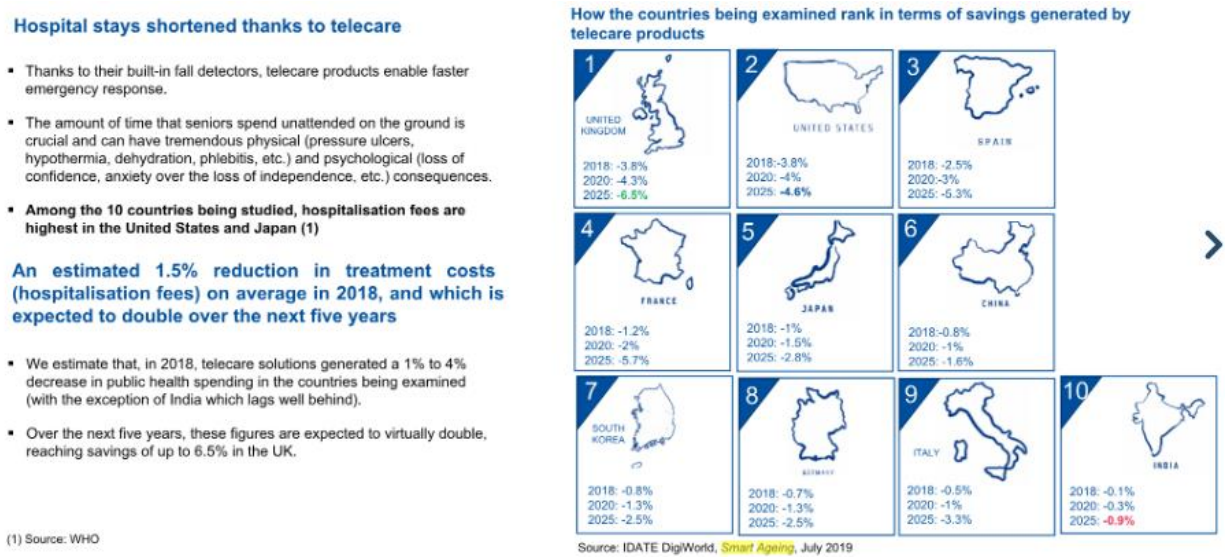


Figure 28: Estimated savings generated by Telecare solutions (Source: WHO)

The report also analysed gains from telemedicine, but most savings will come from reduction of trips. Therefore, savings are related to connectivity rather than IoT data.

2.3.4.2 New revenues for 2025

Connected medical devices

From the IDATE DigiWorld IoT reports, the estimated number of connected medical devices stands at 302.6 million units for 2025, up from 46 million units in 2018.

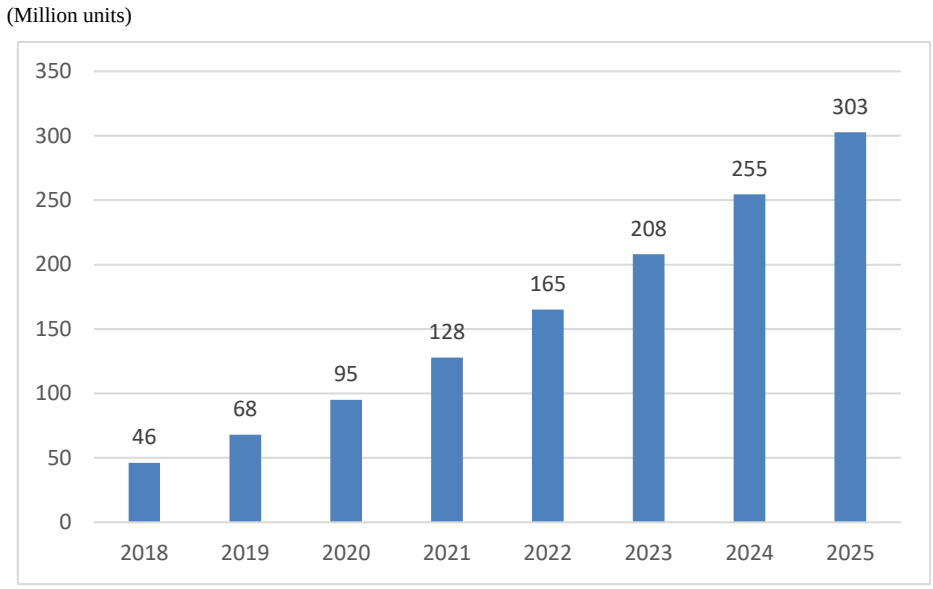


Figure 29: Global installed base of connected healthcare devices 2018-2025 (Source: IDATE DigiWorld [32])

New revenue generation through remote monitoring services

Looking at connected remote patient monitoring services available today, the ARPU varies around 30 EUR per month (it used to be between 30 and 50 EUR a few years ago, but has gone down with the start of the massification), including the device fees as well as the connection fees. The variant in the pricing is largely due to the scope of service coverage, but basic connectivity, i.e. the use of IoT data is included to some extent for virtually all cases.

The difficulty of the health vertical today comes from the fact that it is currently in a time of transition; what was long a vertical belonging to just the medical players is now seeing new digital players coming into play. Going forwards, revolution in traditional healthcare system is indispensable, to provide an adapted environment to accommodate the transformation brought by a variety of eHealth applications. Digital players play a key role in intermediating device-generated data with other stakeholders through open APIs, such as the Apple HealthKit. In addition, they also actively involve in the development of data-driven applications for health-related researches.

Still, medtech giants and traditional telemedicine companies still maintain their strong position in the eHealth realm with comprehensive offerings covering full blocks of value chain, to name a few, Philips, Medtronics and Tunstall. Taking this into account, the health industry as a whole is expected to take 60% of the new services' revenue share, with the rest being split among the digital players, connectivity players, integrators and so on. Within that 60%, a large part of the revenues is related to devices. We assume therefore only half of the RPM revenues can be considered as software/data.

Example of IoT in healthcare: Philips HealthSuite digital platform

The Philips HealthSuite digital platform, supported by Salesforce, is open and cloud based. By leveraging the Salesforce1 platform, it securely collects, integrates and analyses data from multiple sources - a variety of medical and personal devices, as well as health systems such as EMRs, imaging and monitoring data.

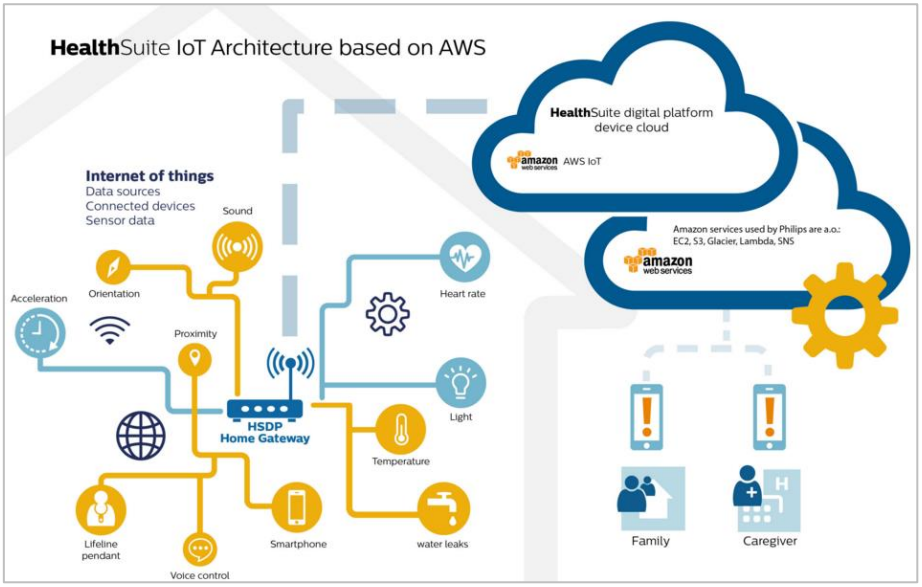


Figure 30: Philips HealthSuite digital platform (Source: Philips, 2017)

Conclusion: IoT data to generate 32.5 billion EUR worth of new revenues, or 0.35% of the total market in 2025 in the health vertical

In conclusion, taking the monthly ARPU of remote patient monitoring services to be 30 EUR on average, and assuming the health industry takes 30% of the revenues, this **provides new revenue generation through IoT data of 32.5 billion EUR in 2025**.

Table 9: Breakdown new revenue calculation through IoT data for health industry

Revenue	As a share of total health industry	Value (billion EUR)
Remote patient monitoring	0.35%	32.5

3. BASELINE REQUIREMENTS FOR AN IOT DATA VALUE CHAIN

Based on the earlier discussion, as well as on, the existing literature expanding on the requirements for an IoT data value chain model, this section will produce a preliminary IoT Data Model. In essence, this chapter suggests using the traits, and the requirements of IoT Data Value Chains, to extract the relevant properties. The discussion presents an abstract model for reliable IoT Data Value Chains aiming to support and further unleash the potential of existing and future IoT value chains. Building on the earlier analysis, the discussion below addresses the attributes of economical and legal relevance.

Note that the attributes below are relevant for the associated IoT value chains across all Large-Scale Pilots (LSPs).

3.1 The benchmark scheme for an IoT Value Chain Data Model

The digital society is converging with a consumer-industrial-business Internet that is based on hyperconnected IoT environments.

The data flow and exchange across the IoT architectural layers is reflected in the IoT Data Value Chain that includes the following processes as presented in delivery D05.03: data acquisition, data transmission/ingestion, data processing, data storage, data filtering, data analysis/analytics, data integration, data discovery, data usage, data exposure (openness), and data monetization.

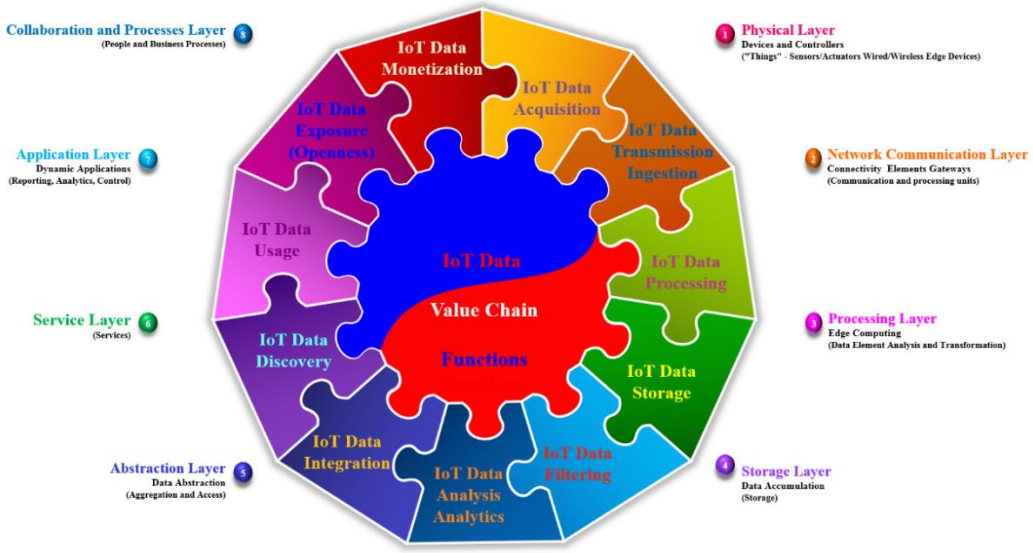


Figure 31: IoT data value chain functions

- **IoT Data acquisition** – addresses the process of gathering, and formatting IoT data before it is transmitted through different channels/pipelines for ingestion and processing. Data acquisition is one of the major IoT data challenges in terms of infrastructure requirements and edge heterogenous devices/nodes/things.
- **IoT Data transmission/ingestion** – addresses the communication channels and pipelines for transmitting the IoT data and the ingestion of data for enabling reliable operation of entire IoT platforms using various file formats and network connections while considering frequency, volume, data rates, neutrality, etc.
- **IoT Data processing** – addresses the processing of IoT data form different sources (sensors, actuators, processes, virtual things, etc.) for transforming the data into a format that facilitates its reuse or enables immediate action based on incoming events and interactions.
- **IoT Data storage** – provides cost-effective ways for distributed IoT data storage with the choice of format or database technology determined by the nature of other stages in the value

chain (i.e. analysis, analytics, nature of applications – safety/mission critical). The IoT data storage assures the persistence and management of IoT data in a scalable way that satisfies the needs of IoT applications that require fast access to the raw or processed IoT data.

