Building the Legal Knowledge Graph for Smart Compliance Services in Multilingual Europe

D1.1 Functional Requirements Analysis Report

PROJECT ACRONYM	Lynx
PROJECT TITLE	Building the Legal Knowledge Graph for Smart Compliance Services in Multilingual Europe
GRANT AGREEMENT	H2020-780602
FUNDING SCHEME	ICT-14-2017 - Innovation Action (IA)
STARTING DATE (DURATION)	01/12/2017 (36 months)
PROJECT WEBSITE	http://lynx-project.eu
COORDINATOR	Elena Montiel-Ponsoda (UPM)
RESPONSIBLE AUTHORS	Jorge González-Conejero (UAB), Emma Teodoro (UAB) and Pompeu Casanovas (LTU, UAB)
CONTRIBUTORS	
REVIEWERS	Georg Rehm (DFKI), Stefanie Hegele (DFKI) and Julian Moreno (DFKI) Pascual Boil (CC), Francesc Muñoz (CC)
**VERSION	STATUS**	V12	Final
NATURE	Report
DISSEMINATION LEVEL	Public
DOCUMENT DOI	10.5281/zenodo.1256836
DATE	31/05/2018 (M6)

VERSION	MODIFICATION(S)	DATE	AUTHOR(S)
01	Table of Contents	09/05/2018	Jorge Gonzalez-Conejero (UAB)
02	Interviews and Focus Group	24/05/2018	Emma Teodoro (UAB)
03	Methodology and conclusions	29/05/2018	Pompeu Casanovas (UAB)
04	Recap	30/05/2018	Jorge Gonzalez-Conejero (UAB)

ACRONYMS LIST

DOA: Description of Action
GDPR: General Data Protection Regulation
KAP: Knowledge Acquisition Process
LA: Legal Advisor
LE: Large Enterprise
REAs: Requirements Elicitation Activities
SME: Small and Medium Enterprises

EXECUTIVE SUMMARY

The main purpose of this document is to collect business requirements from industry related to needs and expectations regarding a smart and multilingual regulatory compliance information system. The Knowledge Acquisition Process (KAP) includes surveys, interviews and focus groups.

We defined the targeted end-users and stakeholders. We identified five different groups: consultancy firms, legal advisor, small and medium enterprise, large enterprise, and public or private agency.

The survey consists of 24 questions, containing different paths based on the profiles of the respondents. The total number of organizations contacted was 120, obtaining 15 responses from different profiles. On the qualitative side, a total number of 6 semi-structured interviews were conducted in different European countries to public agencies; large enterprises; small and medium enterprises and legal advisors. In addition, one focus group was conducted in a law firm organisation.

The document analyses the results to identify the business and legal requirements related to the handling of multilingual digital regulatory compliance documents. Specifically, the KAP focuses on the stakeholders' strategies, their needs and challenging tasks in their daily workload and, finally, on the solutions and services that could be implemented. The conclusions summarise the findings, and proposes a set of functional and general requirements.

It is worth noting that Ethical guidelines for the processing of data in the context of the knowledge acquisition according to the recent EU GDPR are included in Annex VII.

TABLE OF CONTENTS

TABLE OF FIGURES

1 INTRODUCTION

This Deliverable is the main output of Task 1.1 "Industry requirements elicitation". The main target is to collect business requirements from industry related to needs and wishes for a smart and multilingual regulatory compliance information system. The Knowledge Acquisition Process (KAP) performed—following the Value Proposition Canvas[1]—consists of quantitative (survey) and qualitative (interviews and focus groups) techniques. These techniques will be applied to relevant end users and stakeholders outside the project use cases. The profiles of these relevant end-users and stakeholders are listed in Section 1.2.

The KAP are specifically devised to provide the Lynx consortium with the utmost business requirements in three different subjects when dealing with digital regulatory compliance documents:

- **Strategy** for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents. It includes workflows and/or strategies for the analysis, process and management of digital regulatory compliant documents.
- **Pains** when dealing with digital regulatory compliant documents. This subject is focused on collecting anything that could annoy customers/entities before, during and after dealing with digital regulatory compliant documents.
- **Gains** when dealing with digital regulatory compliance documents. This subject is focused on collecting outcomes and benefits from customers/entities when dealing with the analysis, processing and managing of regulatory compliance documents.

