PAPUD   ITEA3 - 16037

Profiling and **A**nalysis **P**latform **U**sing **D**eep Learning

D1.1   Market Analysis

HISTORY OF CHANGES

Version	Date	Changes	Author
V 0.1	08/03/2018	First draft	HI Iberia
V 0.2	22/03/2018	Version with Papud Template	Cornel Crisan
V0.3	23/03/2018	Contributions from IMT to BAREM Use Case	Laurent Goncalves
V0.4	30/03/2018	Contributions from BEIA to Call Center Use Case	George Suciu
V0.5	02/04/2018	Contributions from LORIA to BAREM Use Case	Samuel Cruz-Lara
V0.6	04/04/2018	Contributions from HIB to HHRR Use Case	Elena Muelas
V0.7	03/04/2018	Contributions from ERICSSON to Call Center Use Case	Arda Gureller
V0.8	03/04/2018	Contributions from LILLE to BAREM Use Case	Marius Bilasco
V0.9	05/04/2018	Contributions from BULL to HPC Use Case	Cornel Crisan
V0.10	06/04/2018	Integration and review of all the contributions by HIB	Inmaculada Luengo Elena Muelas
V 0.11	09/04/2018	Contribution from 4C	Veerle Liebaut

V0.12	16/04/2018	Re-organization of Use Cases starting from meting results	Inmaculada Luengo Elena Muelas
V 0.13	17/04/2018	Business model on UC5	Cornel Crisan
V 0.14	24/04/2018	Aggregated contributions for UC2: call centers	Benoit Germonpré (4C)
V0.15	26/04/2018	Review of contributions	Elena Muelas
V0.16	26/04/2018	Add contribution to Market Analysis for UC4	Laurent Goncalves
V0.17	27/04/2018	Add IMT contribution	Samin Mohammadi
V0.18	03/05/2018	Review and comments to contributions	Elena Muelas Inmaculada Luengo
V.019	11/05/2018	Add KocSistem contribution	Ahmet Cagatay Talay /Gunes Soyler /Ozgur Akarsu
V0.20	17/05/2018	Minor modifications in HPC UC	Cornel Crisan
V0.21	17/05/2018	Correction of Market Value Chain format	Sebastien Manfredini
V0.22	28/05/2018	Contribution to UC1 Market Analysis	Cenk Ustabas
V0.23	31/05/2018	Contribution to Introduction and Conclusions. Document review	Elena Muelas Inmaculada Luengo
V0.24	04/06/2018	Update Market Value Chain UC3	Lucian Necula
V0.25	07/06/2018	Document second review. Draft version ready for internal review.	Elena Muelas Inmaculada Luengo
V0.26	11/06/2018	Reviewed by IMT	Samin Mohammadi
V0.27	22/06/2018	Solved minor referencing problems	Lucian Necula
V0.28	22/06/2018	Reviewed by BULL	Cornel Crisan
V0.29	26/06/2018	Integration of contribution from SETUR, BEIA, PRESSINOV after review	Elena Muelas Inmaculada Luengo
V1.0	26/06/2018	Final version ready to be submitted	Elena Muelas Inmaculada Luengo

ABSTRACT

This report is the result of the work performed in *Task 1.1 Business Model and Market Analysis* which consists of a research and analysis of the current market situation and the envisaged business model at the initial stage of the project for the different use cases that will be developed within the project execution.

Table of Contents

List of Figures

List of Tables

1 Introduction

This report is the result of the work performed in *Task 1.1 Business Model and Market Analysis* which consists of a research and analysis of the current market situation and the envisaged business model at the initial stage of the project for the different use cases that will be developed within the project execution.

In PAPUD project, five use cases have been defined in order to demonstrate the project aims within different domains. The use cases envisaged are the following:

- Use Case 1: e-commerce
- Use Case 2: Call centre
- Use Case 3: Recommendation system for Human Resources
- Use Case 4: Behaviour Analysis for Reverse Efficient Modelling (BAREM)
- Use Case 5: HPC

Details for the use cases are included in *D1.2 Use cases Definition* as result of the work performed in Task 1.2. In current deliverable, the analysis performed for each of the use cases of the PAPUD project covers the following details:

- Market Value Chain: At this section the market value chain for the different use cases is defined including the link in the chain of each of the consortium partners involved as well as their participation.

- Market Analysis: This section presents a report on the current tools, technologies and procedures currently available and used for facing the different issues currently.

- Business models: This section shows an initial business plan for the use case for the partners participating in each of them. For sure, this business model is a first idea that will evolve and change with the progress project.

According to changes in the market during the project lifetime this report might be updated as needed.

2   Context

2.1   Project

Businesses currently have to deal with a data set that is more than they can handle. Today's necessity is not the usage of data analytics; it is the utilization of combined technologies in which data analytics are executed to make sense out of the data. The scope of the project is to build a universal model for data analytics using Deep Learning on a proposed set of technologies including HPDA (High Performance Data Analytics) environment that fit best to the data provided

2.2   Work package

Work package 1 deals with the embodiment of the use cases by defining the roadmap and the required technologies. For each use case, a review is performed to identify all constraints related to users or a given domain. This work package will define all technical and functional requirements needed for the specification and the development step. Business-level strategies are also taken into consideration; the innovation will be determining factor for use cases study.

In particular, this document is the output for T1.1. Business Model and Market Analysis. This task will consist of business/market research and analysis that includes evaluation of best practices in the market. At the end of the task, a detailed report of the current market situation will be provided as well as a business model for the initial period of the project. According to changes in the market during the project lifetime this model might be updated as needed.

3 Market analysis and business model for UC1 – e-commerce

3.1 Market Value Chain for UC1

The general market value chain for e-commerce is presented in Figure 1.
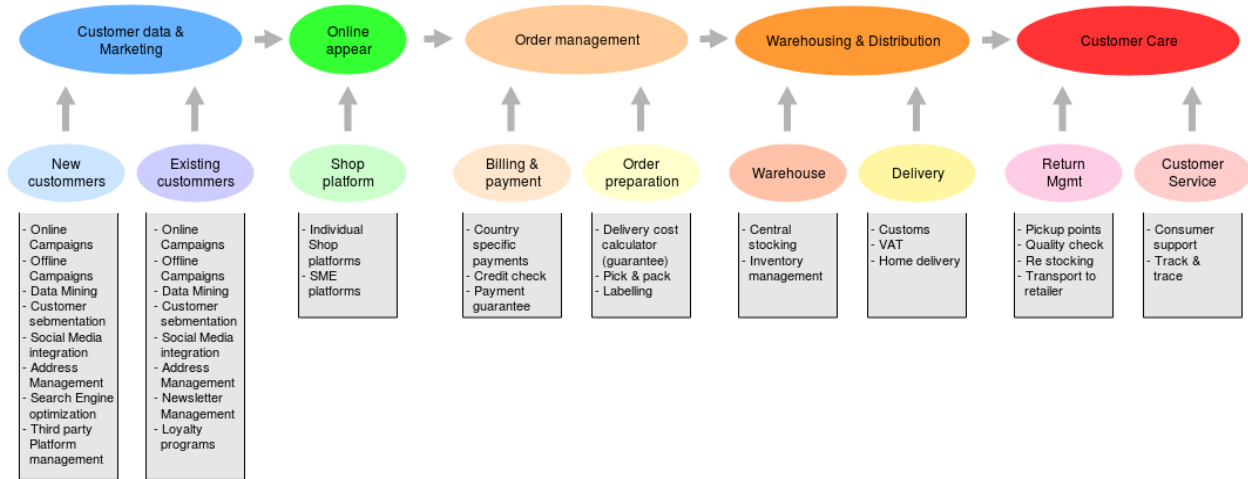
Figure 1 - Market value chain for UC1

The positioning of the consortium is mainly on the three following segments:

- Customer data & Marketing
- Online appear
- Customer care

Here are the different positioning of involved partners:

- Pertimm value proposition is on two main segments in the Online appear segment :
    - On the shop platform, pushing products to maximize the sales, using search engine on production and recommendation using product lists and user data preferences
    - On the Customer care segment, pushing products based on user data preferences
- Setur will work on data provisioning. The context of the data is: products and categories, locations and customer identity information (masked). Information such as age range, gender, marital status is also available. The identity of the customers will be hidden with a masking algorithm. Masked data will be examined, and a shopping profile will be created on individual basis.
- Press'Innov will work on content extraction from structured data using Deep Learning technologies.
- IMT will work on content and user popularity data in order to predict their future popularity using deep learning techniques and will act as a technology provider in the value chain.
- KoçSistem will position in Customer data & Marketing segment. It will work on building a deep learning based recommendation engine. The recommendation engine will serve for marketing better fit products to customers. In addition social media will be surveyed to understand customer needs to improve services.

3.2 Market analysis

3.2.1   e-Commerce Market

According to [1], the global e-commerce market will grow steadily and post an impressive CAGR of more than 19% by 2020. One of the major factors contributing to the growth of the e-commerce market is the diverse portfolio in several product segments such as electronic gadgets, apparels, books, and cosmetics. These categories in the e-commerce websites are regularly updated with latest offerings, which not only attract new customers but also retains customer loyalty. Also, these shopping websites allow customers to make multiple payment options through credit and debit cards, electronic wallets, net banking, cash on delivery (COD), and cryptocurrency. To withstand the growing competition in the market and maintain customer allegiance, several e-commerce companies are introducing incentives like free shipping to customers. They are also offering hassle-free exchange and return policies to increase their brand proposition. Social media advertising is one of the latest developments in the e-commerce market, which not only allows advertisers to use display banner ads but also sponsored content to attract customers.

Vendors in the market are switching from website platforms to app-only services owing to the increase in online shopping transactions through mobile apps. Shifting to app-only can be a risky situation for vendors as many developing countries, such as Africa, India, and South Korea face grave issues of internet connectivity, which restricts the customers from making an online purchase. The e-commerce vendors are coming up with a lighter version of the apps that work even with slow internet connectivity. They are focusing to develop lighter versions of their mobile sites that support browsing even at moderate internet connectivity like Google, Facebook, and LinkedIn.

The global e-commerce market is highly competitive because of the presence of many large established players. Intense competition prevails in this market where the vendors contend in terms of product portfolio, pricing, delivery and payment options, return policies, and discounts and offers. To increase their profit margins and extend their geographic presence, vendors are making investments in planning, designing, and developing their services and acquiring new players. The strong positions of the established vendors and long break-even periods will restrict the entry of new players in the market over the predicted period.

In market research analysts predict that the global duty-free retailing market will witness impressive growth and will post a CAGR of more than 8% over the forecast period. Duty-free retailing is mostly concentrated on the sales of luxury brands of fashion goods, accessories, hard luxury items, cosmetics, fragrances, wines, and spirits. However, luxury items, fashion goods, and accessories account for the majority of revenue generated in the global duty-free retailing market. Moreover, the exemption of excise or custom duties at airports, cruise lines, and international borders further lower the prices of these otherwise expensive premium products. With premium products becoming affordable to the price-sensitive consumer segment, the market for duty-free retailing is expected to have a positive outlook in the coming years.