- **IoT Data filtering** – concerns the active management of IoT data over its life cycle to ensure it meets the necessary data quality requirements for its effective usage in various IoT applications across different industrial domains. The IoT data filtering processes include different activities i.e. content creation, selection, classification, transformation, validation, preservation, etc.
- **IoT Data analysis/analytics** – addressed at every layer in the IoT architecture and every step in the IoT data value chain, allowing the generation of new insights and actions based on the IoT data from various sources and enabled by the tools and IoT platforms used in different applications. IoT Data analysis transforms the raw IoT data into data for use in the decision-making as well as domain-specific usage, through exploring, transforming, and modelling IoT data with the goal of extracting the relevant "smart" data, synthesising and extracting useful "invisible" information with high potential from an IoT application point of view.
- **IoT Data integration** – combining a variety of IoT data sources to provide new insights, with IoT data integration as a key element for any IoT application.
- IoT Data discovery – addresses the localization and identification of IoT data sources need and the evaluation for different attributes, relevance, quality, integrity, security, privacy, cost, coverage, etc.
- **IoT Data usage** – considers the IoT data-driven applications that need access to IoT data, its analysis, and the tools and IoT platforms needed to integrate the data analysis within the different IoT applications and use cases. IoT data usage in use cases/applications/scenarios decision-making enhances effectiveness through reduction of costs, increased added value, portability, etc.
- **IoT Data exposure (openness)** – addresses how IoT data that are exposed to the other IoT applications and IoT ecosystems stakeholders in a way that makes them useful for value co-creation in order to generate value from IoT data from various edge and platforms sources.
- **IoT Data monetization** – addressing the IoT business models that support IoT ecosystems for determining the value of IoT data provided by different sources and available from different IoT applications and creating new opportunities for growth and economical and social benefits.

Note that the identification of the aforementioned processes should be placed within the scope of the IoT Policy Framework presented under D05.01 IoT Policy Framework and D05.02 IoT Policy Framework Evaluation & Final IoT Policy Framework.

3.2 Liability and transparency

Liability and transparency that can be of critical significance for an IoT Data Value Chain Model that aims for at a trustworthy and above all human centred IoT ecosystem.

3.2.1 Liability

Liability forms a legal concept closely linked to the notion of responsibility used across disciplines, yet substantially different.

The notion of liability implies an IoT stakeholder's legal responsibility for his actions and/or omissions. Depending on the point of view, liability may also be perceived as burdensome or discouraging for the stakeholder. As it has been argued in this respect, "Liability is the legal obligation (either financially or with some other penalty) in connection with failure to apply the norms" [18]. Interestingly and liability links to the notion of responsibility, although being held liable does not necessary presume actual responsibility.

At this moment of convergence of technologies, markets and stakeholders, the attribution of liability becomes more complex to answer compared to the physical society. For example, a manufacturer of certain objects has to accept and address its respective and proportionate responsibility in the IoT ecosystem its objects are deployed [1]. IoT will bring more responsibility for each stakeholder in the market, and each of such stakeholders will have to think and arrange for those effects in a transparent, diligent and ethical manner. Another example is a security breach in an IoT ecosystem as per insecure coding of software somewhere in the multi-angled value chain [1]. As long as related software companies cannot be held liable, a solid and stable digital economy and society will be difficult to create.

Overall, though, it should be noted liability is of paramount importance with the sphere of rule of law, as it forms the emerging consequence of legal obligations created within the context of legal and contractual relationships; the latter are discussed under D05.01 IoT Policy Framework, while being at the heart of the research falling under Work Package 5 on "IoT Policy Framework - Trusted, Safe and Legal Environment for IoT".

3.2.2 Transparency

Transparency has been defined as "the property of a system, organization or individual that provides visibility of its governing norms, behaviour and compliance of behaviour to the norms." [17]. Furthermore, transparency forms one of the essentials of good governance as it allows individuals, organizations, society at large to assess risks and benefits and proceed in making the appropriate decisions.

Transparency can either relate to sharing of information or, more broadly, to the adoption of a certain behaviour, before the occurrence of an unwanted event or it may concern informing on the associated consequence, after a certain event has already taken place. In view of materializing transparency in the IoT ecosystem, there are different mechanisms that can be of help to this end, including, the publication of transparency reports by the cloud stakeholders on annual basis and the conclusion of contractual agreements meeting certain criteria that will be briefly discussed below.

In particular, contracts regulating the relationships between IoT stakeholders should provide for a Data Management Service Level Objectives Overview, along the lines of the Service Level Objectives (SLOs) introduced by the Cloud Standardization Guidelines [20]. The appropriate data management SLOs could be assigned with a complementary function to the applicable security and data protection certifications afforded to the IoT stakeholders. Such an approach would mandate the provisioning of SLOs linking to four distinctive categories, namely, a) Data Classification, b) Data Mirroring, Backup & Restore, c) Data Lifecycle and d) Data Portability that can be further subdivides to further categories.

Data classification refers to the detailed description of the classes of data involved to the provisioning of a specific service by a specific cloud stakeholder that may include, among other-cloud service customer data, cloud service provider data and cloud service derived data[10]. Should Service Level Agreements provide in a clear manner for the relevant SLOs linking to the relationships between the IoT stakeholders, consumers and organizations would be better equipped to make informed decisions, thus, competition would be significantly boosted.

As far as the data mirroring, backup and restoration is concerned, this category refers to the actual mechanisms guaranteeing the online or offline availability of data, in case of failures impeding access to it. These mechanisms falling under the scope of this SLO can be further divided in two widely used categories (i) data mirroring, (ii) backup/restore.

[10] Note that the relevant definitions are provided under the above-mentioned Cloud Standardization Guidelines.

Data Mirroring refers to the difference between the time data is placed on primary storage and the time the same data is placed on mirrored storage. Furthermore, data back up and restoration refers (primarily) to the list of method(s) employed to backup data and to the time of completion of the backups.

As to the data lifecycle, it refers to the effectiveness of the IoT stakeholder's data lifecycle practices, with a special focus on the practices and mechanisms for data handling and deletion.

Finally, data portability, which is highly relevant for the free flow of data within the EU and the strengthening of the Digital Single Market, involves the specification of the data portability format, of the data portability interface and of the data transfer rate [21], as further discussed under the previously mentioned Cloud Standardization Guidelines.

The discussion above aimed at serving as an example on how to increase transparency of IoT Data Value Chains through the provisioning of specific SLOs in the contractual arrangements regulating the relationships between the various IoT stakeholders. Although the discussion has been largely inspired by the cloud environment, it remains relevant for the IoT ecosystem as well.

Bearing in mind the aims of the present deliverable, it should be, thus, noted that the importance of SLOs with respect to transparency of IoT Data Value Chains does not lie so much in the concrete SLOs [20] deemed relevant in the context of specific relationship between cloud stakeholders, but rather on their determination and incorporation per se under a specific contractual agreement.

3.3 Open data marketplaces

Policies to support open marketplaces are highly relevant, and a series of IoT events related to data marketplaces have been arranged. The Workshop on Policies to Support Open Data Marketplaces - Data Sharing in IoT Ecosystems, Data-supported Services Concepts & Best Practices [39] is a part of this series. The event is organised by the CREATE-IoT project as part of the IoT Large-Scale Pilots Programme with support from DG Connect and Alliance for IoT Innovation (AIOTI) as the co-hosts of this event. The workshop will look into opportunities for data-driven services in different sectors/domains like mobility, farming and energy. Use cases from the IoT LSP projects will be exploited and key elements of a striving data economy in those sectors extracted, most important to present the underlying architecture design and setting the standards for fair level playing field whilst supporting innovative business models.

It is important to identify the main challenges for IoT data value chain, regarding data-driven services, data flows including data classification, data sharing, data life cycle, digital rights management, personal data protection and present best practices from different IoT LSP projects and sectors/domains [39]. Special attention should be paid to building consensus among stakeholders in vertical and horizontal value chains on principles of data sharing, possibly paving the way for the establishment of a code of conduct applicable across all sectors/domains. Regulatory issues across different sectors/domains including links to national initiatives to achieve unified solutions. Examples on national initiatives are: The Data Market Austria project is partially founded by the "ICT of the Future" Programme of the Austrian Research Promotion Agency (FFG) and the Austrian Ministry for Transport, Innovation and Technology (BMVIT) [40]; and the Dutch Digitalisation Strategy - Dutch vision on data sharing between businesses, from the Ministry of Economic Affairs and Climate Policy [41].

Both governments and the industry has taken action by putting forward appropriate instruments, in response to the changing market dynamics [39]. With legislation as the GDPR and the Free Flow of Non-Personal Data Regulation, as well as by updating sector specific regulations such as the Payment Service Directive 2 (PSD2) relevant for the finance sector governments have taken first steps to foster a vibrant data economy that provides for equal opportunities for all market players. Also, certain markets have demonstrated interest by taking sector specific initiatives, such as Energy Data Access Alliance - Connecting Europe's energy data [42], and European

Agricultural Machinery Association - the adoption of a Code of Conduct on agricultural data sharing [38][43] (see section 3.4).

3.4 Principles of data sharing, code of conduct models

Free flow of non-personal data is a key element for a competitive data economy within the DSM and there is a need to provide the mechanisms and the regulatory framework to ensure a free flow of data, allowing companies and public administrations to store and process non-personal data wherever they choose in the EU.

An example of attempt to define such a framework is the "EU Code of conduct on agricultural data sharing by contractual agreement" by Copa and Cogeca, CEMA, Fertilizers Europe, CEETTAR, CEJA, ECPA, EFFAB, FEFAC, and ESA [37][38]. Accurate agricultural data availability is vital to develop digital farming enabling farmers to produce more using less resources. In order to fully utilize the benefits of digital farming, sharing data between different partners in the agro-food chain must be conducted in a fair and transparent way.

Access to the necessary data will facilitate and accelerate data driven business models. This framework aims to create trust among partners by setting the transparent principles and clarifying the responsibilities. This framework is not a legal document, but works as a guideline combined with a check list [38]:

- Is there an agreement/contract in place?
- What obligations are there? What warranties and indemnities are there for each party?
- What data is collected?
- Who owns/controls access to the data?
- What services are delivered?
- Will my data be used for purposes other than providing me, the data originator (e.g. farmer) a service? Is it clear what these are? Can I agree/disagree? What are the benefits/value for me as data originator?
- Is the data shared with other parties? What rules do the external parties adhere to? Can I agree/disagree with sharing data with other parties?
- Can the service provider change the agreements unilaterally?
- What happens when the service provider changes ownership?
- Can I retrieve my dataset from the system in a usable format?
- Will I be updated on security breaches?
- Can I opt out of the service and have my data deleted from the system?
- Is there a contact point to assist me with any question that I may have?
- Do I need insurance?
- What are the confidentially terms?