Moreover, Section 2 addresses the survey design, the electronic participant consent, the questionnaire and the results obtained with this knowledge acquisition technique. On the other hand, Section 3 introduces the methodology, informed consent and results obtained from the interviews and focus groups. Finally, Section 4 points out the analysis as a conclusion of the KAP.

In addition, Annex VII reviews the ethical guidelines considered during the design and performing of the KAP.

1.1 PURPOSE OF THIS DOCUMENT

The aim of this Deliverable is to complement Task 4.1 "Pilot Use Cases and Requirements Study". Task 4.1 is aimed at collecting all technical requirements from the use cases within the project. Thus, Task 1.1 main goal is to complement Task 4.1 activities with business requirements provided by relevant end-users and stakeholders (SMEs, Legal advisors, Public and private agencies, among others) from sources outside the project. These business requirements will be focused on end-users and stakeholders needs and wishes regarding a smart and multilingual regulatory compliance information system such as Lynx.

According to Lynx's Description of Action (DOA), Task 1.1 should contain meetings and open discussions to identify requirements and defining priorities for each one of them. However, a more complex strategy has been designed to improve the main outcomes of Task 1.1. This strategy includes a survey, interviews, and a focus group. This is a prospective task to be performed before carrying out the modelling requirements. This Document is not based on polls, but focused on a distributed consultation to stakeholders and potential end-users, respecting the boundaries and guidelines of recent European General Data Protection Regulation (GDPR) requirements. Data is not collected to reflect any ecological validity (a larger amount of data would be needed to make such a claim), but to build up a useful scaffold to complement and support Lynx functional business requirements. The following sections depict the design and conduction of this strategy.

---

[1] Value Proposition Canvas was introduced by Alex Osterwalder: https://strategyzer.com/canvas/value-proposition-canvas

1.2   TARGETED END USERS AND STAKEHOLDERS

The end-users and stakeholders targeted for the knowledge acquisition process (surveys, interviews and focus groups) are described in Table 1. In this table the description and requirements are listed. The list of requirements is not exhaustive; any institution with relevant knowledge or know-how for the Lynx project is suitable to participate in the REAs.

End user / stakeholder	Description	Requirements
Consultancy firms	Enterprise that provides advice to another entity	Domains: • Big data • Legal • Semantics • Internationalization
Legal Advisor	Law firm or lawyer	Domains: • Legal • Experience with the regulatory compliance scenario
SMEs	• Less than 250 staff headcounts • Less or equal of 50m euros turnover • Or less or equal of 43 m euros balance sheet total	Domains: • Enterprise that develops software related to one of the following topics: big data, semantics, natural language processing. This list is not exhaustive. • Internationalized enterprise • Enterprise in process of internationalization
LEs	• More than 250 staff headcounts • More of 50m euros turnover • Or more of 43 m euros balance sheet total	Domains: • Enterprise that develops software related to one of the following topics: big data, semantics, natural language processing. This list is not exhaustive. • Internationalized enterprise • Enterprise in process of internationalization
Public or private agencies	Public or private agencies in the internationalization domain and professionally involved	Domains: • Public or private agency that helps companies in the internationalization process.

Table 1. Targeted end-users and stakeholders.

2 SURVEY

This section addresses the quantitative REAs, the survey, and it is composed of four main stages: Section 2.1 describes the design process carried out during the survey design; Section 2.2 lists the questionnaire; and Section 2.3 introduces the results gathered from the survey and the contributions achieved.

2.1 SURVEY DESIGN

The Lynx survey design process relies on two main pillars: (i) the identification of relevant end-users and stakeholders and the requirements that make them suitable for the Lynx scenario (Table 1); and (ii) the Value Proposition Canvas for the design of the questionnaire.

2.1.1 Value Proposition Canvas

The Value Proposition Canvas helps to design products and services that end-users and stakeholders really want because it allows to focus on what matters most to them. Jobs to be done by end-users and stakeholders is one of the main inputs since jobs describe the things that end-users and stakeholders are trying to get done in their work or in their life. A job could be the tasks they are trying to perform and complete, the problems they are trying to solve, or the needs they are trying to satisfy.