---

1   https://www.technavio.com/report/global-media-and-entertainment-services-global-e-commerce-market-2016-2020

In the duty-free market, the conventional e-commerce concept is structured a bit differently, because regulations do not allow duty-free flow through e-commerce. We think e-commerce is a productive way to increase the awareness of luxury brands. Because flight and waiting times for passengers are very important, e-commerce (pre-order system) offers them the opportunity to easily pre-examine and shop for luxury products.

Leading vendors in the market are: Alibaba**,** Amazon.com**,** Apple**,** eBay, Google and PayPal.

Other prominent vendors in the e-commerce market include ASOS.com, Barnes & Noble, Best Buy, Costco, GameStop, Groupon, J. C. Penney, Kohl's, Liberty Interactive, LL Bean, Lowe's, Macy's, Newegg, MasterCard, Sears Holdings, Softcard, Staples, Target, The Home Depot, Visa, and Walmart.

Segmentation by application and analysis of the e-commerce market

The home appliances segment is the largest revenue segment in the e-commerce market and accounted for more than 32% of the total market share in 2015. The segment includes consumer electronics, houseware, and other kitchen appliances. Much of the segment's growth can be attributed to the increasing credibility of the brands online and the frequent tendency of consumers to upgrade their homes. Moreover, the growing trend towards recycling and the use of smart energy in consumer electronics will be a major factor augmenting the growth prospects for this segment in the coming years.

Geographical segmentation and analysis of the e-commerce market

Among all geographies, APAC (Asia and Pacific) will be the fastest growing region in the e-commerce market during the forecast period due to the increasing internet penetration and high purchasing power of the middle class population. The local vendors dominate the region's market and restrict the influence of foreign players, thereby maintaining the revenue growth. With the growing adoption of smart devices and latest digital technologies, the region is anticipated to grow at a tremendous CAGR of more than 25% over the next four years.

In terms of geography, the APAC region dominated the global duty-free retailing market and is expected to lead the market during the predicted period. South Korea, China, Hong Kong, Singapore, and Thailand are some of the key countries that drive the sales and growth of the duty-free market in the region. The growth of duty-free retailing in the region is attributed to the increasing number of Chinese outbound tourists, who contributed nearly 30% of the global duty-free retailing market's revenue in 2016. Additionally, a consistent increase in the middle-class population in the region is also one of the major factors contributing to this market's growth in APAC.

It is one of the goals that aims to make the pre-order system a global point of sale in the future by opening it for all passengers, not at certain points.

3.2.2 Internet of Things Data Market

MarketsandMarkets forecasts [2] the global IoT data management market size to grow from USD 27.54 Billion in 2017 to USD 66.44 Billion by 2022, at a Compound Annual Growth Rate (CAGR) of 19.3%.

The IoT data management market is gaining traction, due to widespread adoption of end-to-end data management platforms for capitalizing IoT data more efficiently and flexibly across organizations. The major factors driving the growth of the market include the modernization of data warehouse architecture, increasing adoption of data encryption for IoT device security, and rising need for data security and data traffic management. Opportunities such as the adoption of mature content models in enterprises, emerging unified metadata services, and efficient processes for increased Return on Investment (ROI) have boosted the adoption rate of IoT data management solutions.

The trending technologies, such as data analytics, blockchain, machine learning, and artificial intelligence, play a vital role in IoT data management. These technologies enable an enterprise to manage and share data with ease, remove single points of failure, and reduce costs. The emerging semantic interoperability of IoT data management solutions enables the interpretation and integration of large volumes of data formats using meta-tagged data.

The IoT data management market includes various vendors providing IoT data management solutions inclusive of data integration, data analytics and visualization, metadata management, data security, data migration, data quality, data governance, and data orchestration. Companies such as International Business Machines (IBM) Corporation (US), PTC Inc. (US), Teradata Corporation (US), Dell Technologies, Inc. (US), Cisco Systems, Inc. (US), SAS Institute Inc. (US), Hewlett Packard Enterprise (HPE) Company (US), Fujitsu Limited (Japan), Oracle Corporation (US), Google Inc. (US), and SAP SE (Germany) have adopted various growth strategies, including partnerships, agreements, and collaborations, to enhance their market presence.

According to [3], the global IoT market will grow from $157B in 2016 to $457B by 2020, attaining a Compound Annual Growth Rate (CAGR) of 28.5%. According to GrowthEnabler & MarketsandMarkets analysis, the global IoT market share will be dominated by three sub-sectors; Smart Cities (26%), Industrial IoT (24%) and Connected Health (20%). Followed by Smart Homes (14%), Connected Cars (7%), Smart Utilities (4%) and Wearables (3%). Figure 2 presents the size of the IoT market worldwide in 2014 and 2020, by industry (in billion U.S. dollars) [4].

---

[2] https://www.marketsandmarkets.com/Market-Reports/iot-data-management-market-53767032.html
[3] https://growthenabler.com/flipbook/pdf/IOT%20Report.pdf
[4] https://www.statista.com/statistics/512673/worldwide-internet-of-things-market/

Size of the Internet of Things market worldwide in 2014 and 2020, by industry (in billion U.S. dollars)
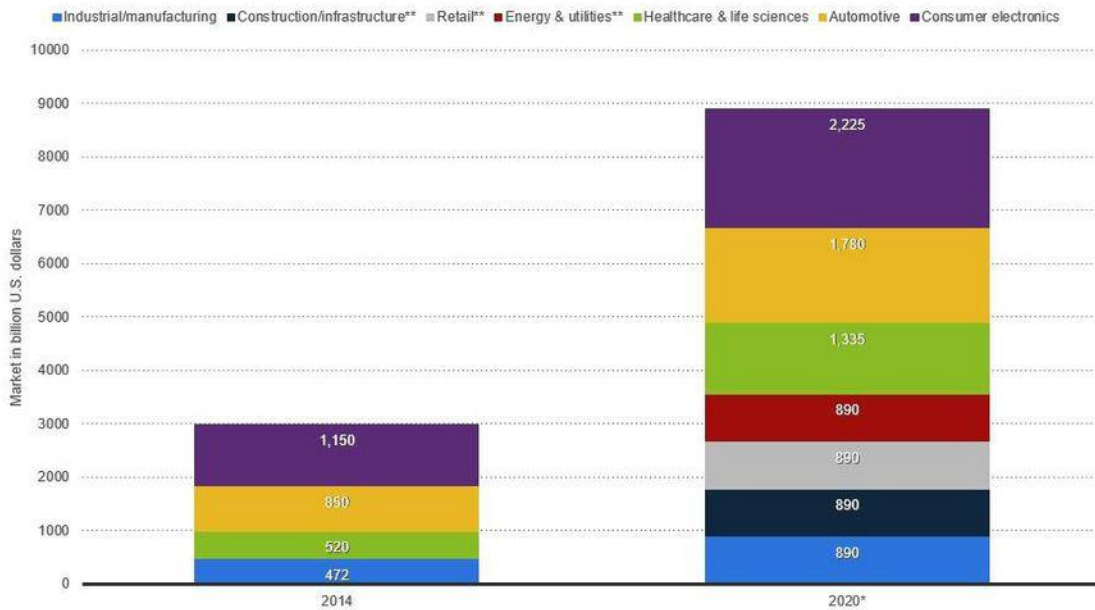
Figure 2 - Size of the IoT market worldwide in 2014 and 2020, by industry (in billion U.S. dollars)

Bain predicts [5] B2B IoT segments will generate more than $300B annually by 2020, including about $85B in the industrial sector. Advisory firm Bain predicts the most competitive areas of IoT will be in the enterprise and industrial segments. Bain predicts consumer applications will generate $150B by 2020, with B2B applications being worth more than $300B. Globally, enthusiasm for the Internet of Things has fuelled more than $80B in merger and acquisition (M&A) investments by major vendors and more than $30B in venture capital, according to Bain's estimates.

3.3 Business model

The results of the project can be valuable in several ways:

1. Improving the current B2B/B2C modules or website/webapps:

   - Learn and improve the quality of the products descriptions, thus helping the e-commerce community increase their selling by providing better targeted products.
   - Learn and improve the quality of the visitors' knowledge thus helping the e-commerce community increase their selling by providing products that match better the user needs.

2. Provide a B2B solution that will be the result of the research in Papud.

---

[5] http://www.bain.com/publications/articles/choosing-the-right-platform-for-the-industrial-iot.aspx

3.3.1   Concrete examples

Partner: Pertimm

Pertimm is providing a search solution using product descriptions, and its solutions is all the more relevant as the content of the products description is well structured. This leads to a better services for our current and future customers.

Pertimm is also providing a recommendation solution using products descriptions and user information to give recommendation that can be automatic, real-time and personalized. Learning better relations between products and customers using machine learning technologies will be a major improvement over a manual or semi-manual approach.

Partner: Press'innov

Press' Innov is a company which objective is to develop new technologies for media analysis, create new innovations, and value added for media content. Press' Innov has worked and still working on several European projects where we were among the main players. Through its own Research & Development team, Press' Innov develops core technologies and the majority of its own product concept.

Partner: Setur

Setur is the Turkey's most established tourism companies. Setur is a company which objective is to develop new technologies, create new innovations, and value added for tourism and duty-free market.

By using the most up-to-date and pioneering technologies in the field of tourism, we are making new projects to become a global player in digital environment. We aim to improve our operational strength, know-how and experience in both institutional and individual travel management.

Partner: KoçSistem

KoçSistem as a system integrator is one of the major market players in Turkey and has access to a wide span of markets including finance, automotive, energy, retail, telecommunication. The outcome of the project will be to develop a deep learning based recommender system dedicated to retail sector. The project will provide KoçSistem new opportunities according to its 2020 strategic digital conversion technology road-maps. As an R&D department, it will be able to perform innovative studies by using the new deep learning technologies based on e-Commerce use-case.