3.5 GAIA-X initiative

Addressing the data value chain implies focusing as well on connected data infrastructure that deliver on the non-physical level that will connect with the IoT devices used in different industrial sectors. In principle, a digital data infrastructure consists of three architecture levels[36]:

- Network level – the data transfer networks and related hardware.
- Data level – the data storage level, including operating system and databases for storage.
- Service level – the data processing and use level, including applications systems, functions and services.

The global public cloud computing market is dominated by non-European players like Amazon Web Services (47,8%), Microsoft Azure (15,5%), Alibaba (7,7%), Google (4%) and IBM (1,8%)

[34]. However, geopolitical tensions are making the European politicians and companies cautious about storing data in the cloud on non-European servers.

GAIA-X is a highly ambitious cloud project initiated in Germany [35]. Germany together with France plan to launch a European cloud service platform in 2020, open for both European and foreign countries following the projects rule on data sovereignty [34].

GAIA-X would connect various cloud providers across Europe using open standards, allowing businesses and consumers to move their data around freely which is important for Europe's digital and technological autonomy and a vibrant European Ecosystem [34].

The deployment is guided by the following principles [36]: European data protection; Openness and transparency; Authenticity and trust; Digital sovereignty and self-determination; Free market access and European value creation; Modularity an interoperability; and User-friendliness.

Figure 32 illustrates the overall GAIA-X data infrastructure and ecosystem [36].

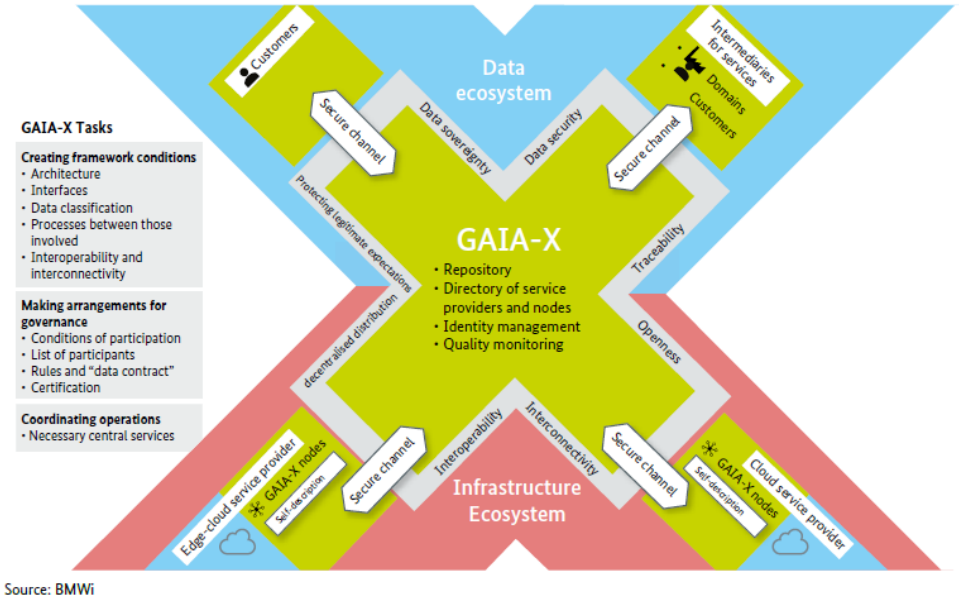


Figure 32: GAIA-X data infrastructure and ecosystem [36]

The ecosystem is based on a federated data infrastructure that promote cooperation through integrating specific strengths of various participants, and the users obtain access to relevant products/services [36]:

- The infrastructure integrates existing digital and cloud-based state-of-the-art products/services;
- The infrastructure offers full transparency by providing authentication on verified data protection and regulatory criteria of the products/services offered;
- The infrastructure simplifies the management of interfaces/integration, especially regarding multi-cloud strategies and data-pooling;
- The infrastructure allows the user to retain command over particularly sensitive data, while simultaneously sharing other data with partners for joint use;
- The infrastructure creates the preconditions for optimising the users' data strategies, (decentralised/centralised cloud infrastructure nodes can be linked-up with one another and generate options regarding how data and algorithms can be used securely);
- The infrastructure enables digital ecosystems in various user domains to make the transfer from bilateral individual-project solutions to marketplace solutions, (standardised contracts/procedures reduce transaction costs, data markets can emerge, and data availability is improved).

4. CONCLUDING REMARKS

Valuable insights to the IoT Data Value Chain Model Evaluation and Final IoT Data Value Chain Model creation are presented. Recognising the complexity of IoT environments, this document has discussed the emphasis placed by IoT data value chains on the potential of data for the economy and society at large. Although the title of the deliverable suggests the discussion of data value *chains*, the actual discussion revealed that the focus should rather be on the data *value ecosystems,* given that data value chains are basically converging to value networks and more broadly to IoT ecosystems. Furthermore, the discussion has reaffirmed the *n*-dimensional nature of the subject which also accounts for the challenges faced when capturing and documenting its individual features and properties. While examining the subject from an architectural as well as model perspective, it has become clear that the notion of *context* is of paramount importance with respect to the perception of data in the IoT environment.

Furthermore, it has been highlighted how the particularities of the IoT Value Chains based upon the dynamic flows of information render the attribution of responsibility across the supply chain is highly complex, thus, further challenging traditional concepts, such the concept of liability. Regulations and directives such as the General Data Protection Regulation (GDPR), the Directive on security of network and information systems (NIS Directive), the Regulation on the Free Flow of Data, and the Public Sector Information (PSI Directive) are creating/will create an extensive impact for IoT Data Value Chains. In response to the changing market dynamics, both governments and the industry have taken action by proposing legislation such as the Free Flow of Non-Personal Data Regulation, as well as by updating sector specific regulations to foster a vibrant data economy that provides for equal opportunities for all market players. In addition, certain markets have taken sector specific initiatives, such as the adoption of a Code of Conduct on agricultural data sharing.

The IoT Value Chain and data monetization for different industrial vertical sectors has been addressed. The enablers of data monetisation solutions are technologies such as business intelligence, data mining, smart data and deep analytics through artificial intelligence (AI). Smart data is the most innovative set of technologies among analytics technologies. IoT data monetization in automotive and healthcare are elaborated through market overview, data availability, drivers, barriers, cost savings and new revenue opportunities. These sectors are represented in CREATE-IoT and are two important examples of verticals that have potential benefits of utilizing IoT data.

Free flow of non-personal data is a key element for a competitive data economy within the Digital Single Market (DSM) and there is a need to provide the mechanisms and the regulatory framework to ensure a free flow of data, allowing companies and public administrations to store and process non-personal data. The "EU Code of conduct on agricultural data sharing by contractual agreement" (by Copa and Cogeca, CEMA, Fertilizers Europe, CEETTAR, CEJA, ECPA, EFFAB, FEFAC, and ESA) is an example of attempt to define such a framework.

Addressing the data value chain implies focusing as well on connected data infrastructure that deliver on the non-physical level that will connect with the IoT devices used in different industrial sectors. The global public cloud computing market is dominated by non-European players like Amazon Web Services, Microsoft Azure, etc., and geopolitical tensions are making the European politicians and companies cautious about storing data in the cloud on non-European servers. GAIA-X is a highly ambitious cloud project initiated in Germany, and together with France plan to launch a European cloud service platform in 2020, open for both European and foreign countries following the projects rule on data sovereignty.

Also, based on the earlier analysis, the discussion produced a first set of attributes for an IoT Data Value Chain Model drawing links with the overarching IoT Policy Framework discussed under D05.01 and the concrete processes entailed. In this context, transparency surfaced as a key attribute given the labyrinth of contracts and the need of IoT stakeholders to be properly informed through

the proposed introduction of specific Service Level Objectives (SLOs). Transparency, though, also, implies the clear allocation of the roles of controllers and processors to those entities handling personal information, as being emphasized as well by the GDPR.

Overall, it should be noted that the creation of an IoT Data Model does not constitute a theoretical exercise; it rather forms a challenge of high practical significance and of value from a governance standpoint, as it can increase control within the fluid IoT ecosystems and augment the IoT benefits for the entire spectrum of the IoT stakeholders.

5. REFERENCES

[1] O. Vermesan and J. Friess (Eds.), *Digitising the Industry - Internet of Things Connecting the Physical, Digital and Virtual Worlds*, ISBN: 978-87-93379-81-7, River Publishers, Gistrup, 2016.

[2] A. van der Wees, J. Breeuwsma and A. van Sleen, "IoT Societal Impact – Legal Considerations and Perspectives" in O. Vermesan and J. Friess (Eds.), *Digitising the Industry - Internet of Things Connecting the Physical, Digital and Virtual Worlds*, ISBN: 978-87-93379-81-7, River Publishers, Gistrup, 2016, River Publishers Series in Communication, Volume 49, Chapter 7, page 220.

[3] IDATE DigiWorld.

[4] IDATE DigiWorld, Connected Healthcare, June 2016.

[5] A. van der Wees, J. Breeuwsma and A. van Sleen, "IoT Societal Impact – Legal Considerations and Perspectives" in O. Vermesan and J. Friess (Eds.), *Digitising the Industry - Internet of Things Connecting the Physical, Digital and Virtual Worlds*, ISBN: 978-87-93379-81-7, River Publishers, Gistrup, 2016, River Publishers Series in Communication, Volume 49, Chapter 7, page 233.

[6] Proposal for a Regulation of the European Parliament and of the Council on a framework for the free flow of non-personal data in the European Union.

[7] Cloud Accountability Project, "D-4.4" Remediation guidelines and tools, 2015.

[8] G. Noto La Diega and I. Walden, "Contracting for the 'Internet of Things': looking into the Nest", in European Journal of Law and Technology, Vol 7, No 2, 2016.

[9] G. Noto La Diegaand I. Walden, "Contracting for the 'Internet of Things': looking into the Nest", in European Journal of Law and Technology, Vol 7, No 2, 2016.

[10] M. E. Porter, Competitive advantage: Creating and sustaining superior performance. New York: Free Press, 1985. doi:10.1182/blood-2005-11-4354.

[11] European Commission DG CONNECT, A European strategy on the data value chain, online at: ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=3488, 2013.

[12] European Commission, Towards a thriving data-driven economy, Communication from the commission to the European Parliament, the council, the European economic and social Committee and the committee of the regions, Brussels, 2014.

[13] O. Vermesan and J. Bacquet (Eds.), *Cognitive Hyperconnected Digital Transformation Internet of Things Intelligence Evolution*, ISBN: 978-87-93609-10-5, River Publishers, Gistrup, 2017.

[14] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation).

[15] Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union.

[16] Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products.

[17] http://www.a4cloud.eu/lexicon/glossary/letter_t.

[18] Cloud Accountability Project, "D:C-2.1 Report detailing conceptual framework", 2014.

[19] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

[20] The Cloud Select Industry Group, "Cloud Service Level Agreement Standardisation Guidelines", 2014.

[21] SMART 2016/0032 Study, IDC and Arthur's Legal, "Switching between Cloud Service Providers", 2017 (Ongoing).