What are the stepping-stones? What are the contexts? How do the activities change depending on these contexts? What functional problems are end-users and stakeholders trying to solve? These are some of the questions involved in the Value Proposition Canvas. As a result, Figure 1 depicts the survey design scheme developed for the Lynx survey.
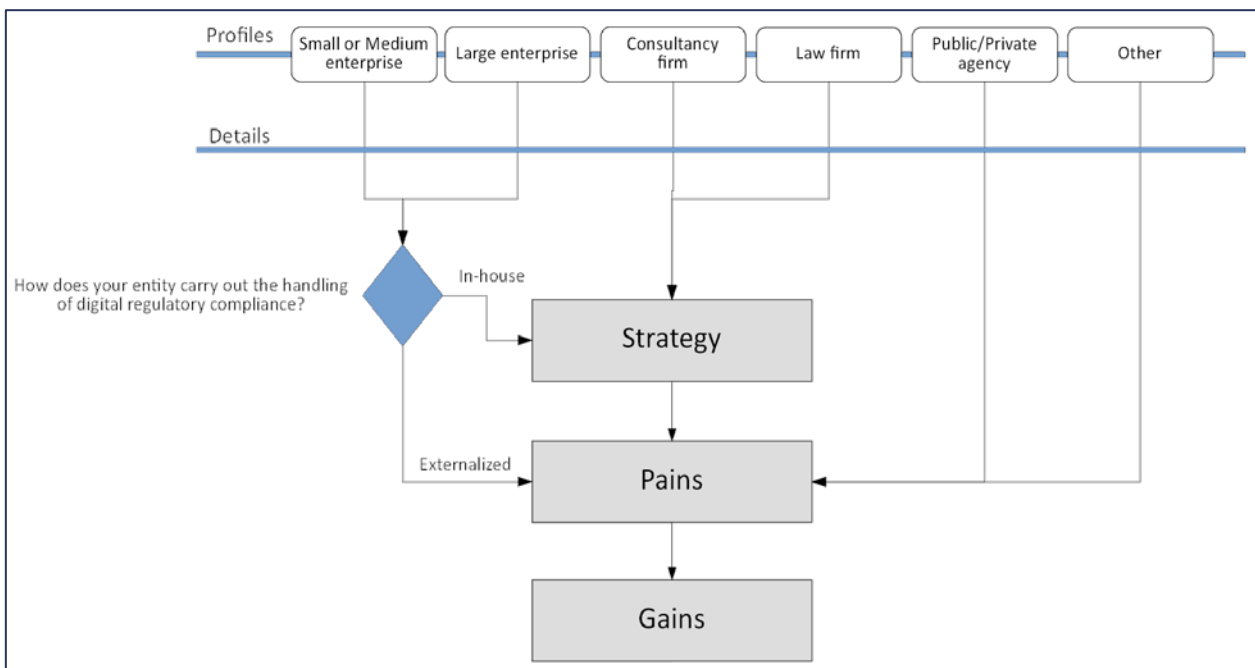
Figure 1. Survey design scheme.

2.2 QUESTIONNAIRE

The final questionnaire obtained from the survey design process depicted in Figure 1 is listed in Table 2. It consists of the Electronic Consent (Annex I); the organization's profile; the strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents; pains and gains.