4   Market analysis and business model for UC2 – call center

4.1   Market Value Chain for UC2

We will describe the impact of UC2 in the PAPUD project on the market value chain based on the following diagram (Figure 3):
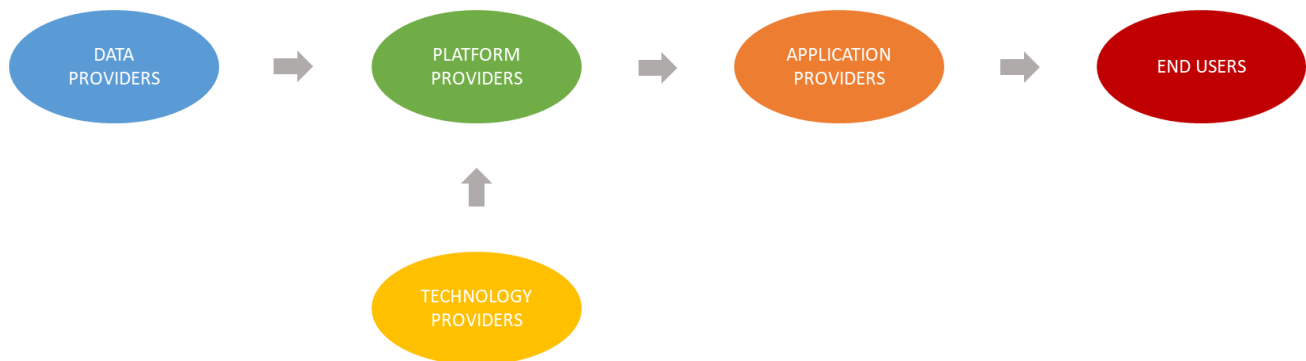
Figure 3 - Market value chain for UC2

The contribution of each partner to the value chain blocks is described below:

- **Data providers**: the main data source for this use case will be call center data. These can be audio recordings of customer calls, textual logs of the calls, but also e-mails or other textual communication between a customer and a service center. 4C and Ericsson can provide such data.
- **Technology providers**: the focus of all partners involved in UC2 will be to contribute as a technology provider. We will all contribute and collaborate in the development of deep learning algorithms on call center data. These deep learning models should enable companies to extract more insights from the interaction with the customer during a customer service touchpoint. On top of that, we want to develop models that improve the efficiency in handling customer cases.
- **Platform providers**: as a telecom infrastructure vendor Ericsson will focus on building solutions for communications service providers (CSPs). As such, they will create a telecom data model to be introduced into the platform. The data model to be introduced will be based on Mobile network metrics and consumer service usage details based on 3-4G networks.
- **Application providers**: The deep learning models developed during the PAPUD project are intended to be used as an add-on to existing service center applications. Building an entire service center application from scratch is complex, and out of the scope of this project. However, 4C will build a visualization app, with reports and dashboards for their developed NLP models. Turk Telekom will also build a user interface and a reporting application to facilitate the implementation of the NLP models into call center work processes.

- **End users**: the end users will be business experts and service agents. Business experts will benefit from the insights that are extracted from the large volumes of unstructured call center data. Using these insights, they will be able to better understand the needs of their customers, and guide future decision making. Service agents will also benefit from the developed models, as they will be able to work more efficiently. They e.g. be assisted in case logging, routing, etc. and will thus be able to handle cases faster.

4.2 Market Analysis

4.2.1 Call/Contact Center Market

According to [6] the global contact center market will grow steadily during the 2016 – 2020 horizon and post a CAGR of almost 11% by 2020. This market research analysis identifies the digital transformation of contact centers to enhance customer services and revenue generation as one of the primary factors that will have a positive impact on the growth of the call center market in the coming years. The increased adoption of mobile devices such as smartphones and tablets and the rise in usage of social media platforms, induces enterprises to adopt next-generation IT solutions. Focusing on reducing their ownership costs and enhancing customer services, customer contact centers will start preferring cloud computing solutions and will also concentrate on implementing advanced biometrics systems. Moreover, end-to-end digitization in several industries such as retail, banking, telecom, and healthcare will also compel contact center service providers to offer improved services using advanced technologies.

This global market is characterized by the presence of several vendors spread across the globe. Major vendors such as Avaya, Cisco Systems, Genesys, and Huawei Technologies have a wide portfolio of products, services, and solutions and dominate the market. Market players are mainly focusing on providing advanced technologies that help in integrating customer engagements across all channels and devices. One of the major contact center service providers Avaya, is a leading provider of real-time business collaboration and communications solutions and provides contact centers, unified communications, networking, and other related services to several companies across the globe.

The leading vendors in the market are: Avaya, Cisco Systems, Genesys and Huawei Technologies.

The other prominent vendors in the market are 3CLogic, 8x8, Aastra, Altitude Software, Aspect Software, Connect First, Convergys, CRMXchange, Five9, Fujitsu, HP, IBM, inContact, Interactive Intelligence, Mitel Networks, NICE, Noble Systems, Oracle, Presence Technology, SAP, ShoreTel, Syntel, Teleopti, Unify, Verint Systems, Verizon Communications, Vocalcom and Salesforce.

According to this market study, the inbound systems segment accounted for the major shares and dominated the call center services industry during 2015. The inbound systems contact center is one of the most flexible, scalable, and comprehensive contact centers and enables agents to easily communicate with customers through various channels such as emails, telephone calls, chats, and

---

6   https://www.technavio.com/report/global-it-spending-region-and-industry-global-contact-center-market-2016-2020

social media. Benefits such as resource utilization, optimal customer experience, and automatic call distribution routing, will drive the growth of the market segment during the 2016 – 2020 timeframe.

Technological advancements have led to the increased adoption of hosted services in the recent years. The adoption of hosted services eliminates the need to spend a huge capital on on-premises infrastructure, reducing the total cost of ownership of IT systems. Additionally, hosted services offer several other advantages including improved technology uptime, access to a single integrated customer contact platform, and increased market agility. Owing to these benefits, there will be a major shift in customer preference towards hosted solutions, which will induce on-premise solution providers to offer hosted solutions.

According to [7], the contact center software market size is estimated to grow from USD 13.27 Billion in 2016 to USD 29.13 Billion by 2022, at a Compound Annual Growth Rate (CAGR) of 13.8% during the forecast period (2017–2022). The base year considered for the study is 2016 and the market size has been projected from 2017 to 2022. A major factor driving the growth of contact center software market is to deliver enhanced experience to customers through multiple channels including voice, video, web, and social media.

The contact center software ecosystem comprises vendors, such as Genesys Telecommunications Laboratories, Inc.(US), Cisco Systems (US), Avaya Inc. (US), Mitel Networks Corporation (Canada), Enghouse Interactive (US), SAP SE (Germany), Five9, Inc (US), Huawei Technologies Co., Ltd. (China), Alcatel-Lucent Enterprise (France), Oracle Corporation (US), 8x8, Inc. (US), Unify Inc. (US), Drishti Soft Solutions Pvt. Ltd. (India); software providers; cloud service providers such as Salesforce (US); IT consulting service providers; and system integrators.

Professional services, including consulting, training and support, and integration and implementation of contact center software, are a widely used service type in the contact center software market. Managed services have a higher adoption rate than professional services. Managed services empower organizations to focus on their core business functions while delegating contact center operations to professionals. Managed service providers ensure the efficient management of contact center processes and simultaneously deliver enhanced customer experience through multiple channels.

Contact center software and services are being increasingly adopted across various industry verticals, such as banking, financial services, and insurance (BFSI); retail and consumer goods; travel and hospitality; and healthcare. The increasing adoption rate is expected to drive the global contact center software market. The retail and consumer goods vertical is expected to grow at the highest CAGR during the forecast period. Figure 4 depicts the Contact Center software market (2017 – 2022).

---

[7] https://www.marketsandmarkets.com/Market-Reports/contact-center-software-market-257044641.html

Interactive Voice Response (IVR) solution is expected to grow at a slow pace in the contact center software market during 2017–2022

Largest Market in 2017 (estimated)

Market with the highest growth rate for 2017-2022 (projected)

Customer Collaboration is the fastest-growing solution in the contact center software market during 2017–2022

Figure 4 - Contact center software market (2017 – 2022)

North America is expected to hold the largest market share and dominate the market from 2017 to 2022. This is mainly because there is a high focus on innovations and technology adoption, especially in the developed economies of North America. The high adoption of cloud, mobile, analytics, and social media among enterprises across various industry verticals has helped contact center software market to grow in North America. APAC is expected to be the fastest-growing region in the global market. The growth in the region is attributed to the presence of a large number of SMEs and huge customer base.

However, the high initial investment incurred while the integration of new age solutions with legacy systems and the risk of information loss are anticipated to restrain the market growth. The implementation of contact center software helps organizations to communicate effectively with customers through various channels and to manage customer queries efficiently.

4.2.2 Speech recognition market

According to [8], the speech and voice recognition market is expected to be valued at USD 6.19 Billion in 2017 and is likely to reach USD 18.30 Billion by 2023, at a CAGR of 19.80% between 2017 and 2023. The increased demand for speech-based biometric systems for multifactor authentication is a driving factor for this market. The speech and voice recognition market has been segmented on the basis of technology into voice recognition and speech recognition. The market for voice recognition is expected

---

[8] https://www.marketsandmarkets.com/Market-Reports/speech-voice-recognition-market-202401714.html

to grow at a higher CAGR during the forecast period. The need for biometrics authentication leads to the growing integration of voice authentication in mobile banking and creates a growing demand for multifactor authentication in enterprises and industries. The market in APAC is expected to grow at the highest CAGR during the forecast period. The increasing awareness among masses and lowered costs of biometric devices drive the growth of the speech and voice recognition market in countries such as India, China, and Japan.

The major restraint for the recognition market is the lack of accuracy in speech and voice recognition systems in noisy and harsh working environments. The performance of speech recognition systems degrades drastically in noisy and reverberant environments. This is especially true when the background noise is itself speech. The effect worsens as the distance between the talker and the microphone increases.

The key players in this market focus on strategic partnerships and collaborations, acquisitions, agreements, contracts, and product launches to increase revenue. Nuance Communications, Inc. (US) is a leading player in the speech and voice recognition market, followed by Microsoft Corporation (US) and Alphabet Inc. (US). In March 2016, Nuance Communications, Inc. signed an agreement with TalkTalk Communications Ltd. (UK) to access Nuance's voice biometrics technology and improve customer experience in the automated phone system. This innovative authentication solution is being integrated with TalkTalk's intuitive interactive voice response (IVR) system, which uses Nuance's natural language understanding (NLU) technology to allow customers to speak naturally instead of having to navigate through a phone menu.

4.2.3 SWOT

To support the market analysis for this use case, we have performed a SWOT (Strengths, Weakness, Opportunities, and Threats) analysis. A brief summary can be found below.

Strengths

- Market knowledge: the partners included in the use case have a lot of combined expertise, ranging from telecommunication services to service center optimization.
- Market access: the partners included in the use case already have an existing customer base that fits the target group.
- Market trust: the support of KU Leuven to our use case will stimulate market trust since they have very deep expertise in NLP.

Weaknesses

- Training data: to build deep learning NLP models, a large set of training data is required. Access to these training datasets is not always straightforward, especially since various partners in the use case will build models in languages other than English.

- Sharing models: it will be a challenge to collaborate on NLP models built in separate languages (e.g. Dutch and Turkish). We need to find a way to optimally share results and findings given this challenge.

Opportunities

- AI hype: AI and deep learning are amongst the top technology trends in the world
- Market maturity: 4C has done research on the market maturity amongst its customers. All interviewed customers are collecting call center interaction data and are interested in using this data in a more automated and objective way through NLP.
- Market size: our use case is not restricted to a specific industry. Customer interaction through a service center occurs in practically every industry, such as banking, retail, telco, etc.

Threats

- Service center competitors: since deep learning is such a technology trend, it is not surprising that there are several competitors in the domain of service centers as well, such as: Salesforce service cloud Einstein, Zendesk Explore, Digital Genius, Thinknext, etc.

4.3 Business models

The results of the project can be valuable in several ways:

- **Internal valorization**: partners contributing to the use case can deploy the developed models to improve the interaction with their own customers. By doing this, they can increase their customer experience and increase customer loyalty.