[22] IoT Policy Framework (D05.01). CREATE-IoT, October 2017. Online at: https://european-iot-pilots.eu/wp-content/uploads/2017/10/D05_01_WP05_H2020_CREATE-IoT_Final.pdf
[23] IDATE DigiWorld, Data monetisation, October 2016.
[24] IDATE DigiWorld, The digitisation of the automotive industry, November 2016.
[25] Otonomo-Edison Research Consumer Survey, June 2018.
[26] Edison Research for Otonomo, June 2018.
[27] Marketing Land, October 2019.
[28] Deloitte, Automotive Data Treasure, 2017.
[29] IDATE DigiWorld, State of OTT markets worldwide, November 2019.
[30] IDATE DigiWorld, Connected Healthcare, December 2017.
[31] IDATE DigiWorld, Data monetisation, October 2016.
[32] IDATE DigiWorld, December 2019.
[33] IDATE Digiworld, Smart Ageing report, 2019. Online at: https://fr.idate.org/produit/smart-ageing/
[34] Gyles, S. (Main author). European Cloud Service Gaia-X in the Making. VPN Overview, November 2019. Online at: https://vpnoverview.com/news/european-cloud-service-gaia-x-in-the-making/
[35] Will Gaia X Become THE Public Cloud for the European Market? Hosting Journalist, November 2019. Online at: https://hostingjournalist.com/cloud-hosting/will-gaia-x-become-the-public-cloud-for-the-european-market/
[36] Project GAIA-X, A Federated Data Infrastructure as the Cradle of a Vibrant European Ecosystem. German Federal Ministry of Economic Affairs and Energy and Federal Ministry of Education and Research, October 2019. Online at: https://www.bmwi.de/Redaktion/EN/Publikationen/Digitale-Welt/project-gaia-x.pdf?__blob=publicationFile&v=4
[37] Food and Agriculture Organization of the United Nations. Family Farming Knowledge Platform. EU Code of Conduct on agriculture data sharing. Xxx. Online at: http://www.fao.org/family-farming/detail/en/c/1127623/
[38] EU Code of conduct on agricultural data sharing by contractual agreement, 2018. Online at: https://www.copa-cogeca.eu/img/user/files/EU%20CODE/EU_Code_2018_web_version.pdf
[39] IoT European Large-Scale Programme. Policies to Support Open Data Marketplaces - Data Sharing in IoT Ecosystems, Data-supported Services Concepts & Best Practices. Online at: https://european-iot-pilots.eu/2020-january-hague/
[40] Data Market Austria project. Online at: https://www.ait.ac.at/en/research-topics/data-science/projects/dma/
[41] Dutch Digitalisation Strategy - Dutch vision on data sharing between businesses, February 2019. Online at: https://www.government.nl/documents/reports/2019/02/01/dutch-vision-on-data-sharing-between-businesses
[42] Energy Data Access Alliance. Connecting Europe's energy data. Online at: https://www.dataalliance.eu/
[43] European Agricultural Machinery Association - EU Code of Conduct on agricultural data sharing. Online at: https://cema-agri.org/publications/19-brochures-publications/37-eu-code-of-conduct-on-agricultural-data-sharing

6. ANNEXES

6.1 The ecosystem backgrounds

IoT applications are seen as highly complex supply chains which connects an unlimited number of various devices together making it possible for the devices to communicate and operate through different infrastructures across various supply chain layers. As the supply chains extend across borders and industry sectors and domains, this supply chain in the existing and rapidly developing hyperconnected world is no longer linear and they are transformed in value chains and value networks.

In this context, the relations between the developers, vendors, consumers and other stakeholders of the digital economy and society (including but not limited to IoT enabled devices, systems or services) are becoming non-linear and changing dynamically.

This creates an extensive multi-dimensional system which can also be referred to as the *chain ecosystem*. Every participant within this multi-dimensional ecosystem is relevant and plays an important role in the design, engineering, manufacturing, deploying and functioning of both a connected device, system and service, as well as hyperconnected (IoT) ones [1]. Figure 33 presents a supply value chain in two dimensions (2D):

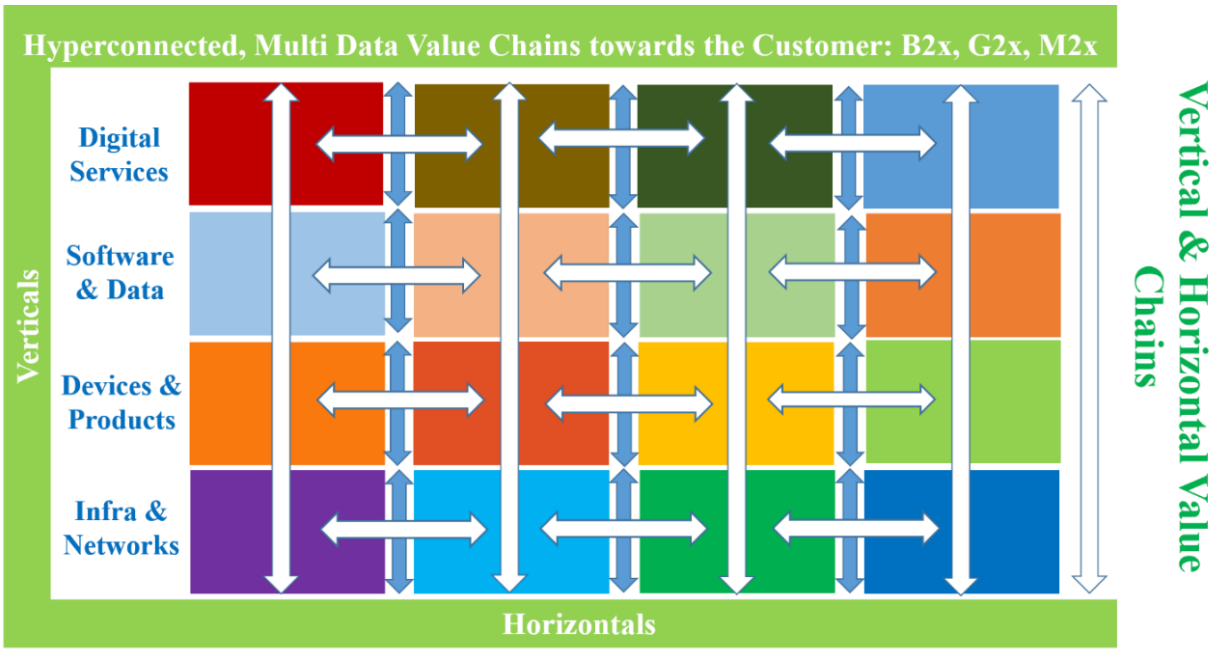


Figure 33: Hyperconnected, vertical and horizontal value chain [1]

The Figure 33 illustrates the complexity of the ecosystem. If a consumer is placed at the very end of the supply chain (downstream), it becomes apparent that numerous different parties located up the stream participate in manufacturing and assembling of software and hardware parts, as well as the functioning of the device on the digital network.

In addition, IoT devices themselves also represent highly complex value chains, connecting various hardware and software components together, while being connected to and communicating with one or more networks and other devices.

This accounts for situations in which a user may not always have a good and complete understanding of what actions the device carries out, how it works and more importantly, what a device is capable of doing and the way it works within the ecosystem. As a result, situations may

arise in which damage occurs without the consumer even knowing how the damage has occurred and what the cause has been.

The data value chain reflects not only the existing value, but also the value that can be derived for economy and society at large. In other words, data value chains place emphasis on the potential of data, rather than on the data per se, that may result to the addition of several layers of value on top of the original raw data, both private and public.

Given the central role of data for IoT data value chains, it should be noted that this document endorses the definition of data under ISO/IEC 2382-1, considering data as "a reinterpretable representation of information in a formalized manner, suitable for communication, interpretation or processing."

It is, thus, needed that *"Data should not be treated as a four-letter word. The concept of data encompasses data of any form, nature or structure, that can be created, uploaded, inserted in, collected or derived from or with cloud services and/or cloud computing, including without limitation proprietary and non-proprietary data, confidential and non-confidential data, non-personal and personal data, as well as other human readable or machine-readable data"* [20].

Moreover, given that data value chains are non-linear, there can be continuous use and re-use, which creates a series of challenges of legal and strategic relevance to be partly discussed under this deliverable as well as under the forthcoming deliverables due under "Work Package 05: IoT Policy Framework - Trusted, Safe and Legal Environment for IoT." This aspect of non-linearity allowing continuous use and re-use is captured by the concept of "lifecycles" that will be extensively used in the present analysis in relation to devices, stakeholders, data and law.

In particular, the **IoT Device/Product Life Cycle** is used to capture whether and how long a device/product can remain connected to an IoT ecosystem in a secure, safe and compliant manner. It also refers to what the user/customer expects, and how both the device/product and the user/customer are able to keep up to date with (at least) the state of practice. The different stages of the IoT device lifecycle are further captured in the figure below.

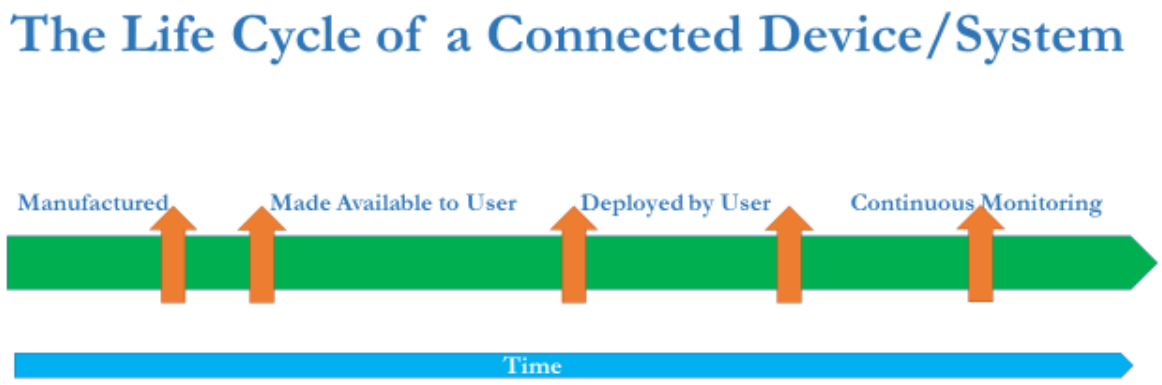


Figure 34: The Life Cycle of a Connected Device/System

The **Stakeholders Life Cycle** refers to the entire spectrum of stakeholders with a role in relation to an IoT device/product. This lifecycle provides the ground for attribution of responsibilities – also, allowing for changes in initial dynamics - and, consequently, what happens in case an incident occurs within an IoT ecosystem. Stakeholders Lifecycle also implies keeping stakeholders up to date.

The **Data Life Cycle refers to the data collected,** created or otherwise concerned and to the way that these data can be segmented, minimised and isolated. The data lifecycle further paves the ground to determine what happens if data have multiple classifications or if these classifications change. The figure below captures the distinctive phases of the personal data lifecycle.

Contextual Life Cycle refers to the specific context that a device/product is being used and the persona under which a stakeholder acts with respect to the specific device/product. The contextual lifecycle is taken into account when allocating responsibilities (e.g. who is accountable in what context) and, similarly, how to make stakeholders aware of changes in best practices, rights and obligations when the context changes, as well as how to secure the rights and obligations of the other relevant stakeholders.