ID	Question	Possible answers
	Electronic consent	

1	Electronic consent	Mark only one: ○ Agree ○ Disagree
	Organization's profile	
2	Your organization's profile (If your organization's main activity is that of providing advice or counselling to other entities in the field of "regulatory compliance", please select "Legal Advisor" or "Consultancy agency", regardless of the size/legal nature of your organization.)	Mark only one: ○ Small or Medium Enterprise (<250 staff headcount; ≤ € 50m turnover or ≤ € 43 m Balance ○ sheet total) ○ Large Enterprise (LE) ○ Legal Advisor (Law firm or Lawyer) ○ Consultancy agency (Enterprise that provides advice to another entity) ○ Public/Private Agency (BAR Association, Chambers of Commerce, European-National-Regional Trade Organization) ○ Other (Open text box)
3	Country of your entity	Mark only one: ○ List of the European Economic Association ○ Other (Open text box)
4	Economic sector	Mark all that apply: ○ Aeronautic industries ○ Automotive industry ○ Biotechnology ○ Chemicals ○ Construction ○ Cosmetics ○ Defence industries ○ Digital economy ○ Electrical and Electronic Engineering Industries ○ Food and drink industry ○ Gambling ○ Healthcare Industries ○ Maritime industries ○ Mechanical Engineering ○ Medical devices ○ Postal services ○ Pressure and Gas ○ Professional services ○ Raw materials, metals, minerals and forest-based industries ○ Social economy ○ Space ○ Textiles, Fashion and Creative Industries ○ Tourism ○ Toys ○ Other (Open text box)

Enterprise details

5	How does your entity carry out the search, analysis, processing and handling of digital regulatory compliance documents?	Mark only one: o This process is externalized to another entity (Consultancy agency or Legal Advisor) o This process is carried out in-house

Public/Private Agency details

6	Type of legal entity	Mark only one: o Public o Private
7	At What level does your entity work?	Mark all that apply: o European o National o Regional o Other (Open text box)

Strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents

8	Which strategy/workflow for the search, analysis and processing of relevant digital regulatory compliance documents do you use?	Open text box
9	If applicable, Does this strategy/workflow include any specific IT tool? Which are the features of this IT tool? (For instance, an IT tool named X from company Y that provides document analysis, keyword highlighting, topic classification, etc.)	Open text box
10	Which tool for the storing and sharing of relevant digital regulatory compliance documents among your team do you use? (For instance; a cloud service; a shared folder in your computer, a database; etc.)	Open text box
11	How do you monitor changes in regulatory compliance?	Open text box
12	When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, How do you inform about these changes? (For instance, I send a report with these changes every week/month/... by e-mail, I provide a summary of these changes every week/month/..., etc.)	Open text box
13	How do you provide your customers/entity's responsible with the relevant regulatory	Open text box

	compliance in their scenario? (For instance, I send the whole relevant documents; I send a summary of selected relevant documents; I manually/automatically define a set of simplified rules to follow; etc.)	
14	If applicable, does any of the specific IT tools that you mentioned include semantic capabilities? (Semantic capabilities seeks to improve results accuracy by understanding the intent and contextual meaning of terms as they appear. It considers context, location, intent, variation of words, synonyms, concept matching and natural language, among others)	Open text box
Pains		
15	How do your customers/entity define "too costly"? (For instance, our main concern is that a certain task: requires too much time and/or costs a lot of money and/or requires a substantial effort to be accomplished.)	Mark all that apply: o Takes a lot of time o Costs too much money o Requires substantial efforts o Other (Open text box)
16	Which features are your customers/entity missing?	Mark all that apply: o Translation of digital regulatory compliance documents o Summarization of digital regulatory compliance documents o Recommendation of digital regulatory compliance documents based on their content o Alert of changes in digital regulatory compliance documents o Topic classification of digital regulatory compliance documents o Linking of digital regulatory compliance documents based on their content o Smart search among relevant digital regulatory compliance documents o Smart referencing of case law o Other (Open text box)
17	Are there performance issues or malfunctions your customers/entity cite? (For instance, the collecting of regulatory compliance documents takes many time/resources; the collected documents are not related enough to their scenario, the access to the collected documents is slow, etc.)	Open text box
18	What are the main general challenges your customers/entity encounter? (For instance,	Open text box