- **External valorization**: a second option is that partners will use the results of the project in a B2B offering. That way, they can enable other companies to benefit from the developed models. There are 2 approaches when choosing this external valorization option:

  o **Consultancy mode**: The NLP models developed during the project can be offered to the market in a consultancy mode. In this mode, we can use the intelligence available on the platform for smaller and well-defined business problems. The results of this type of project will be a "one shot" (interactive) report with the requested business insights. In this mode, we don't deliver code for re-use or updates of the report. We believe the consultancy mode is an easy way for customers to have their first experience with AI. If they want to explore more, they can take the next step and let us set-up a platform of their own.

  o **Product mode**: In this mode, new inputs are analysed continuously, (interactive) reports are updated automatically and the models are updated and refined on a regular base. Necessary connections between the platform and the customers' databases and operational systems (such as e.g. Salesforce Sales, Service or Marketing Cloud; ERP systems, etc.) will be developed. In this product mode, we also foresee a

part where we focus on the processes of the customer, change management and training. Installing an AI platform is one thing; however, this does not guarantee the optimal usage.

The 2-pronged approach is introduced because they tend to serve another purpose and different target audience. The goal, purpose and target audience are presented in Table 1.

Table 1. Goal, purpose and target audience

	Consultancy mode	**Product mode**
Goal	Provide insights	Provide Insights and automation
Purpose	Make AI accessible for smaller organizations and deliver them insights. Gain trust of AI for larger organisations	Focus on process automation and optimization
Target audience	Small and midsize enterprises	Midsize and large enterprises

Table 2 – Initial business Model envisaged for UC2

4.3.1   Concrete examples

Partner: Ericsson

In today's fast changing market, operators need to evolve infrastructure frequently to cater for new demands and increase competitive advantage. Meanwhile the ever-changing market offerings is increasingly attractive for customers to churn. The ultimate goal of telecom vendors such as Ericsson, is to ensure high performing telecom operations that deliver high quality end-user experience for their subscribers.

Leading operators strive to handle churn by using the complexity of technology. Ericsson combines technology leadership with a unique mix of global and local expertise to enable operators and enterprise customers to meet their promised service level towards their subscribers. So Ericsson takes advantage of Ericsson global support insights and innovative approaches to avoid high number of churns which has a big negative affect of the OPEX (Operating Expenses).

Since churn can be triggered from different points where customer is in touch, any improvement in the service level is very crucial. Ericsson works on predictive analytics and deep learning techniques on call centre data from the mobile operators to understand and prevent the potential churn candidates. Ericsson's support services are presented in Figure 5.
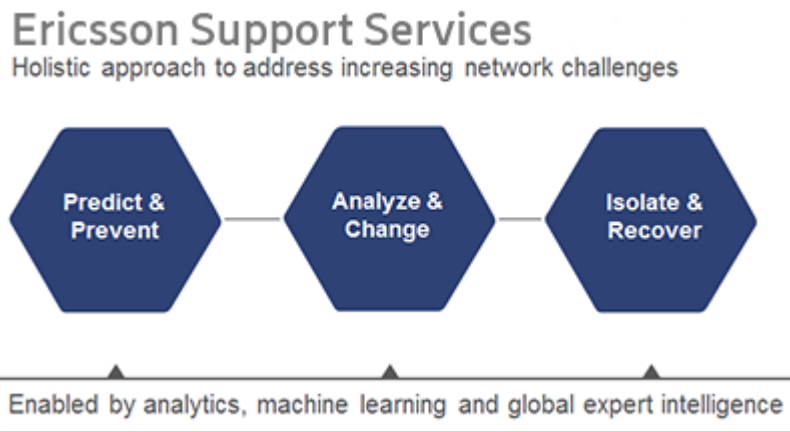
Ericsson Support Services
Holistic approach to address increasing network challenges

Predict & Prevent — Analyze & Change — Isolate & Recover

Enabled by analytics, machine learning and global expert intelligence

Figure 5 - Ericsson Support Services

With the capabilities that Ericsson will gain from PAPUD project, our operator customers will achieve a predictive solution that can be used in different call centers of the operators that Ericsson serves more than 3 Billion end users. We deliver through close collaboration with our customers and building on local presence with focus on business performance and risk mitigation. Our goal is to help our customers get close to Zero-churn services and products in the market.

Partner: 4C

4C will use the models developed during this project in their AI platform, TellMi. TellMi is a service of 4C in which they offer AI and NLP services to extract valuable insights out of unstructured data. Even though the TellMi platform is use case independent, a tangible project scope needs to be defined. Accordingly, we will focus on two specific use cases: actionable customer feedback system and service centre efficiency improvement. Here we will only discuss the service center use case.

When customers contact companies for e.g. more information, complaints, etc., these interactions are often recorded and stored to further improve their services. By determining the underlying fine-grained topic and sentiment with its corresponding reason, TellMi can automate the logging process and define the urgency of incoming messages, across English, French and Dutch. The system can also learn to suggest appropriate answers to incoming questions, which will increase the knowledge and efficiency of the contact centres. By identifying re-occurring complications/reasons of contact regardless languages, the business can take pro-active actions to reduce the number of contacts and increasing customer satisfaction.

4C will use both the consulting as the product mode to valorise the results of the PAPUD project. For companies who have collected a large set of call logs, e-mails, web messages, etc. who would like to dive into it in a quantitative way, 4C will perform an NLP consulting project. With the results, possibly linked with other customer data, businesses can take pro-active actions to improve customer service and increase the efficiency of the contact centres.

For companies with a large contact centre, who receive many incoming messages on a regular basis, 4C will set-up the TellMi platform in a product mode. The information hidden in call logs, e-mails, web message, etc. will be extracted in real time. The advantage of the product approach is the automation of the process, enabling urgent topics and/or unhappy clients to be treated first, reducing the average handling time, etc.

Partner: Turk Telekom

A Network Service Provider can receive different network related problems and complaints from various communication channels regarding their service activity in certain regions.

One of the major and most important communication channels is customer call center. Detecting customer related problems that customers are notifying during these calls are significant in order to provide solutions and increase customer satisfaction. However, due to sheer volume of the call records that are converted to text, it is quite difficult to analyze whole data using traditional approaches.

In this use case, we'll study a topic modeling approach for detecting customer related problems from call center text data. The analysis results will be used from a TT's customer call center company AssisTT's call center reports of all calls related to Türk Telekom's customers. These communication channels help customers to pass their comments, suggestions and complaints information about a particular service and product into the service provider's agents. At the service provider's premises, all the information gathered from each communication channel is first converted into text based documents and then stored inside service provider's data center. After all data are gathered at data center, analysis of service related problems is performed as shown in Fig. 1.

During analysis, customer related problems can be detected using deep learning. The methodology aims to increase the customer satisfaction via accurate analysis of all the customer related data received daily from various communication channels. This will also help in reducing the man-month efforts spent for customer analysis inside service providers.

We consider applying text mining and deep learning approaches in the telecommunication domain using real data set of customer call records. In this use case, we'll investigate topic modeling techniques for extracting customer related problems using the call center text data of AssisTT.

5 Market analysis and business model for UC3 – Recommendation system for Human Resources

5.1 Market Value Chain for UC3

The Market Value Chain envisaged for the Recommendation System for Human Resources departments in PAPUD project is presented in Figure 6.
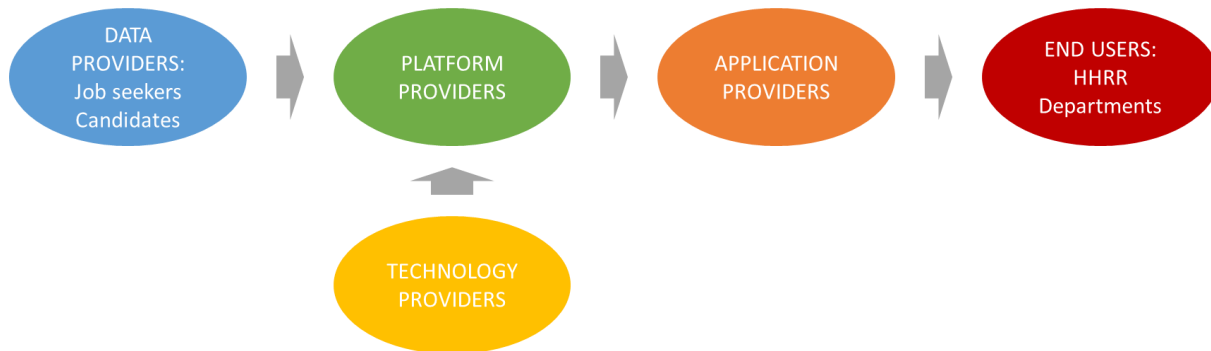
Figure 6 - Market value chain for UC3

In particular, the demonstrator for this use case will be planned in Spain and Romania, particularly in HI-Iberia HHRR Department. In consequence, HI-Iberia will participate in all the market value chain as follows:

- **Data Providers**: HI-Iberia HHRR Department will provide the data available within the company data bases about candidates and applicants for job vacancies. This data will be provided in the languages available (Spanish and English), for the data processing of this data the information will be anonymized so just ids (tokens) will be available.

  IMT will collect and provide the data from Online Social Networks such as Twitter and Facebook, to be used by technology providers in this use case.

  Within this use case, BEIA will provide data from online social networks which will be used as input by technology providers.

- **Platform Providers**: HI-Iberia will contribute to the development of PAPUD Platform by applying the technologies envisaged in PAPUD project in the areas of NLP, Big data, deep learning to the raw data.

  Beia will contribute to the development of the PAPUD platform by integrating technologies for Natural Language Processing (NLP), Big Data analysis and Machine Learning.

- **Technology Providers**: HI-Iberia will provide deep learning and big data algorithms and techniques to the platform for the data processing and information extraction.

Beia will contribute to the development of technologies for Natural Language Processing (NLP), Big Data analysis and Machine Learning. Regarding NLP, Beia will focus on the development of instruments for Romanian language processing.

- **Application Providers**: HI-Iberia will provide the relevant applications and interfaces particularly for the Human Resources Use Case. The output of this use case will be a web application for this department.
To facilitate Big Data processing, Beia will provide an application for the analysis and extraction of relevant information from various sources such as CVs and Social Media.

- **End User**: HHRR Departments will be in charge of the usage of the PAPUD Platform obtained along the project execution and so, it will be defined taking into account their particular needs and requirements.
Beia will provide information regarding the needs and requirements specific to the Romanian recruitment market. In order to maximize the market penetration rate during exploitation, the PAPUD platform needs slight adaptations for specific markets. Beia will adapt the solution mainly for the Romanian market to enhance its local socio-economic impact.

5.2 Market analysis

5.2.1 Recommendation system for recruiting purpose

In the recruiting and staffing industry, fast and accurate candidate matching is an enormous competitive edge. In many cases, presenting the right candidate before your competitor identifies the same person in their candidate database is crucial. Given the large amounts of candidates that most recruiters deal with, it is easy to lose track of the best matches between the requirements and the candidates' skills.