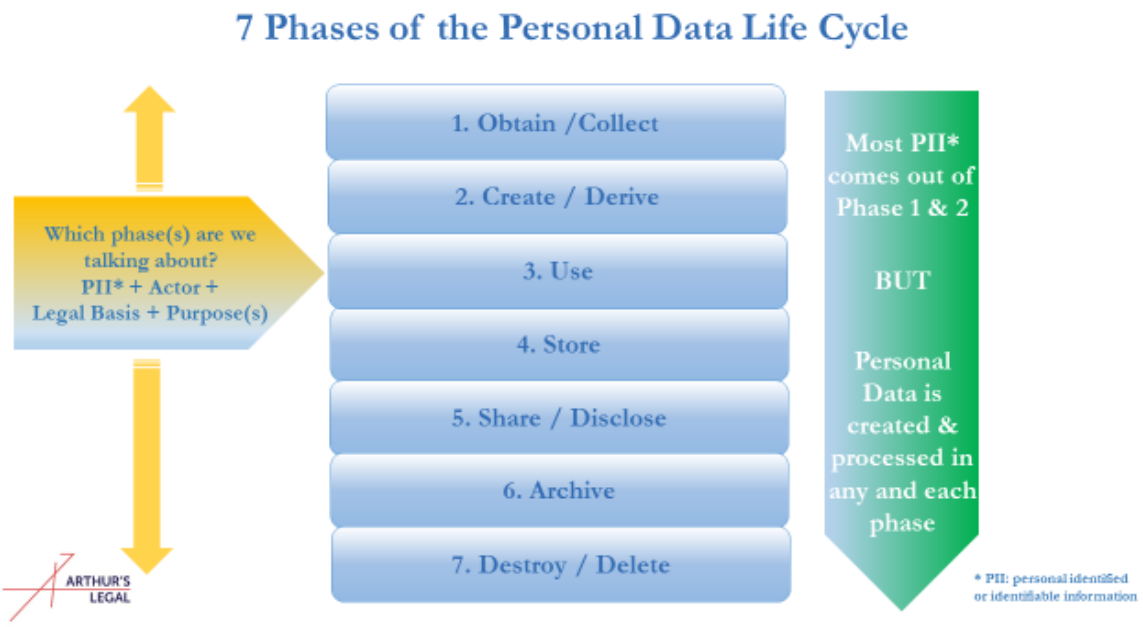


Figure 35: 7 Phases of the Personal Data Life Cycle

Finally, **the Legal Life Cycle should be conceived on the basis of all the earlier mentioned lifecycles.** It refers to person or legal entity, with whom another person or legal entity wishes to engage and, if this is the case, it will include the steps taken to assess, prepare, negotiate, conclude, execute, operate, update, amend, escalate and terminate such an engagement and, in essence, the legal relationship.

Note that there are natural interdependencies in the IoT environment between all the earlier stated lifecycles, in the sense that the issues that impact upon on the IoT data lifecycle are relevant for the stakeholders' lifecycle; changes in data classification may create an impact on the distribution of responsibilities between stakeholders. The same applies to the responsibilities emerging from the meta-data and derived data.

6.2 The main traits of the IoT value chain

This part presented in the delivery D05.03 deconstructs the IoT Value Chain into its constitutive chains, expanding on the markets involved revealing the economic value chain and identifying the set of actors involved.

The analysis explains the data relation flows involved and briefly touches upon how these chains may interact.

6.2.1 Chain of markets

As mentioned earlier, the value chain of data within the IoT field can be observed from two different angles; a horizontal approach and a vertical approach.

The former approach could be thought of as the conceptual framework for IoT data (in this case the data being created through IoT platforms), which provides the foundations for how data is passed along the chain.

The latter approach looks at the specific value chains in in any given vertical. Both approaches will be considered further in the paragraphs that follow.

6.2.1.1 The horizontal value chain (generic use of IoT data)

6.2.1.1.1 Vast ecosystem mainly from IT and software sectors

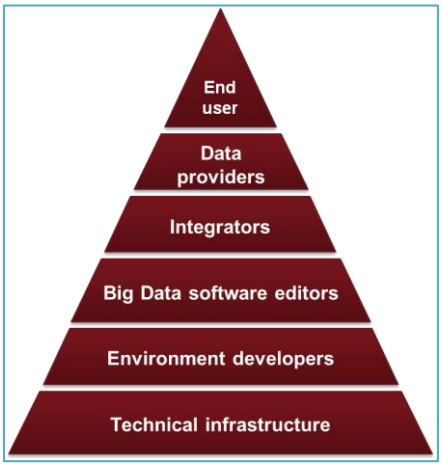


Figure 36: Big data value chain [3]

In relation to the conceptual framework for IoT data, different categories of actors are involved in implementing the technology:

- **Data providers**: These are companies and organisations that provide free and/ or paid data sources in the form of single or multiple streams. It can include social networks, public administrations or private businesses that provide access to some of their own data. These players do not necessarily send their data; they can simply make them available through APIs.
- **Integrators** offer businesses the opportunity to build an application which meets their own needs and then 'integrate' or install it on the server of the customer. These applications typically operate with multiple elements of a company IT system. For example, an application can automatically extract data from the customer database and subsequently analyse it with a big data 'integrated' application.
- **The big data software editors**, generally using a development environment, will offer different types of applications to analyse or 'draw' the data. In addition, some vendors are developing business intelligence software that enables the end client to make strategic decisions. For example, analysis of data may show that a particular demographic tends to gather at a given location at a given time, aiding marketing campaigns and decisions.
- **Environment developers**: In order to develop an application, it is often necessary to use a development environment. This facilitates developing the application and is intended to allow it to perform such specific tasks as parallel computing and management of very large databases. Hadoop is one example of a big data environment based on distributed computing.
- **Technical infrastructure providers** include all stakeholders providing infrastructure for big data technologies. It can typically be a telecom operator providing Internet access and a server manufacturer, on whose equipment big data applications will be installed. Data centre builders are also included in this category.

6.2.1.1.2 Major players are IT and software companies rather than pure big data players

Today, over one hundred players claim to be working in the big data field.

The pioneers of big data are actually players from scientific research and major Internet companies, including Google, Amazon and, more recently, Facebook who have used very large amounts of

data as a core part of their businesses since inception. Among the major names covering the complete big data value chain are Oracle, SAP, IBM and Microsoft.

They have either developed proprietary technologies for mastering such data or they are specialised players in data processing through data mining, business intelligence and database management.

Figure 37: Big data landscape [3]

Competition is already considerable on lower stacks of big data due to standardisation (Hadoop's open source platform has contributed to this) and the extension of traditional infrastructure providers and integrators. No single player has a clear leadership position.

Despite this considerable competition in the lower stacks, most of the value is captured by the data providers, who generally keep control of their most valuable data. Internet giants dominate this field of owning data, plus to a lesser extent, retailers, telcos and banks.

6.2.1.2 The vertical value chain (vertical specific value chain)

Following are examples of vertical value chains of the IoT markets in "connected healthcare" and "connected vehicles".

6.2.1.2.1 The connected healthcare value chain

Across the value chain of connected healthcare, the data life-cycle is a core issue, since data silo and security concerns have historically been severe challenges for the highly-regulated health sector. Connected devices and services hold significant potential to generate value by developing and placing "new generations" of sensors on the market for both caregivers and patients, and by providing new business opportunities around emerging services such as telemedicine, remote monitoring and home-care delivery respectively. The reduction of readmission and overall care costs, and the optimisation of the care path and efficiency are also of great interest, not just to the public health authorities, but to private health sector stakeholders too.

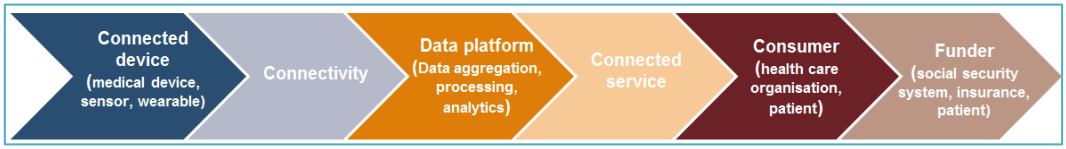


Figure 38: Healthcare value chain [4]

Where connected healthcare differs from other connected markets is that the funder, which may be a public health authority or a private health provider or insurer (i.e. not necessarily the end user/ patient), plays an essential role in propelling the market's progress. In some markets, such as home monitoring, the technology adopter may be the carer of the patient.

6.2.1.2.1.1 Medtech companies reign supreme

The position of the medtech companies in health industries is firm, and not at all likely to be shaken in the near future, as their initiatives are spanning out across full range of products and services. Medtech companies have inherent advantages over others in medical device/sensors development and in their knowledge of regulatory affairs, care delivery paths and care providers' working modalities. By comparison, Internet players are intervening in this market in a relatively subdued way. Their main purpose here is to seize opportunities out of their core business – Apple and Samsung are all seeking industry partners for health-related research by intermediating on the hardware-generated data. Google goes further by developing its own biotechnological products and services through its Verily and Calico divisions.

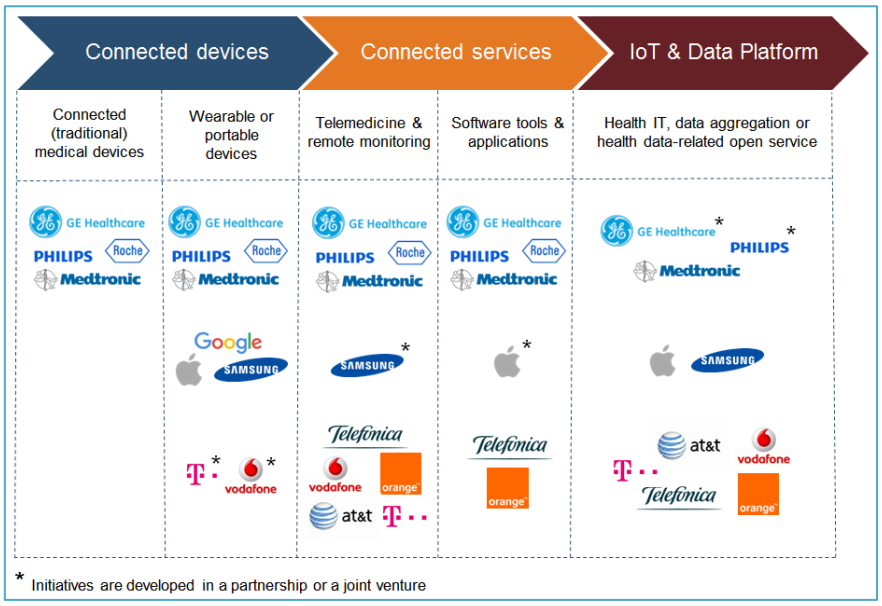


Figure 39: Initiatives by giant players in connected healthcare market

Telcos are active in this promising market and regard connected healthcare as an essential part of their IoT strategies. They have, though, adopted a wait-and-see approach. Their main initiatives refer to the "connected hospital" topic. The new and rising application is in remote patient-monitoring products, through a wholesale approach.

More generally, telcos are developing services for healthcare often through a specific division. Vodafone appears to be the most advanced player, while DT positioned itself early on and has interesting credentials. The majority of models are logically B2B2C solutions.

Security and fluidity of health data among different stakeholders is another key element to propel the connected healthcare market, driving more corporation between telcos, other ICT actors and medtech players to build open data platforms and services.

6.2.1.2.2 The connected car value chain

The figure below illustrates the whole automotive value chain including the main players.

Figure 40: Initiatives by giant players in connected healthcare market

6.2.1.2.2.1 Premium brand car manufacturers leading the way

It is no surprise that the premium brands are leading the way in the connected cars market. Their service portfolio is broader and more innovative by far. Moreover, their clientele is the most willing to pay for such services. The brands have implemented an embedded architecture to provide the best services in terms of QoS (Quality of Service).