	to analyse a huge amount of documents in a short period of time; to increase the accuracy of the documents considered as relevant; to find out documents related to already relevant regulatory compliance documents; etc.)	
19	Do your customers/entity face specific difficulties getting certain tasks done? (For instance, to access and share relevant documents; to monitor changes in regulatory compliance documents; to determine time references within documents; etc.)	Open text box
Gains		
20	Which savings would make your customers /entity satisfied?	Mark all that apply: o Time o Money o Effort o Other (Open text box)
21	Which specific features do your customers /entity would enjoy?	Mark all that apply: o Translation of digital regulatory compliance documents o Summarization of digital regulatory compliance documents o Recommendation of digital regulatory compliance documents based on their content o Alert of changes in digital regulatory compliance documents o Topic classification of digital regulatory compliance documents o Linking of digital regulatory compliance documents based on their content o Smart search among the relevant regulatory compliance documents o Smart referencing of case law o Other (Open text box)
22	What would make your customers'/entity's day to day workload easier?	Mark all that apply: o More services (For instance, translation of documents; topic classification, summary of relevant documents) o Lower costs of money when dealing with regulatory compliance documents o Lower efforts when dealing with regulatory compliance documents o Other (Open text box)

23	If applicable, Which services would make your customers'/entity's day to day workload easier? (Please, specify)	Open text box
24	What level of performance do your customers/entity expect? (For instance, a 100% accuracy when setting relevant documents in their scenario; a high-speed access to relevant documents; a high-speed update process when a change in regulatory compliance occurs; etc.)	Open text box

Table 2. Questionnaire for the survey.

2.3   RESULTS

During the task 1.1 activities, 120 e-mails were sent out with invitations to answer the questionnaire. As a result, 15 of the contacted organizations answered the survey. The distribution by country and organization profile is listed in Table 3.

Countries / Profiles	Austria	Germany	Italy	Netherlands	Norway	Latvia	Spain	Total
Consultancy firm	--	--	--	--	--	--	--	**0**
Legal Advisor			1				1	**2**
SME	1	3				1		**5**
LE			2		1			**3**
Public/Private Agency							3	**3**
Other		1		1				**2**
Total	**1**	**4**	**3**	**1**	**1**	**1**	**4**	**15**

Table 3. Participation in the survey by country and by profile.

Regarding the economic sector of the respondents, Figure 2 shows that "Digital Economy" is the most common. Then, "Professional services" is the second option. Moreover, Question 5 shows that from 8 SMEs and LEs, 5 handle digital regulatory compliance documents in-house and 3 have externalized the process. Thus, in the questionnaire when a SME or LE have externalized the process, the part of the strategy is skipped.
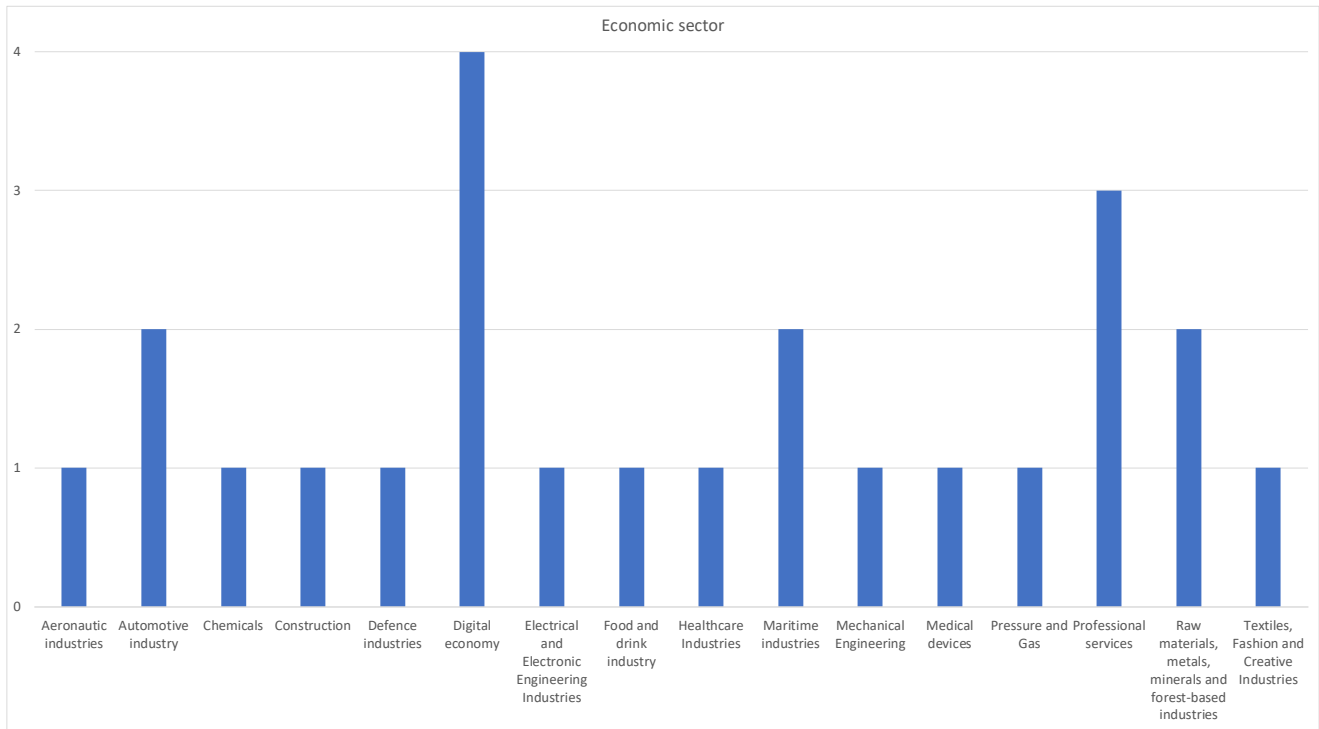
Figure 2. Economic sector of the respondents: Question #3 of the survey.