Nowadays there are many "intelligent" candidate search and match solutions for the recruiting industry, delivered either as stand-alone products or embedded in other applicant tracking systems. These can provide convenience, but are limited to basic resume parsing and matching the corresponding metadata to a job posting. Thus, most of these embedded search systems failed to significantly improve fill rates or provide a simpler, better search experience for recruiters.

This is where powerful big data recruitment applications can enable recruiters to know more and hire better. The approach of combining advanced search techniques and the processing power of big data analytics to provide sophisticated statistical and linguistic capabilities for understanding the applicants' profiles and identifying the best candidates. This accurate, automated big data platform will bring a highly interactive search environment to recruiter as well as the reduction of the workloads in the department. Regarding the success of the application, it is ensured as every company in the world needs new staff continuously so the recruitments process takes place every day.

CV parsing, also called resume parsing or CV extraction, is a process that uses a powerful parsing engine to extract key information from a candidate CV or résumé file to automatically create and populate relevant areas of a candidate record within your recruitment software database. Although complicated; this process is normally very quick - much faster than you or your colleagues could cut

and paste the information from the CV themselves. The overall benefit of this smart technology is a dramatic reduction in the amount of time your entire agency spend manually entering data onto your recruitment database and therefore increasing the overall productivity of your entire team.

Current tools in the market are mainly focused in the parsing of CVs and the creation of new records for future searches. Some of these competitors in the area of CV parsing are:

- RecruitmentWorx: It offers a range of powerful and time-saving features designed to make recruitment process quick, efficient and so simple by providing on demand access to your talent pool wherever and whenever you need it, by means of some form of CV parsing functionality built into the core system. This makes it possible to use the data extracted from a candidate CV to automatically create new candidate records and populate your recruitment database in just a few seconds.
- Textkernel: It offers high quality multilingual CV parsing and semantic search, sourcing, lead generation, matching tools and labour market statistics to accelerate and improve the process of matching demand and supply in the job market.
- HireAbility: It provides cloud-based CV/Resume parsing software and Job Order parsing software. It is a resume parser that uses pattern recognition, language structure, and artificial intelligence to extract the content of a résumé (CV) or a job order (vacancy) into a predefined format (e.g. HR-XML, JSON)

Social sourcing, also called social recruiting, is a type of e-recruitment in which recruiters and HR managers search for job candidates by using social media sites, predominantly Facebook, Twitter and LinkedIn. It can also refer to the process of gathering referrals from employees to staff a position or posting an opening on social media websites.

Social sourcing can help recruiters find candidates who are not actively looking for jobs but who might consider the "right" opportunity. Such candidates, who are sometimes referred to as passive candidates, are often employees with skills that are in high demand.

Below also are some more tools regarding profiling of users and social sourcing in job market. However these tools do not dig deep in semantics and they do not follow the tracks of the users on the web to analyze their interest, competencies and behaviors in full manner.

- Talentbin: It is a large database with millions of profiles. It focuses on finding passive candidates through boolean search and social media recruiting. They provide a lot of information from candidates' social media accounts to help you reach out to them.

- Zillionresume: It is an aggregator database, gathering resumes from thousands of other sources. It gives you an opportunity to discover 'hard-to-find' candidates. The platform can provide you with lists of resumes that match your criteria.

- Hired: It is a platform designed to bring employers and job seekers together. Employers can create a company profile and browse candidates. They can express their interest in candidates through the platform, where, according to Hired's website, candidates answer 95% of requests.

Starting from this analysis, PAPUD project intends to provide qualified search and analysis to job market with extensive semantics produced from different sources such as web tracks, interviews,

online profiles, CV's on the system and such, and focuses on producing a complete profiling of a candidate comprising its interests, competencies and recommendations based on these.

5.2.2 Social Media Monitoring for enterprises

The prominence of Social Media as a mass of communication is part of today's daily lives for citizens worldwide. Communities and personal relations of all sorts are now inextricable from the Internet tools that have appeared over the last decade. Impacts are felt everywhere: from news spreading through personal relationships and artistic movements; Social Media has grown to cover a majority of the spectrum of human activities. Thus, the use of social media has moved from experimentation phase to a more mission-critical activity, placing significant pressure on social programs to efficiently monitor, analyze and engage this variety of information.

The number and quality of Social Media Monitoring (SMM) tools available in the market has also changed considerably with the SMM technology maturing rapidly and delivering unique and more sophisticated analytical capabilities. New tools and services are emerging continually, while already established SMM technologies are frequently improving their products by introducing new features and coverage to accommodate their clients' needs or transforming their point solutions into a complete social suite to address multiple features within one single platform.

The market of Social Media Monitoring constitutes an emerging market due to the fast-growing of social media and the need of these tools to adapt to it. Although currently there are a lot of tools for tracking social media for different purposes, this market will become greater in the following years driven by the following trends:

- **Increase emphasis on social media analytics**. The biggest challenge for all is how to manage and act upon the user data collected through social media.
- **Multi-platform usages** of Social Media for both mobile and desktop devices.

Specifically, regarding Social Media Monitoring (SMM) for enterprises, there are a number of available tools and service providers currently in the market. Some of them are described in this section:

- *BoomSoonar* provides an integrated social media business platform, which empowers organizations to monitor, measure, manage, analyze and respond across social media channels and Web. The platform offers a real time web and social media monitoring, analytics and engagement platform with content categorization and social CRM modules that helps organizations gain a complete understanding of the online conversations about their company, products, topics or competitors, gather data for social CRM and engage directly from within the BoomSonar's platform.
- *Brand24* is a technology company that offers social CRM and social monitoring platform, which empowers brands to get real time customer insights and actionable intelligence relevant to their products, brands, competitors, person, or topic of interest across multiple social media platforms including: Facebook, Twitter, YouTube, blogs, message boards and news sites.
- *BrandsEye* provides an online and social monitoring, listening and analytics tool that helps business and organizations track and evaluate their online presence and extract valuable insights out of the data to better inform their strategic business decisions.
- *Cision* is leading global media intelligence company that offers the industry most comprehensive public relations and social software, rich analytics and global insight team to enable communication, social media and content marketing professionals to plan, execute and

measure influencer-oriented campaigns in one integrated platform to improve their marketing and make more data driven decisions.

After an analysis carried out with a set of around 200 social monitoring tools, the conclusion is that most of these tools provide very similar features:

- Extensive media coverage (Twitter, Facebook, Instagram, etc.)
- Real time search and monitoring
- In-depth semantic analysis
- Profiling analysis
- Sentiment analysis
- Trend analysis
- Geo-location, to do precise searches on certain locations
- Filtering of results by media type topic/keyword, time period, demographics, sentiment, custom tags or geography (geo-location based on the combination of self-reported location information and mapping of geo-enabled smartphones).
- Results visualization, extensive range of graphs and charts for interpreting appropriately the results.
- Natural Language Processing available in several languages.
- Data alerts with standard and fully customizable alerts.

However, other features more specialized as easy to use interface, content analysis (multimedia analysis – intelligent audio processing) social analysis or machine learning techniques are difficult to be found in current tools available in the market. In consequence, PAPUD will try to merge the data extracted from social monitory with the information from candidates in order to provide an overview of the Psychosocial profile of the applicant.

5.3 Business model

The idea of providing a Recommendation system for Human Resources starts from the internal need of HI-Iberia HHRR Department when looking for candidates for a specific job offer. Nowadays, the department receives CVs and information from different sources and in many times, merging and linking the skills becomes a hard and time-consuming task. The analysis of candidates and needs for a concrete job vacancy consumes a lot of the company resources. In consequence, the initial idea of exploitation for this use case is to deploy and set up the project results internally in the company to ease and optimize the current recruitment process.

In the near future and once the produce becomes stable and produces tangible outcomes we will consider the option of launching to the market for improving other companies recruitment processes. For sure its use and future exploitation by other PAPUD partners will be also welcome.

6 Market analysis and business model for UC4 – Behaviour Analysis for Reverse Efficient Modeling (BAREM)

6.1 Market Value Chain for UC4

BAREM use case is focused on user's activity on eServices to extract defect and misusing. The market value chain for this use case is presented in Figure 7.
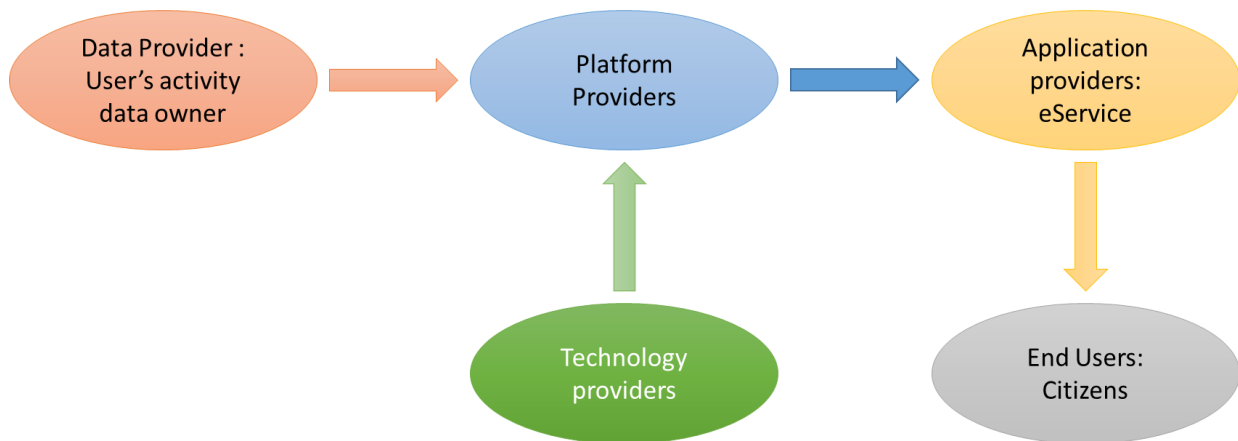
Figure 7 - Market value chain for UC4

The market value chain on user's activity analysis can be described like this:

- User's activity data owner : Organization collecting user's activity
- Technology providers for log analysis : Academic or Industry technology providers
- Platform Providers: Platform provider for deep learning analysis and other technologies
- Application providers : In this use case, application providers are eService designer interested in log analysis to enhance its eService
- End user : Citizen using eServices

In particular the different links of the chain will be covered by the following partners in UC4:

- **Data Providers**: Softeam will provide data for this use case (User's activity data owner)

- **Technology Providers**: The University of Lorraine (LORIA) will analyse e-Citiz users' activity files in order to extract navigation-related information and, by applying deep learning analysis, will try to infer defective or improper use of the e-Citiz application. The University of Lorraine (LORIA) will act as a technology provider.

  IMT will work on semantic analysis of social networks data related to user services using deep learning techniques in order to infer the service's problems stated by social networks users and will act as a technology provider in the value chain.

Lille University will work on sentiment extraction from log files based on video footage recorded during workshops with volunteer users. This work will find correlation between sentiment and user's activity and will enable PAPUD to extract sentiment from log files.
Lille University will act as a technology provider in the value chain.