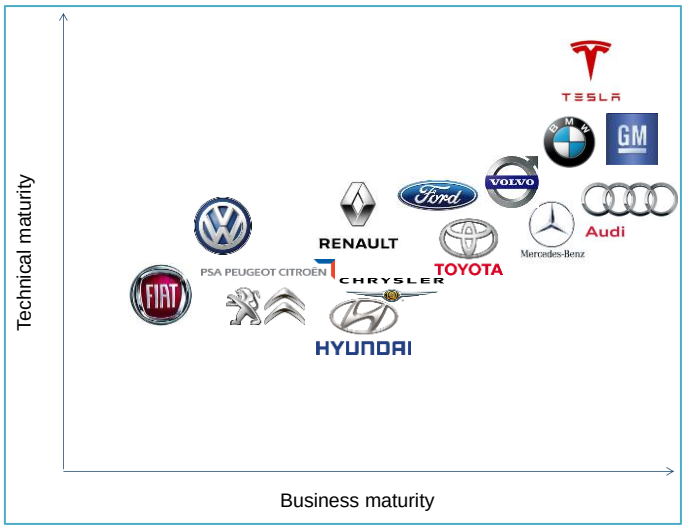


Figure 41: Main positioning of car manufacturers regarding connected-car market [3]

6.2.1.2.2.2 Main connectivity providers

All telcos are very involved in the automotive space. For them, the automotive and related connected-car topic is a top priority driver in their M2M and IoT strategy mix. Many regulations throughout the world will impose the presence of a dedicated SIM card in each vehicle.

In this, AT&T and Vodafone have a clear leadership. If wholesale rules, AT&T stands out with its B2C business model, through different business models including the integration of the car into a share plan.

- Both provide global SIMs which simplifies sourcing for manufacturers as they look at a global approach rather than a series of local SIMs from multiple telecom operators.
- AT&T and Vodafone also both provide vertical services (beyond connectivity services). The AT&T Drive platform and the recent Vodafone acquisition of Cobra has also generated significant interest from the automotive industry: both have multiplied their partnerships in the recent past.

It is also worth noting that China is a specific market and therefore China Mobile is seen as a key partner in this region.

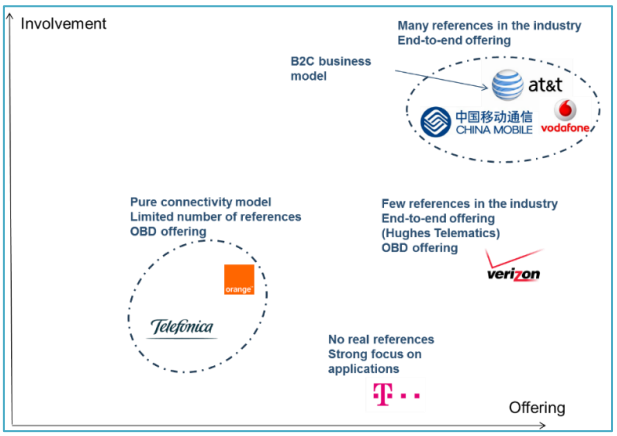


Figure 42: Key differentiation in carrier positioning in the automotive market

6.2.1.2.2.3 Main Internet player strategies

In the automotive market, Google seems to lead among the players under review here, as they provide not only a wealth of building blocks but even a car itself – in addition to being highly involved in the R&D automotive space. In terms of communication, Apple is also very involved with CarPlay and has been at the centre of rumours concerning the potential purchase of a car manufacturer, although at present its offering does not go further. Other OTT (over-the-top) players are strictly positioning on the platform and application layers.

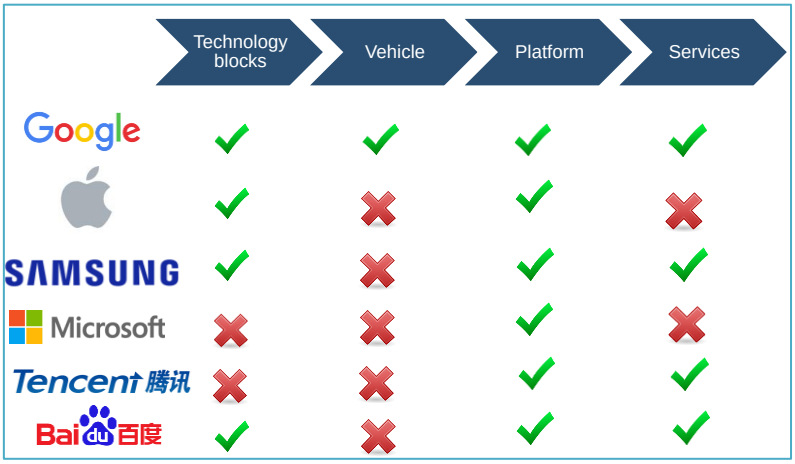


Figure 43: Key differentiation in positioning among Internet OTT players in the automotive market

6.2.2 Chain of data relation flows

On the free flow of data, it can be established that restrictions on the free movement of data within the European Union and unjustified restrictions on the location of data for storage or processing purposes are generally not addressed in generic IoT products and services.

This is understood as most restrictions are only applicable to certain industries, markets or use. It is however a main challenge as hyperconnected ecosystems are borderless and the data therein should be able to flow freely and unrestricted, at least within the European Union.

Quite a few member states have implemented sector-specific rules and regulations that differ per member state, thus hampering the DSM (Digital Single Market) and limit the competitiveness of European manufacturers, service providers and other vendors and their ability to benefit from marketing respective products, services and data across borders.

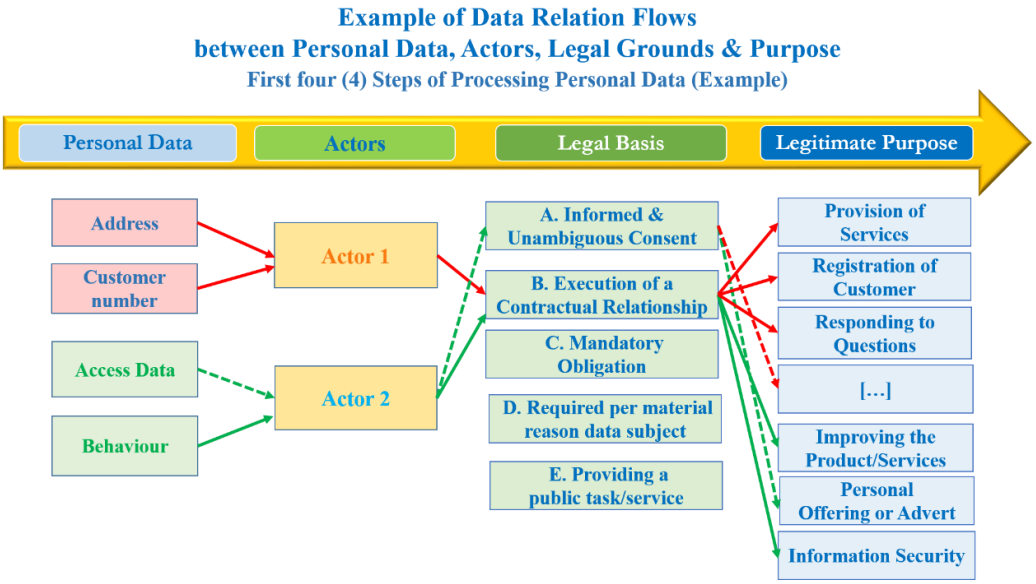


Figure 44: Example of Data Relation Flows between Personal Data, Actors, Legal Grounds & Purpose [5]

Note that before 28 May, 2019 only personal data was allowed to freely flow across EU Member States. The Regulation on the free flow of non-personal data, however, [6] lifted the previously existing barriers and has been discussed under the following chapter of this deliverable.

6.2.3 Chain of triggers

IoT devices are usually not operated in safe environments. It is therefore possible for hackers to access or obtain such devices and attempt to locate the key data on their microchips. In addition, due to the hyperconnected nature of the IoT environments, gaining access into an individual device, hackers may also access other connected devices and the network.

As the devices in an IoT ecosystem are connected to each other, communication between them often takes place rapidly, without human interference. This also means that any incident compromising security of a device may trigger a security compromise of other devices. The extensive degree of interdependencies creates high risk of a domino effect of different consequences potentially relevant for all value chains discussed. In other words, security risk in IoT ecosystems is diffused, as opposed to isolated.

The emerging risks and their multiplication are, of course, of great relevance, also, from a legal standpoint, and are therefore extensively addressed in separate deliverables in WP05. It has been argued, though, that "all other risks, liabilities and other elements, which could not be defined by legislation, standardization and agreement should in best case scenario be covered by insurance" in order to create sufficient guarantees for the IoT market [5]. In practice, insurance covers know risks and this approach is unlikely to afford protection to start-ups and SMEs developing new IoT products and services.

Incidents occurring in the cloud computing domain may well serve as an example of harm being diffused. In case of an intentional or unintentional incident, files containing data may be affected. This may cause downstream applications or systems relying on the affected files may cease to work properly. As a consequence, the consumer is likely to suffer harm. While this harm may in some cases be expressed in monetary terms, certain harms cannot always be adequately remedied by monetary damages, such as the consequences of data protection breaches [7].

An in-depth assessment of security risks inherent to a device which becomes a part of an IoT ecosystem is essential. In carrying it out, one must first carry out a security analysis, e.g. identify the threats, the assets to be protected and the level of desired security also has to be decided. The challenge is that the developer/actuator might not know in which system or environment the device will later be used. This can be countered effectively by introducing a classification (or certification)

system that would certify devices for use in particular use case scenarios depending on the level of risk. With this type of system, it would be possible to prevent devices entering the market without the appropriate security.

6.3 The emerging legal challenges of data flows

This section produces an overview of the relevant challenges from a legal point of view linking to data flows and, thus, relevant for the IoT Data Value Chain. These challenges either relate to this specific momentum of the European regulatory scene or are innate to the very nature of data flows. Note that a more elaborated discussion on the related legal challenges are produced under the deliverables due under "Task 05.03: Legal support, accountability and liability."

6.3.1 The changing regulatory landscape

It is apparent that the EU lawmakers have started taking initiative again and following years of preparation and laws lagging behind technology developments, their efforts are bearing fruit. 2018 witnessed the enactment of three very important pieces of legislation i.e. the revised Payment Services Directive, General Data Protection Regulation and the Directive on security of network and information systems. In 2019, the Regulation on free flow of non-personal data was adopted and substantial progress was also made on the ePrivacy Regulation front. It is important that organizations of all sizes remain vigilant and adjust their approach to the issues in question accordingly. Just as it is important for the EU lawmakers to monitor the implementation process, enforce the new rules and quickly release further updates where needed with due regard to the impact frequent changes can have on SMEs.

Although technology and innovation tend to be considerably ahead of legal regulation, European lawmakers have taken steps to catch up on regulating the recent technology developments. As a result of their efforts, numerous legislative acts entered into force in the first five months of 2018 and 2019, and further updates to the related laws are expected along the way[11]. The discussion below points at the most relevant legislations for the IoT domain.

Figure 45: From 2018, Digital & Data become Highly Regulated Domains[12]

[11] See, also, CREATE-IoT Project, D05.06: Legal IoT Framework.

[12] Ibid.