This section follows with the answers for the "Strategy" stage (Section 2.3.1), "Pains" stage (Section 2.3.2) and "Gains" (Section 2.3.3).

2.3.1 Strategy

Strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents	
Q8: Which strategy/workflow for the search, analysis and processing of relevant digital regulatory compliance documents do you use?	
SME	In individual case according to work-needs
	Exchange of documents between subsidiaries in Germany, Australia, India
	Internal lawyers follow up the latest changes in digital regulatory and apply changes as required
LE	Partly by expert's knowledge, partly by defining (public available) standards and requirement documents
	Internal approval after external partner suggestion
LA	Difficult to answer (talk about strategy is not a clear thing)
	We pay attention to legislative innovations through official bulletins for domestic law and EurLex for European Law
O	Workflow for document digitization, text recognition (OCR), indexing and online publication

Table 4. Answers related to Q8.

Table 4 shows the results related to the strategy for searching, analysis and processing relevant digital regulatory compliance documents by targeted end-users and stakeholders. In general terms the results reveal that different strategies are used depending on the typology of end-users and stakeholders.

For instance, SME identified the need of adapting the strategy depending on working needs related to a particular case. The exchange of documents with subsidiaries in other countries or the use of internal lawyers to perform the identification and implementation of latest changes are also highlight as two different proceedings to deal with digital regulatory compliance. However, LE based the strategy on experts consulting. Besides, LA mention the fact of paying attention to legislative innovations through consulting national and European legal databases, such as Official Journals or Gazettes in the first case or EurLex in the second.

Strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents	
Q9: If applicable, Does this strategy/workflow include any specific IT tool? Which are the features of this IT tool?	
SME	Lexis Nexis; provides topic classification
	No, mainly via mail
LE	Yes; inclusion of metadata in an in-house production tool
	No IT tool
LA	Expert System Cogito Discover
	Online database
O	Kitodo (http://www.kitodo.org/)

Table 5. Answers related to Q9.

The use of IT tools in the context of the strategy of searching, analysis, processing, monitoring and handling digital regulatory compliance documents appear to be commonly used by different end-users and stakeholders according to the responses received. Details on the specific IT tools used by end-users are provided in Table 5.

Strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents	
Q10: Which tool for the storing and sharing of relevant digital regulatory compliance documents among your team do you use?	
SME	A shared folder
	Shared folder (e.g. dropbox)
	Shared folders, data bases
LE	Database
	Shared folder
LA	NetDocuments (Document Management System on the cloud)
	Legal database
O	Fedora (http://www.fedora-commons.org/)

Table 6. Answers related to Q10.

Question 10 of the online survey offers a very common response on which tool is used for storing and sharing relevant digital regulatory compliance documents. In this regard, shared folders are borne out by end-users and stakeholders as the most common tool for the purposes mentioned above. However, LE and LA identified the use of databases and particularly LA mention the use of document management systems on the cloud.

Strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents	
Q11: How do you monitor changes in regulatory compliance?	
SME	IT Tool (LexisNexis)
	Manually, via mail
	Monitor the governmental
LE	Delegation to expert (teams) who establish shared overviews
	Our internal legal team
LA	We have a specialized KM team, but we normally use BOE alerts and specific tools like Lexis Nexis, VLEX…that provide us legal contents and it has the possibility to define some alerts
	Through official buletins

Table 7. Answers related to Q11.

The monitoring of changes related to regulatory compliance is performed differently depending on the typology of the end-users and stakeholders. Therefore, SME pointed out both the use of IT Tools such as LexisNexis and the proceeding used to monitor changes. In this regard, they perform this task on manually basis by their own staff. LE put in practice different proceedings to perform this task: some of them prefer the externalization and delegation of this task in expert teams while others use internal legal teams. LA also reported the use of Official Journals Alerts or specific tools such as LexisNexis as well as the use of specialized teams that are in charge of providing legal contents and defining some alerts.

Answers related to Q13 show that the most frequent method to provide regulatory compliance used by end-users is a summary of selected relevant documents adding some comments. Nevertheless, LA mentioned also the use of specific automated subscription services, social media and blogs.

Answers provided by end-users to Q14 suggest that no IT tools with semantic capabilities are frequently used in providing regulatory compliance to customers or entities. Only one of the responders, a LA mentioned the use of an IT tool, which is Expert System Cogito.

2.3.2 Pains

The "Pains" stage in the survey is devised to find out the challenges that respondents experience in their daily work. In this sense, Question 15 in the survey asks for how customers/entities define "too costly". Figure 3 collects all the responses in a bar chart. This chart points out that "Costs too much money" is the main concern for customers/entities when dealing with digital regulatory compliance documents. In addition, "Takes a lot of time" is highly considered with equivalent results.
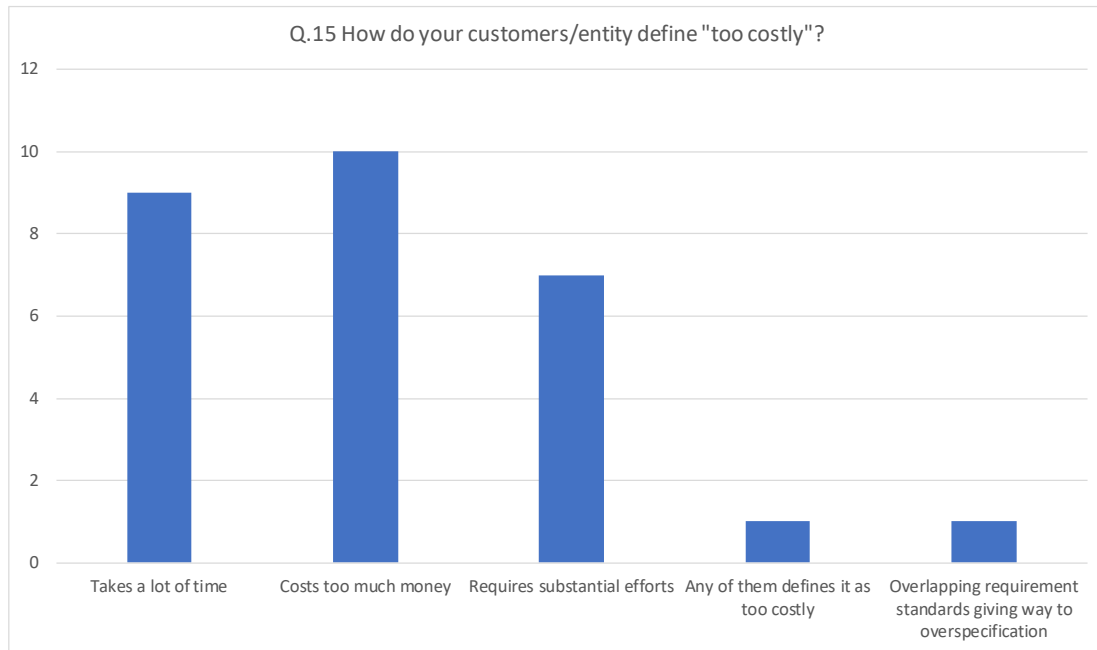
Figure 3. Answers related to question 15.