Note: Performetric may participate in this use case on the sentiment analysis part by proposing its tools to get more inputs from the user like mouse tracking, keyboard listener …

- **Platform Providers**: Atos will provide the Deep Learning platform Infrastructure and act as Platform provider.

- **Application Providers**: Softeam publishes several e-Service application (with its tool named e-Citiz) used by citizens. These applications generate a lot of user activity trace describing their application use. Softeam will provide eService to enhance (eService provider).

- **End Users**: Citizen using eServices

6.2 Market analysis

6.2.1 EServices Market Analysis

EServices are everywhere in every aspect of our day-life. It can be the eService for booking a ticket for sporting event or a concert, it can be the eService for ordering food for lunch time, or it can also be the eService for declaring your income to the government to compute tax amount.

In the eGovernment sector, eServices offer a lot of enhancement but it can make massive cost savings. In Denmark, for example, electronic invoicing saves taxpayers €150 million and businesses €50 million a year[9]. Estimates suggest that improving online information would cut the time and resources needed by public administrations to answer requests and potentially yield a reduction in annual costs for national administrations of as much as €500 000. For businesses the potential savings at EU level just from cutting red tape by means of better online information are between €4.4 billion and €50.4 billion. Full exploitation of public sector data has the potential to reduce administrative costs significantly: for the 23 largest EU Member State governments there are estimated savings of 15 % to 20 %, equivalent to €150 to €300 billion and accelerating annual productivity growth by 0.5 % over a decade. The main benefits are greater operational efficiency (owing to improved transparency), increased tax collection (owing to customised services), and less fraud and fewer errors (owing to automated data analytics). Moreover, the full implementation of a 'digital by default' strategy in the public sector (all services being provided only digitally) could result in annual savings of around €10 billion at EU level. All these beneficial effects are likely to be multiplied by their cross-border availability as this increases administrative efficiency still further[10].

---

[9] https://ec.europa.eu/digital-single-market/en/policies/egovernment
[10] http://www.europarl.europa.eu/RegData/etudes/BRIE/2017/608706/EPRS_BRI(2017)608706_EN.pdf

In 2016, in OECD countries [11], about 36% of individuals submitted online forms via public authorities' websites. Since 2006 the number has tripled [12]. The number of people using eGovernment services just between 2014 and 2016 increased by 8%, as presented in Figure 8.
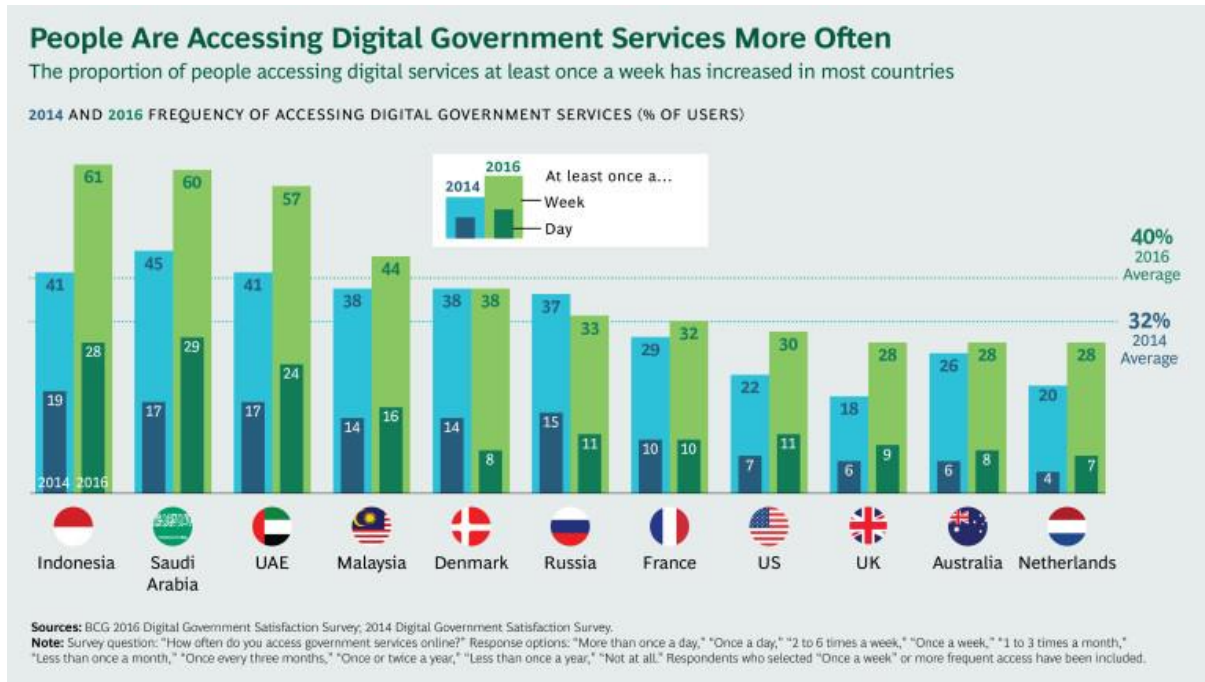
Figure 8 - Statistics on the number of people using e-Government services

In 2014, MARKESS estimates more than 5.2 billion euros in the French market for software and IT services associated with dematerialization[13]. McKinsey&Company estimated in 2014 that capturing the full potential of government digitization could free up to $1 trillion annually in economic value worldwide, through improved cost and operational performance [14].

---

11 http://www.oecd.org/about/membersandpartners/list-oecd-member-countries.htm
12 http://www.europarl.europa.eu/RegData/etudes/BRIE/2017/608706/EPRS_BRI(2017)608706_EN.pdf
13 http://blog.markess.com/2015/07/dematerialisation-un-marche-de-plus-de-5-milliards-deuros-en-france
14 https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/public-sector-digitization-the-trillion-dollar-challenge

As eServices usage is growing, citizen expectations rise. A 2016 study[15] shows that more than 60% of users still experience a problem when accessing services, so governments can continue to improve. Users' problems when accessing services are presented in Figure 9.
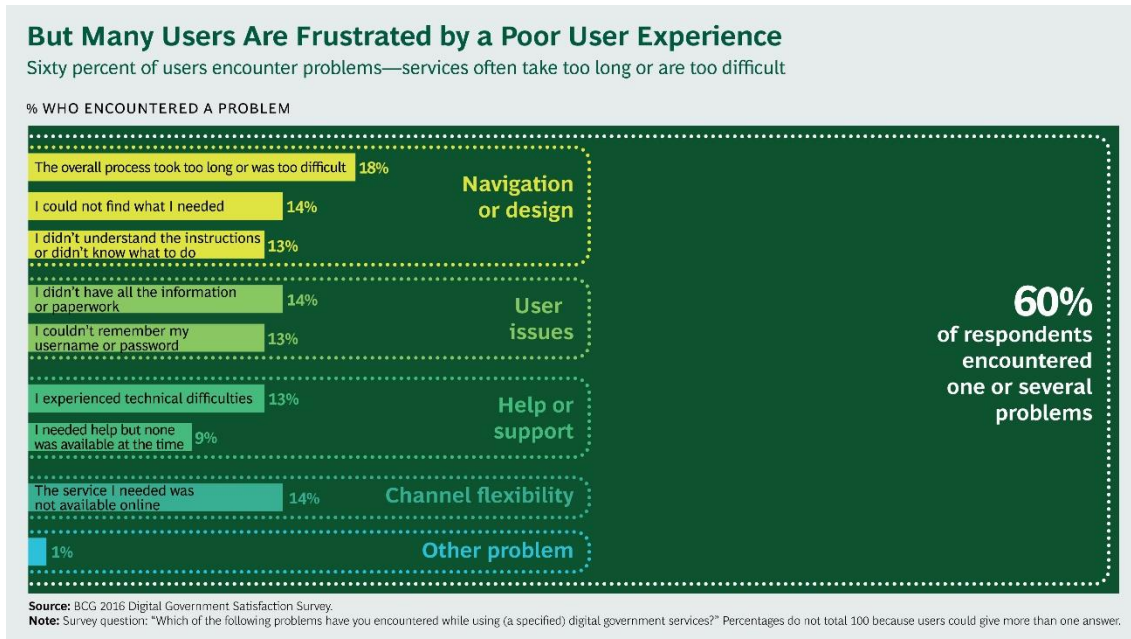
Figure 9 - Users problems when accessing services

PAPUD with the BAREM Use case will try to address the main problems when citizen accessing eGovernment eServices.

6.3 Business model

Softeam plans to develop and include a new feature in its e-Citiz product enabling PAPUD analysis in generated e-Services. The main benefit of this feature will be to provide the eService owner with a way to improve the user experience without the need for a survey or bug tickets and moreover, with the checking tool some error fixing will be done automatically. In a market where user feedback is rare this type of tool can be of strong interest to our customers.

A detailed business model will be studied at the end of the project.

---

[15] https://www.bcg.com/publications/2017/government-digital-services-by-numbers.aspx

7 Market analysis and business model for UC5 – HPC

7.1 Market Value Chain for UC5
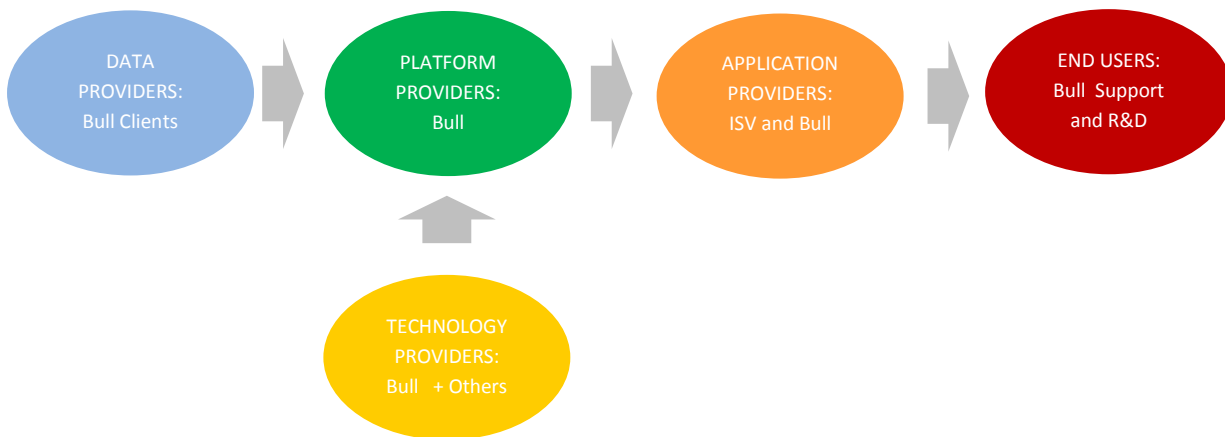
Figure 10 - Market value chain for UC5

Bull is providing large HPC systems mainly to European clients and is facing the maintenance issues linked to the size of this equipment. An HPC cluster can include up to thousands elements (cluster up to 64000 nodes) including compute nodes, switches, storages systems, administration nodes and so on all subject to failures that are impacting the reliability and availability of the clusters.