6.3.1.1 General Data Protection Regulation (GDPR) [14]

The GDPR which came into force on 25 May 2018 received much attention from organizations as well as governments around the world. This is not surprising, since it places obligations upon everyone handling personal data of EU residents, irrespective of where the data is collected, stored or processed. Given its rather consumer-focused character, the Regulation considers privacy and processing of personal data of natural persons a fundamental right. While emphasizing key principles related to the processing of personal data, including lawfulness, fairness, transparency, data minimization and accountability, the Regulation grants numerous rights to users (data subjects). These include the right of access by the data subject and the right to be forgotten.

GDPR takes a more stringent stance towards data protection and security requirements by requiring organizations to assess the level of protection from a wider perspective. The GDPR offers an equation for finding the appropriate level of protection, per purpose, per impact assessment, et cetera. The level of having state of the art security measures (both technical and organizational) in place is the benchmark in the GDPR, where (i) the related cost of implementation, (ii) the purposes of personal data processing and (iii) the impact on the rights and freedoms of the data subject (also good, bad and worst-case scenarios) need to be taken into account, whether one is either data controller or data processor. We call this the appropriate dynamic accountability (ADA) "formula":

State of the art security – Costs – Purposes + Impact

The GDPR aims to ensure up-to-date levels of protection by requiring the levels of data protection and security to continuously meet the ADA formula. Failure to comply with the Regulation may result in some hefty penalties which may to amount to several billion Euros for some large enterprises, since the Regulation allows for penalties of up 4% of the total worldwide annual turnover.

6.3.1.2 Directive of Security of Network and Information System (NIS Directive) [15]

While the GDPR focuses on privacy, NIS Directive aims to achieve a high common level of security of network and information systems within the EU by improving cybersecurity capabilities at national level, increasing EU-level cooperation, and setting out risk management and incident reporting obligations for operators of essential services (banking, energy, transport, financial market infrastructure, health, drinking water and digital infrastructure) and digital service providers (online marketplaces, online search engines and cloud services). Operators of essential services and digital service providers are tasked with ensuring the security of their networks and systems to promote a culture of risk management and ensure that serious incidents are reported to national competent authorities and computer security incident response teams (CSIRTs).

In case of a “significant” impact on the provision of the operator’s service, the operators of essential services will have to notify the national competent authorities. Digital service providers will have to notify any incident having a “substantial” impact on the provision of the service. Notification processes have also been put in place to ensure effective communication of incidents across members states’ CSIRTs. In order to support and facilitate strategic cooperation and the exchange of information between member states, the Cooperation Group has been established, consisting of representatives of Members States, European Commission and the European Union Agency for Network and Information Security.

The complex institutional ecosystem set out by NIS Directive entered into force in August 2016. Therefore, Member States were given time till 9th May 2018 to transpose the Directive into national laws. In addition, they were offered additional six months (until 9 November 2018) to identify the operators of essential services.

6.3.1.3 Payment Services Directive 2 (PSD2)

PSD2 entered into application on 13 January 2018 addresses better integration of internal market for electronic payments. It puts in place comprehensive rules for payment services, with the goal of making payments between Member States as easy, efficient and secure as payments within a single country and equating costs. By setting out strict security requirements, transparency and the rights and obligations of users and providers of payment service, PSD2 seeks to open up payment markets to new entrants leading to more competition, greater choice and better prices for consumers.

In addition, the Directive aims to open up the EU payment market to companies offering consumer- or business-oriented payment services, in particular account information services allowing users to have an overview of their financial situation, and payment initiation services allowing consumers to pay via simple credit transfer for their purchases.

From the consumer point of view, it is also important to note that the Directive significantly reduces consumers' liability for non-authorised payments, introduces an unconditional refund right for direct debits in euro and puts in place an obligation to remove surcharges for the use of a consumer credit or debit card. A user-friendly leaflet on all consumers' rights was published by the European Commission in September 2019.[13]

6.3.1.4 Proposal for Regulation on Privacy and Electronic Communications (ePrivacy Regulation)

Finally, it is worth noting that both the European Commission and the European Parliament took the initiative of updating rules relating to privacy and electronic communications and reinforcing trust and security in the Digital Single Market. Having identified areas to be addressed (including stronger protection online, simpler rules on cookies, and transparency on direct marketing, to name a few), the Commission released a Proposal for the Regulation in January 2017 ("Proposal").

In June 2017, the Parliament's Committee for Civil Liberties, Justice and Home Affairs (LIBE) published a draft report on the Commission's Proposal. Overall, the resulting text of the Regulation strengthens privacy protection for individuals. Amongst others, it provides clarity regarding what legitimate grounds for processing prevail if both the GDPR and the ePrivacy Regulation apply to a processing operation and prohibits all further use of electronic communications data collected under ePrivacy rules. In addition, significantly stronger obligations for privacy by default are proposed, including end-to-end encryption (with no backdoors) proposed as a security default measure for ensuring confidentiality of communications, and a national Do Not Call register for opting out of unsolicited voice-to-voice marketing calls. Finally, the amendments provide for an extension of the principle of confidentiality of communications to machine-to-machine communications as well as enhanced definitions of "electronic communications metadata" and "direct marketing".

With respect to IoT devices, the Proposal contained a recital regarding IoT devices stating that machine to machine communications involved conveyance of signals over a network, thereby constituting electronic communications. Therefore, in order to promote secure IoT devices, it was stated that the Regulation would be made applicable to machine to machine communications. This provision was considered problematic as it would require IoT device manufacturers to obtain the end users consent so as to perform its most essential function of transmitting data from a connected device to another device. Unlike the GDPR that permitted processing of personal data for performance of a contract, the Proposal does not recognise "performance of contract" as a lawful basis for processing. An amendment to the proposal which followed later came as a relief to IoT

[13] European Commission, Your rights when making payments in Europe, also available at: https://ec.europa.eu/info/sites/info/files/business_economy_euro/banking_and_finance/documents/leaflet-your-rights-payments-eu_en.pdf

device manufacturers as Recital 12 was modified to make a distinction between the application layer and the underlying transmission layer of machine to machine services. It was clarified that since transmission services require conveyance of signals through an electronic communications network, it would qualify an electronic communication service and therefore, would be subject to the provisions of the ePrivacy Regulation.

Since the Proposal was first published, several discussions have taken place and resulting amendments have been shared by the Council of the European Union including the recent amendments proposed by the Finnish Presidency on 4 October 2019 and later on 15 November 2019. However, as the current deliverable was being drafted, the Permanent Representatives Committee of the Council of the European Union (COREPER) rejected the Council's position on the draft of the Regulation. Difference of opinion between the member states on topics such as cookie walls, unsolicited commercial advertising and the like is one of the major reasons why the draft failed to garner support. In 2020, as Croatia will take up EU presidency, it will have to decide whether to withdraw the proposal entirely or to re-draft the proposal in a new attempt to gather support from the member states.

6.3.2 The removal of localization restrictions within EU

The discussion on IoT data value chain, and consequently data flows, is closely inter-linked to the discussions at EU level regarding the free flow of data. The free flow of data is highly relevant for the IoT domain, as data is what actually keeps the IoT moving and alive. The free flow of data concerns data in any form, nature or structure, that can be created, uploaded, inserted in, collected or derived from or within the IoT. It includes, amongst others, proprietary and non-proprietary data, confidential and non-confidential data, non-personal and personal data, as well as all other human readable or machine-readable data.

Prior to 28 May 2019, the legal regime governing data flows was highly rigid and fragmented. More specifically, and as has been mentioned earlier, a significant number of Member States had implemented their own sector-specific rules and regulations. These rules imposed, in essence, data localization restrictions inhibiting the free flow of data within EU and will in effect stop the Digital Single Market from becoming a reality.

In particular, there were a plethora of data location restrictions within the individual Member States, as well amplified sets of diversified approaches at national level, which were often largely unreasonable or highly disproportionate. The said data localization restriction resulted from the absence of well-defined standards and practices at the level of the European Union, while the absence of well formulated standards fosters further the implementation of data localization restrictions, thus, catching market players in a vicious circle.

All these instruments of European law due to their nature as Directives are transposed in the national orders by virtue of legislative instruments of all kind, thus, adding complexity and discouraging companies expanding their business in other Member States. Those restrictions mostly relate to the handling of financial data, tax data, health data, and book keeping data, gambling data, banking, as well as public procurement at national & local level. For instance, in the Netherlands public records -both paper and electronic- have to be stored in archives in specific locations in the country. Furthermore, there is often a lack of common understanding and culture in key matters across sectors and Member States. Data localization restrictions bring about the absence of a harmonized understanding as companies processing data across different Member States face increased administrative burdens and need to comply with different legal systems.

On November 14, 2018, the Council of the European Union and the European Parliament backed the Regulation for free flow of non-personal data ("Regulation")[14] thereby prohibiting data

[14] Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union OJ L 303/59

localization restrictions[15]. The intention behind enacting the Regulation was to boost the EU data economy by establishing a single market for data storage and processing services. From a technology perspective, the new framework will augment the development of the data economy and new technologies like Internet of Things which function by gathering and processing data, therefore requiring free flow of data so as to achieve data-driven growth and innovation. The Regulation was published in the Official Journal of the European Union on November 28, 2018 and became applicable to all Member States from 28th May 2019.

On May 5, 2019, the European Commission published a Guidance to throw light on the interactions between the Regulation and the GDPR, as both regulations approach the free movement of data within the EU from two distinct angles. While the Regulation prohibits localisation of non-personal data, the GDPR aims at maintaining a high level of protection of personal data. However, there can be situations involving mixed data sets i.e. Data sets which consist of both personal as well as non-personal data, where separating the two could either be extremely tedious or economically unfeasible. As per the guidance, normally, such a mixed data set would be subject to the provisions of the GDPR and the obligations of data controllers and processors provided thereunder.

6.3.3 The outdated definitions: the consumer protection paradigm

The Directive 85/374 [16] constitutes the European legal framework proving for the protection of individuals in case of liabilities incurred by the use of defective products. Notwithstanding its noble goal of consumer protection, the previously mentioned directive was clearly written with consumers who only make use of standalone material products. There is a real need for it to be adapted into the current consumer environment in which software forms part of smartphones, as computers and as a (key) component of (domestic) devices.

In this respect, there is a series of challenges concerning the liabilities incurred by the use of the devices and services in the IoT domain. In particular, first, IoT devices and services can be subject to unauthorised access by third parties (individuals or machines) which could tamper with such device or service; second, sensitive (personal) data could be stored, processed or exchanged in an unauthorized manner and thereby breach the rights of a data subject (person); third, malicious data can be created/deducted by one IoT device and be exchanged with other IoT devices (and may there cause further harm); finally, this malicious data could lead to a malicious decision which could cause damage to other (IoT) devices or a human being.

Consequently, these triggers a series of questions regarding:

- What or who actually caused the damage and
- Who is liable for such damages?

Providing answers to these questions is complicated as the related applicable definitions are essentially not fit for the IoT environment. In particular, under the product liability directive, all movables are considered as products (even if incorporated into another movable or immovable) including electricity. The damages link to death or personal injury, while the injured person is requested to prove the damage, the defect and causal relationship between defect and damage.