Today's maintenance is based on statistics on elements failures and replacement or repair of the elements in case of failure.

An analysis of the logs joined with the failures could give rules for predictive maintenance and allow to replace the cluster components before the potential failure and increase the availability of the whole computing cluster.

DATA PROVIDERS: In the UC5 the data provider is one major HPC Bull client bringing a whole set of data including logs (NLP) and timestamps.

PLATFORM PROVIDER: Bull is the HPC Hardware platform provider and the deep Learning framework will be based on our own packaging over market solutions (Tensorflow …). We except feedback from our use case to enhance our hardware and software platforms in order to increase our competitivity in the HPDA market

TECHNOLOGY PROVIDERS: Bull uses his own technology based for the hardware based on Intel and Nvidia components for the Hardware part and Bull uses the different framework of the market for DL with our own integration.

APPLICATION PROVIDERS: The DL models will be created with the help of **LORI**A and used to enhance our support strategy . The support uses applications froml the market and Bull aplications.

END USERS: The support and the Bull R&D will use the ouput of the DL Training to enhance the support applications, modify the support strategy and enhance the platforms conception to increase the platforms availability thus reducing the exploitation costs.

7.2 Market analysis

7.2.1   The HPC Market

According to [16], the global high performance computing (HPC) market is expected to grow from USD 32.11 Billion in 2017 to USD 44.98 Billion by 2022, at a Compound Annual Growth Rate (CAGR) of 7.0% during the forecast period. Data center operators face increased needs for data center architecture which supports complex applications, in every business process and research. The HPC technology provides excellent computing power and enhanced performance for running advanced applications. It uses parallel supercomputers and clusters of computers where aggregation of computing power is needed for delivering higher performance. These solutions are fuelling the growth of the HPC market, as organizations are adopting them to cater to their complicated business needs. The deployment type for the high performance computing market consists of on-premises and cloud. The cloud deployment type is expected to grow at a higher CAGR. The cloud deployment type helps minimize an organization's operational costs by eliminating the need to buy an additional computing resource. Moreover, the cloud deployment type comes with cost benefits, which has enabled small enterprises to take advantage of its benefits. The market is also segmented on the basis of organization sizes into Small and Medium-sized Enterprises (SMEs) and large enterprises. The large enterprises segment is expected to have a larger market size, due to a high demand for enhanced HPC systems in their facilities. The growing investments in data center technologies across the world and increasing number of data centers are further fuelling the demand for HPC solutions and facilitating the growth of the high performance computing market.

The market in North America is expected to hold the largest market share, due to the presence of major vendors, and increasing adoption of associated services. The APAC region is expected to provide several growth opportunities in the HPC market and is projected to grow at the highest CAGR during the forecast period.

However, the cost for setting up an HPC and maintaining it is high and it may restrain the market growth. Various vendors provide HPC solutions to help enterprises reduce their Capital Expenditure (CAPEX). HPE is one of the major vendors of HPC solutions. It offers solutions for solving large scientific, engineering, and data analysis problems. Other major vendors in the high performance computing market include AMD (US), Atos (France), AWS (US), Cisco Systems (US), Cray (US), DDN (US), Dell (US), Fujitsu (Japan), Hitachi Vantara (US), Huawei (China), IBM (US), Inspur (China), Intel (US), Lenovo (US), Microsoft (US), NEC (Japan), NetApp (US), NVIDIA (US), Oracle (US), Panasas (US), Penguin Computing (US), Spectra Logic (US), Sugon (China), and Xilinx (US). These market players have adopted various strategies, such as partnerships, collaborations, and expansions, to remain competitive in the HPC market. The worldwide HPC server market status is presented in Figure 11.

---

[16]   https://www.marketsandmarkets.com/Market-Reports/Quantum-High-Performance-Computing-Market-631.html
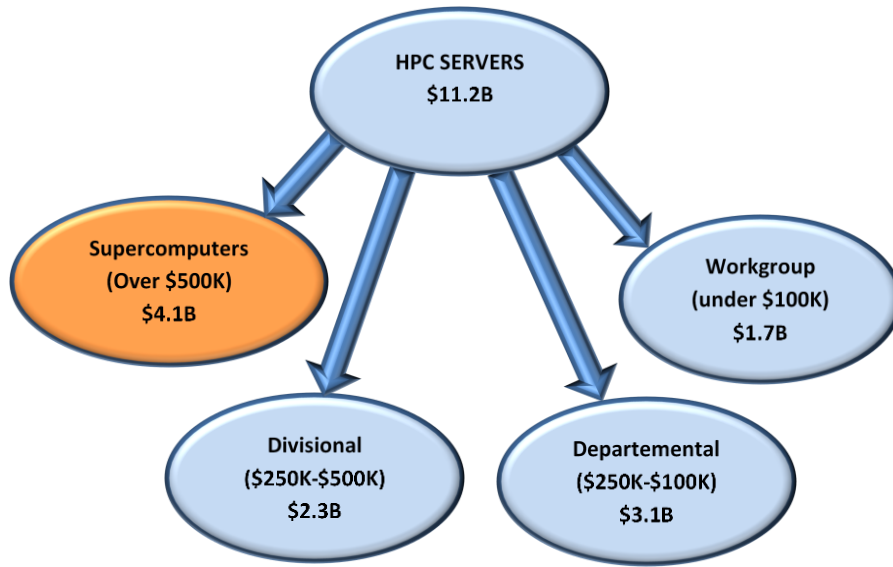
Figure 11 - Worldwide HPC server market

Table 3 depicts the 1H-2017 HPC Market by segments ($K), Table 4 depicts the 1H-2017 HPC Market by Shares ($K), Table 5 depicts the 1H-2017 HPC Market by Regions ($K), Table 6 depicts the 1H-2017 HPC Market by Processor Type, Table 7 depicts the 1H-2017 HPC Market by Coprocessor Type, Table 8 depicts the 1H-2017 HPC Market by Vertical ($ Millions), Table 9 presents HPC Market Forecasts ($ Millions), while Table 10 presents forecasts for the Broader HPC Market ($ Millions).

HPC Server Revenues					
Competitive Segment	Q117	Q217	1H2017		1H17/ 1H16
Supercomputer	948,900	1,062,626	2,011,526		17.3%
Divisional	477,178	518,868	996,046		-14.1%
Departmental	681,172	730,732	1,411,904		-7.3%
Workgroup	329,933	363,125	693,058		-22.9%
Total	2,437,183	2,675,351	5,112,534		-3.5%

Table 3 - 1H-2017 HPC Market by segments ($K)

Mftr	Q1-2017	Q2-2017	Q2 Market Share
HPE/HP	896,027	984,974	36.8%
Dell	488,628	547,582	20.5%
Lenovo	190,521	209,995	7.8%
IBM	97,125	131,682	4.9%
Sugon (Dawning)	71,796	76,369	2.9%
Cray	21,100	47,000	1.8%
NEC	38,572	32,182	1.2%
Fujitsu	42,692	25,801	1.0%
Bull Atos	26,938	19,862	0.7%
Other	563,784	599,905	22.4%
Grand Total	**2,437,183**	**2,675,351**	**100.0%**

Table 4 - 1H-2017 HPC Market by Shares ($K)

Data	2017	
	Q1-2017	Q2-2017
North America	1,163,472	1,298,153
EMEA	648,976	730,000
Asia/Pacific	449,927	484,618
Japan	146,194	131,404
Rest-of-World	28,615	31,175
Total	**2,437,183**	**2,675,351**

Table 5 - 1H-2017 HPC Market by Regions ($K)

WW Processor Package Volume						
	2016				2017	
CPU Type	Q116	Q216	Q316	Q416	Q117	Q217
RISC	21,852	22,725	21,765	19,789	16,264	18,826
x86-64	783,921	827,036	864,815	895,454	799,794	805,708
Custom		40,960				
Total	**805,773**	**890,721**	**886,580**	**915,243**	**816,058**	**824,534**

Table 6 - 1H-2017 HPC Market by Processor Type

Data	Q117	Q217
Sum of Nvidia Board Count	25,455	33,553
Sum of Intel Phi Board Count	4,086	4,204
Sum of Other Acce/CoProc Board Count	1,324	1,338
Total	30,865	39,095

Table 7 - 1H-2017 HPC Market by Coprocessor Type

2016 Market Results	
Bio-Sciences	1,049
CAE	1,251
Chemical Engineering	183
DCC & Distribution	692
Economics/Financial	624
EDA / IT / ISV	823
Geosciences	844
Mechanical Design	57
Defense	1,125
Government Lab	2,059
University/Academic	1,934
Weather	490
Other	70
Total Revenue	11,200

Source: Hyperion 2017

Table 8 - 1H-2017 HPC Market by Vertical ($ Millions)

	2016	2021
Supercomputer	4,091	5,356
Divisional	2,273	2,902
Departmental	3,147	4,274
Workgroup	1,689	2,287
Total	11,200	14,819

Source: Hyperion 2017

Table 9 - HPC Market Forecasts ($ Millions)

	2016	2021	CAGR 16-21
Server	11,200	14,819	5.8%
Storage	4,316	6,269	7.8%
Middleware	1,277	1,786	6.9%
Applications	3,739	5,071	6.3%
Service	1,907	2,309	3.9%
Total Revenue	22,439	30,253	6.2%
Source: Hyperion 2017			

Table 10 - Forecast: The Broader HPC Market ($ Millions)

7.2.2 The HPC/HPDA domain

The Hyperion consulting group has led a study on the HPDA domain with the following objectives:

- Consult experts to develop a taxonomy matrix that matches advanced analytics (HPDA) application problem types with user's preferred algorithms.
- Identify the hardware-software requirements of the applications and the attributes of the algorithms that generate those requirements.
- Present findings in a report designed to be used as a tool for HPDA users (especially non-HPC specialists) from the broad spectrum of application domains

HPDA Key Findings

1. The HPDA Market is still formative
- HPDA was born decades ago in the HPC market
- Only in the past 5-6 years have large commercial firms turned to HPC for advanced analytics

Figure 12 - HPDA market adoption

2. There are almost as many HPDA context as HPDA Users

- Lack of HPDA standards forces users to operate on their own.
- This leads to substantial experimentation and innovation
- But it also means that HPDA users must often fend of themselves
- We hope the present study will provide useful guidance based on the experiences of other users.

3. Not all HPDA algorithms are equal

- Smart algorithms need less data/fewer iterations to home in on useful solutions.
- But applying more brute force computing reduces the need for smart algorithms.
- Talented algorithms developers are in short supply and command high salaries.