An evaluation by the European Commission in September 2016 revealed that since the adoption of the Directive in 1985, the Directive had not been subject to any formal evaluation. The evaluation considered different aspects of the Directive including effectiveness, efficiency, coherence and relevance. Pursuant to the formal evaluation, the Commission launched a public consultation on the Directive that ended on 26 April 2017 with inputs from a range of stakeholders. One of the fundamental questions of the consultation was whether the Directive was fit-for-

[15] See, also, CREATE-IoT Project, D05.06: Legal IoT Framework.

purpose vis-à-vis new technological developments like Internet of Things. On 7th May 2018, the European Commission published a report which acknowledged the exponential technological developments and the change in concepts such as 'product', 'defect', 'damage' etc. Subsequently, the Commission set up an expert group to assess the Directive and is expected to produce a report relating to the liability framework for artificial intelligence and internet of things. In addition, the Commission stated that if deemed necessary, it will update the definition of concepts like 'product', 'defect', 'damage' and 'producer'.

In addition, establishing liability while dealing with IoT devices may be a herculean task for a user. This is because the current definition of the term "defect" focusses on the safety which a person is entitled to expect from a product. The first reason why this is problematic, is the fact that consumers generally have relatively low expectations when it comes to the quality and safety of software and IoT devices. Secondly, the capability of IoT devices to act autonomously makes it very hard to describe or foresee what kind of safety level they have, let alone what kind of safety level people are entitled to expect. Moreover, while dealing with IoT devices, it may be difficult to hold a specific party liable for the resulting damage. IoT devices, on their own account, can be defined as highly complex value chains, linking different hardware and software components together, communicating with one or more networks and other devices. As a result, a consumer does not always have a good and complete understanding of what a device does, how it works and more importantly, what a device do and how it works. This lack of insights might result in the situation where damage is caused, but the consumer does not even know how and by what.

6.3.4 The contractual complexities

This section gives an overview of the emerging contractual complexities in the IoT environment. Note that an extensive analysis of those are provided under the deliverables falling within the scope of "Task 05.03: Legal support, accountability and liability."

IoT devices encompass and consist of numerous layers including hardware, software, data and service. This multi-layered structure often requires numerous different manufacturers and providers to participate in the production of the device as well as in the provision of services during its life time. This setting accounts for a large number of contractual documents, licences, notices, declarations and/or reports to be in place and effective, not only between the supply-side actors themselves, but also vis-à-vis the customer.

The resulting relationships tend to be very complex and bear a great deal of challenges in achieving transparency in allocating responsibilities and risks, as well as issues concerning jurisdiction and remedies. This section aims to address some of the most relevant challenges while noting that it only serves as an indicative overview as it will be treated more extensively in other deliverables related to this topic.

One of the main challenges that customers are repeatedly faced with is the difficulty to understand applicable contracts, agreements and other legal documents. Numerous reasons account for this issue, but for purposes of further discussion it is mainly worth noting that, aside from the European versions of contracts often being verbatim reproductions of their US counterparts, (which may not be necessarily suitable), identifying all the applicable documents may be a challenge in itself. For example, in the case of Nest connected thermostat produced by Nest Labs owned by Google, this challenge is illustrated by about 13 legal documents which a user has to read in order to get a "clear" picture of the rights, obligations and responsibilities in the supply chain.

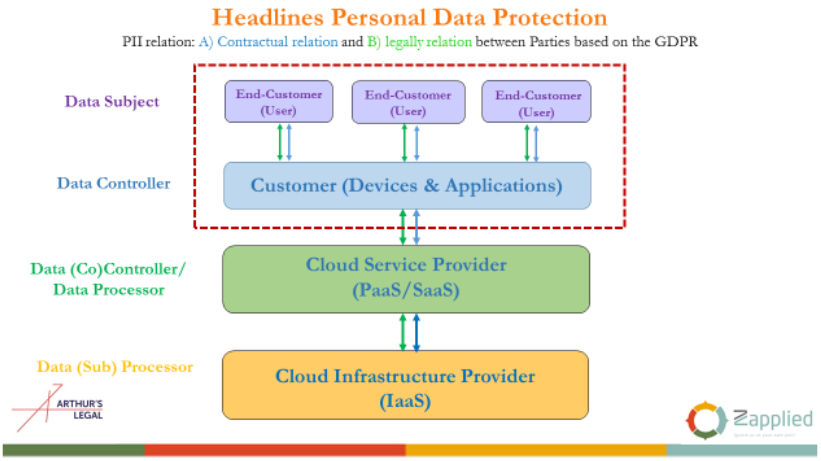


Figure 46: Headlines Personal Data Protection

Having a clear picture of legal relationships between the consumer and individual actors in the supply chain is also challenging from the perspective of the scope of the documents. While they may claim that they are only applicable to one separate part of the IoT device, due to the connected nature of IoT ecosystems it is difficult to imagine a part of the system or a separate layer functioning irrespective of the remaining parts or other layers, i.e. without affecting the whole ecosystem. However, in order to provide the customer with a sufficient amount of transparency, it is essential that the customer has a true and honest account of how the layers (and the respective contractual documents) interact and what supplier becomes relevant (not only active) in what layer. Just as the customer should be able to identify the parties upon whom the service is dependent and who are the processors and sub-processors of data. Not only does this information provide the customer with greater transparency; it also helps them establish the extent of liability of various suppliers should a problem arise that requires legal redress.

Further questions concerning liability and other complex contractual issues arise in context of IoT devices that have the ability to make autonomous decisions and enter into legally binding agreements with third parties (e.g. connected home appliances purchasing products from third parties). On the one hand, questions of liability for actions of these autonomous devices are inevitable. On the other hand, although our traditional understanding of property is a static one, it is likely that that it will need to change and respond to the dynamic nature of IoT devices which are able to evolve and mature over time.

From a separate perspective, it is also important to consider the status and the role of the customer in the ecosystem. It has been argued that two further distinctions of legal consequence can be made [8]. “First, the end-user may be the contracting customer or a third party, such as a family member. Second, the device itself may be owned by the customer or may be leased to the customer by the supplier (or provided as part of rented or leased premises).” Considering the latter, “the distinction between the device and the associated services becomes critical, because the Nest ToS states that if the device owner does not agree with the terms ‘you should disconnect your products from your account and cease accessing or using the services” [9]. However, in some jurisdictions, a disconnected IoT device would potentially breach the law. For example, according to the Sale of Goods Act 1979 of England and Wales, the purchasers of goods will "enjoy quiet possession", which term would be potentially breached if when the Nest device were disconnected it loses most of its functionality.

Last but not least, complexities also arise in the context of clauses relating to selection of jurisdiction in contracts relating to IoT devices. Most commercial contracts explicitly stipulate applicable law and jurisdiction governing them, to the maximum extent permitted by law. However, in cases where mandatory national laws apply, judges will have to abide by those. As a consequence, cases may arise in which the judge will have to apply different pieces of legislation to the same product. Already in today’s connected world it is not difficult to imagine a scenario in

which a Dutch customer uses a US-manufactured IoT device during their holiday in Tunisia, where the device was purchased in Venezuela, consists of software running in Ireland and uses applications developed by a Chinese company. This presents a very complex setting where different pieces of legislation are likely to apply in respect of a single device.

6.4 Economic feasibility

Companies can revolutionise their business models by cross referencing disparate data from a host of sources (including external sources and IoT generated data) – a process which extends well beyond classic approaches to culling and mulling statistics. The first major applications today are in the area of sales and marketing, by analysing attrition rates and knowledge of consumer behaviour. But the range of possibilities is far vaster: from creating new services and fostering new business models, to optimising processes.

6.4.1 Major opportunities for exploiting IoT data for verticals

The data used in these solutions can come from various sources:

- Internal (such as internal databases).
- From data aggregation (forms, documents, sensors, web browsing), and particularly IoT platforms for this study.
- From open data sources and APIs from third parties exchanging data (automatic data transfer). Open data and data provided through APIs often have limited value without further transformation from third parties, such as developers, start-ups or other SMEs developing new products or services.

Valuable data is generally blocked/controlled by the owners, as there is a competitive advantage attached to the data itself. This is especially the case with major OTT players like Facebook but also with major IoT players (data is generally only shared with explicit user consent on a case by case basis).

Beyond Internet giants, many vertical players have a wealth of information about millions of users. There are several opportunities for vertical stakeholders to use and share data:

- **Internal use**: Data are analysed internally and are not shared with third parties.
- **Intermediation for third parties**: Data are analysed internally, but the result of the analysis is shared with third parties, whether monetised or not – a way to monetise data without disclosing them.
- **Data sales to third parties**: Data are directly shared and monetised with third parties. Data will generally be anonymised and sold in aggregated versions.

6.4.2 Main barriers

Most businesses are still struggling to pin down new business models, due to a lack of high-value unstructured data and go beyond the proof of concept stage.

Added to which there is a plethora of disjointed products available, as IT and IoT companies alike are all trying to ride the wave of media hype, even if it means rebranding their old data processing solutions as big data and/or IoT data products. This plethora of activity only adds to the confusion, although a few major players are starting to emerge, as are concrete examples of a solid return on investment.

The main obstacle to the implementation of data monetisation models is more cultural than technological, fuelled by conflicts between departments and traditional approaches to sharing resources and knowledge, both within and outside organisations. Without a real change in philosophy over how to treat data, the harvest of big data will remain a poor one.

In addition, the fact of centralising data only increases data protection and privacy issues, which means business need to rethink how to manage the plethora of data, but also how to collect data from users who are increasingly reluctant to share them.

Such major barriers can be overcome over time, but this will require organisational and cultural change rather than greater technology literacy. Regulatory and privacy issues will remain important.

The growth of big data market would accelerate significantly if turnkey solutions easily integrating big data technology (likely cloud-based) were provided to extend the adoption from large accounts to SMEs.

6.4.3 The potential of new revenue streams through data monetisation

The two main pillars to generate new revenue through IoT data is by either

- Developing new products and especially new associated services (servitisation – i.e. the capacity to generate recurring revenues not from selling a machine and object but from services); or
- By selling the data to third parties (even where the data remains under the control of the data owner).

With this in mind, there are several potential approaches where vertical players can position to generate more revenues through data:

- More sales of core services. Improvement even by a few percent of conversion rates could bring in significant revenues, especially on expensive products and services.
- Sales of additional paid services on top of existing products and services (i.e. servitisation), sold directly, or commissions of third-party services. Typically, only a small fraction of existing product users will subscribe to these services (between 2 and 20%), but services can reach 5 to 10 EUR per month. The revenues from added-value services may be used over time to reduce the cost of the core product or service and attract more customers (in case of price elasticity).
- Sales of individual data, i.e. one-to-one advertising or marketing, or aggregated data (user data, performance data), namely insights. Players can leverage their entire user base, as they generally rely on anonymised data.
    - Nonetheless, revenues from advertising and insights need to be put in perspective. Revenues per user are generally quite low, even when involving very large amounts of data captured through many services. For instance, Google generates around 3 USD per month per user and Facebook close to 0.5 USD. Newcomers with limited data sets are likely to generate even less, even with brand new data. Insights are just giving patterns of consumption and are therefore even less valuable.
    - Sales of aggregated data allow a player to bypass the most controversial privacy issues, but offer fewer perspectives of revenues, as only patterns can be identified. The approach is therefore not as efficient as targeted advertising and does not allow for real-time interaction.
- Sales of tools leveraging certain types of data such as billing or APIs. Pricing is often independent of the value of the data, but more cost oriented.
- Intermediaries leveraging data and user control to improve reselling third-party products (such as recommendations, or a platform and kiosk approach).

Verticals may position themselves across all types of approaches but will have to arbitrate between the various business models, which may require investment, the expected value to be generated by each approach and the competition between models. Indeed, internal optimisation of core products and services can often have greater impact, due to the lack of competition compared with innovative products.