4. The Matrix tables

Here below are some matrices produced by this study by vertical sector:

MATRIX: SHORT VERSION -- Algorithms / Data Ingestion Requirements

Vertical	Data Fusion	Data Reduction	Data Integration
BIO-SCIENCES	11%	11%	78%
CAE: PRODUCT DESIGN			100%
CHEMICAL ENGINEERING			
COMMERCIAL ANALYTICS	54%	15%	31%
DEFENSE			100%
ECONOMICS/ FINANCIAL	13%	13%	75%
GOVERNMENT LAB	18%	41%	41%
MECHANICAL DESIGN			100%
WEATHER/ CLIMATE	100%		

Source: IDC 2016

Table 11 - ALGORITHMS /Analytics/Data Ingestion

SHORT VERSION: Algorithms/Analytics Requirements: Machine Learning

Application Area	Unsupervised	Semi-Supervised	Supervised	Reinforcement Learning	Pattern Recognition
BIO-SCIENCES		10%	10%	10%	70%
CAE: PRODUCT DESIGN		67%	33%		
CHEMICAL ENGINEERING					
COMMERCIAL ANALYTICS	33%			33%	33%
DEFENSE					
ECONOMICS/ FINANCIAL	30%	10%			60%
GOVERNMENT LAB	30%	22%	26%	4%	17%
MECHANICAL DESIGN					
WEATHER/ CLIMATE					

Source: IDC 2016

Table 12 - ALGORITHMS /Analytics/Machine Learning

SHORT VERSION: Algorithms/Analytics Requirements: Numeric Optimization

Application Area	Continuous	Discrete	Stochastic
BIO-SCIENCES	40%	40%	20%
CAE: PRODUCT DESIGN	50%	25%	25%
CHEMICAL ENGINEERING	33%	33%	33%
COMMERCIAL ANALYTICS	100%		
DEFENSE	50%	50%	
ECONOMICS/ FINANCIAL	50%	14%	36%
GOVERNMENT LAB	26%	47%	26%
MECHANICAL DESIGN			
WEATHER/ CLIMATE	67%	33%	

Source: IDC 2016

Table 13 - ALGORITHMS /Analytics/Numeric Optimization

SHORT VERSION: Algorithms Requirements: Data Mining and Simulation

Application Area	Data Mining			Simulations	
	Query Processing	Pattern Recognition	Network Analysis	Agent-Based	Time Series Analysis
BIO-SCIENCES	43%	38%	10%		10%
CAE: PRODUCT DESIGN					100%
CHEMICAL ENGINEERING					100%
COMMERCIAL ANALYTICS	16%	26%	26%	16%	16%%
DEFENSE		67%	33%		
ECONOMICS/ FINANCIAL	53%	6%		16%	25%
GOVERNMENT LAB	24%	35%	19%	8%	14%
MECHANICAL DESIGN				50%	50%
WEATHER/ CLIMATE					100%

Source: IDC 2016

Table 14 - ALGORITHMS /Analytics/Data Mining and Simulation

SHORT VERSION: General Hardware Requirements by Application Area

Vertical	Processors	Accelerators	Memory	Interconnect	Storage (Live)	Storage (Archival)	On Premise	Public Cloud
BIO-SCIENCES	25%	14%	25%	7%	19%	2%	3%	5%
CAE: PRODUCT DESIGN	33%	7%	30%	10%	17%			3%
CHEMICAL ENGINEERING	100%							
COMMERCIAL ANALYTICS	25%	8%	19%	14%	14%	3%	11%	6%
DEFENSE	23%	23%	31%	15%	8%			
ECONOMICS/ FINANCIAL	35%	11%	25%	11%	5%	5%	4%	5%
GOV LAB	25%	17%	31%	8%	14%		6%	
MECHANICAL DESIGN	33%		33%		33%			
WEATHER/ CLIMATE	33%		33%	33%				

Source: IDC 2016

Table 15 - Hardware Requirements by Application Area

SHORT VERSION: Architectural System Requirements by Application Area

Vertical	Desktop Only	Cluster	Shared Memory System	Massively Parallel Processing System	Public Cloud	Private Cloud	Other
BIO-SCIENCES		43%	10%	43%			5%
CAE: PRODUCT DESIGN	27%	45%		18%			9%
CHEMICAL ENGINEERING		100%					
COMMERCIAL ANALYTICS		25%	8%	67%			
DEFENSE		20%	60%	20%			
ECONOMICS/ FINANCIAL	5%	35%	10%	10%	15%	25%	
GOVERNMENT LAB		17%	25%	50%		8%	
MECHANICAL DESIGN	100%						
WEATHER/ CLIMATE		100%					

Source: IDC 2016

Table 16 - System Architecture Requirements by Application Area

SHORT VERSION: Hardware Accelerator Requirements by Application Area

Vertical	NVIDIA GPUs	Intel PHI	FPGA	Other	None
BIO-SCIENCES	40%	8%	28%	8%	16%
CAE: PRODUCT DESIGN	59%	29%	6%		6%
CHEMICAL ENGINEERING					100%
COMMERCIAL ANALYTICS	24%	41%	18%	6%	12%
DEFENSE	33%	33%	20%	13%	
ECONOMICS/FINANCIAL	27%	39%	17%	10%	7%
GOVERNMENT LAB	40%	25%	10%	10%	15%
MECHANICAL DESIGN	100%				
WEATHER/CLIMATE	33%	33%	33%		

Source: IDC 2016

Table 17 - Hardware Accelerator Requirements by Application Area

SHORT VERSION: Storage Hardware Requirements by Application Area

| Vertical	Internal system storage	Offline disk storage	Offline tape storage	Near line storage	Active archiving
BIO-SCIENCES	38%			44%	19%
CAE: PRODUCT DESIGN	70%	10%	10%	10%	
CHEMICAL ENGINEERING	100%				
COMMERCIAL ANALYTICS	67%			25%	8%
DEFENSE	80%				20%
ECONOMICS/ FINANCIAL	47%	21%		26%	5%
GOVERNMENT LAB	55%			45%	
MECHANICAL DESIGN	100%				
WEATHER/ CLIMATE	100%				

Source: IDC 2016

Table 18 - Storage Hardware Requirements by Application Area

7.2.3 Artificial Intelligence, Machine Learning, Deep Learning Market Update and Challenges
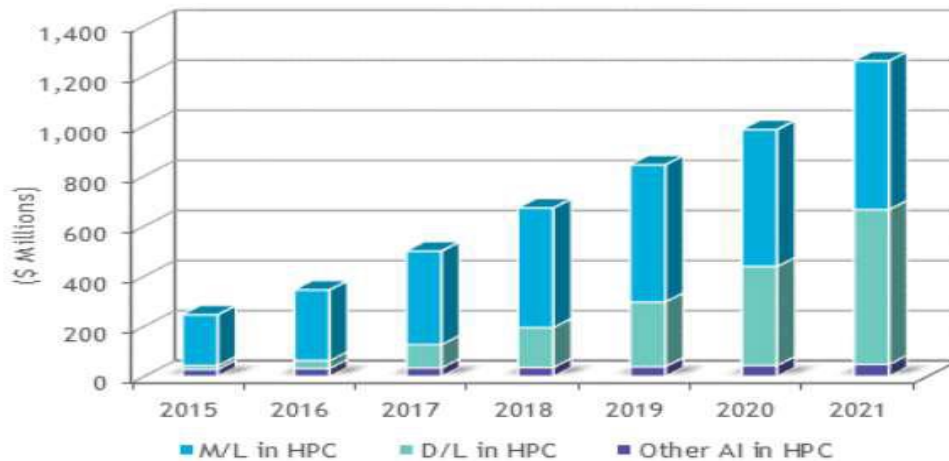
Figure 13 - World Wide ML, DL and AI forecasts

The figure above shows the expected growth of the DL and ML in HPC in the coming years. We can observe that the share of DL becomes more and more important and should be equivalent to the revenues bound to Machine Learning in 2021.

The Deep Learning (DL) challenge

- We can't teach machines to think like humans, because we don't fully understand how humans think.
- DL machines can learn on their own, beyond the instructions humans give them.
- DL machines are capable of learning from each other, I.E., they are capable of "culture".
- Today, DL learning is largely opaque to humans, the basis for DL inferences/decisions is often unclear.
- DL "black boxes" need to be made more transparent.

7.3 Business model

As explained above (Market Value Chain for UC5) we expect some results from the predictive and prescriptive maintenance. The first competitive advantage would be to reduce the downtime or unavailability of the cluster (partial or complete) by the anticipation of failures. We have engagements on the availability rate of the cluster and his components and the predictive maintenance would help us to reduce the downtime and avoid penalties bound to the SLA. The second advantage would be to anticipate on possible failures and replace the potential failing elements in advance thus reducing the

maintenance costs. The cost of the maintenance of an HPC cluster invoiced to the clients is around 5% of the total price every year. On the average lifetime (5 years) of the cluster this adds a revenue of 25% of the initial cost of the HPC cluster. If we can spare a percentage of the costs bound to this maintenance, it will increase our margin in a significant way.

Since we are at the beginning of the project, it's almost impossible to quantify the reduction of costs bound that would bring the prescriptive maintenance to the cluster unavailability rate and to the maintenance costs. This will be evaluated during the project.

8 Conclusion

This report comprises the first output of WP1 and in particular of Task 1.1. It is mainly a market analysis report on the different use cases envisaged for the project starting on the market value chain and concluding with the initial forecast for the business models of the different use cases.

As summary, the following figure shows the overview of the market value chain for all the use cases in order to show and demonstrate the mainstreaming and the cross project developments among the different domains:

	DATA PROVIDERS	TECHNOLOGY PROVIDERS	PLATFORMS PROVIDERS	APLICATION PROVIDERS	END USERS
UC1: E-COMMERCE	PERTIMM IMT SETUR	PERTIMM IMT PRESS'INNOV SETUR KOCSISTEM Lille University	ATOS	PERTIMM	E-COMMERCE AND CUSTOMERS
UC2: CALL CENTER	4C ERICSSON	4C ERICSSON KU LEUVEN TURK TELEKOM TURKGEN BEIA	ERICSSON	4C TURK TELEKOM	BUSINESS EXPERTS AND SERVICE AGENTS
UC3: RECOMMENDATION FOR HHRR	HI-IBERIA BEIA IMT	HI-IBERIA BEIA	HI-IBERIA BEIA	HI-IBERIA BEIA	HHRR DEPARTMENTS
UC4: BAREM	LORIA SOFTEAM	LORIA LILLE UNIVERSITY IMT (PERFOMETRIC)	ATOS	SOFTEAM	CITIZENS
UC5: HPC	ATOS	ATOS	ATOS	ATOS	

Figure 14 - Market value chain summary for the PAPUD use cases

Through this document, the vision of the current Market Analysis for the different envisaged use cases for PAPUD project has been provided. This vision will serve to identify the preliminary gaps and strengths of the market in order to construct a common and powerful platform able to meet the entire value chain of the different processes. For the construction of the platform, it is important to take into consideration that different domains (identified through the different use cases) will have different needs and motivations although the solution will be the same, as same technologies will be implemented to solve the different problems. In consequence, the platform will be used to provide added value to the data and user interaction in different domains.

In addition to this, an initial business model is defined for the different use cases from the perspective of the partners involved and the expected exploitation of the results at this project stage. In general, there is a common feeling on exploiting PAPUD results within the company's current business as the different use cases are aligned with the partners' organizations domains. In parallel with the project evolution, deeper analysis and more concrete plans will be elaborated in order to obtain profit from the project results in their launch to the market.