Moreover, Question 16 in the survey ranks the missing features that respondent's state. In this sense, "Summarization of digital regulatory compliance documents", "Alert on changes in digital regulatory compliance documents" and "Smart search among relevant digital regulatory compliance documents" obtain equivalent results as the most common options. Figure 4 depicts the whole results obtained in this question.
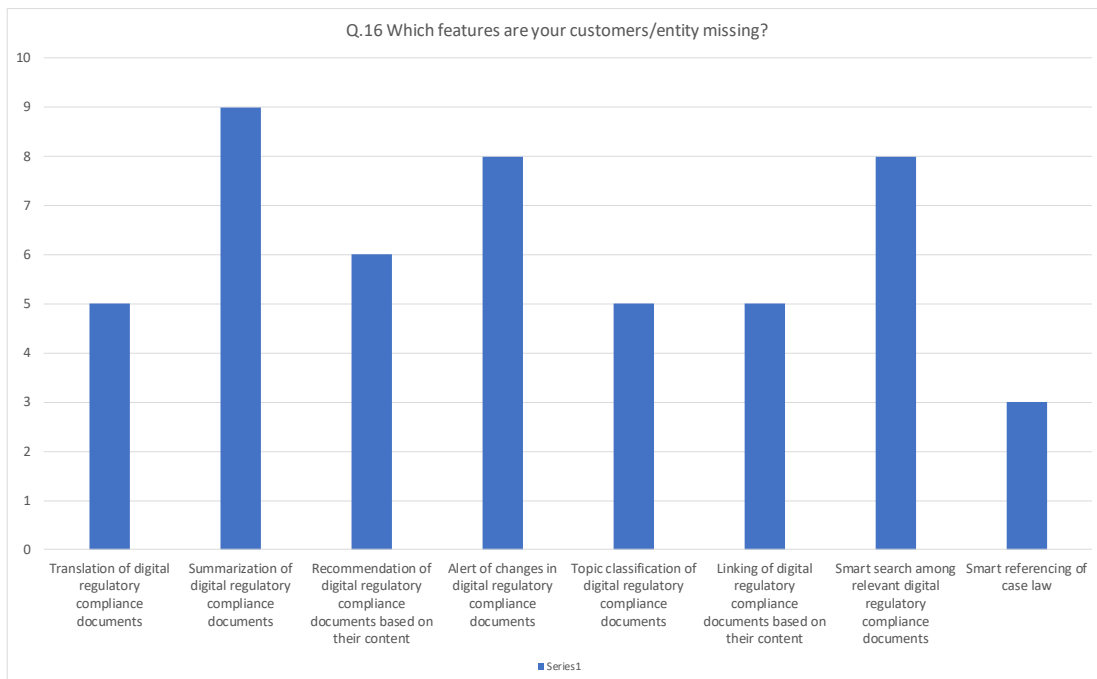
Figure 4. Answers related to question 16.

Table 8 provides an overview of some performance issues or malfunctions highlighted by customers or entities.

Pains
Q17: Are there performance issues or malfunctions your customers/entity cite?

SME	Collected documents are not related enough to their scenario
	No
	Not clear how to introduce regulation
LE	Challenges to control change management (from revision to revision)
	No
	No
LA	I do not know
	Collected documents are too complicate to be clearly understood without a specific legal support
O	The quality of links that are derived between documents and other external knowledge bases is often too unspecific
PPA	Norms and standards ignorance
	Lack of summarized information

Table 8. Answers related to Q17.

Table 9 shows the main general challenges identified by customers and entities with regard to providing regulatory compliance.

Pains
Q18: What are the main general challenges your customers/entity encounter?

SME	To analyse a huge amount of documents in a short period of time
	Not aware of any
	Translation and transfer of texts for state and financial authorities of different countries
	Huge amount of documents, accuracy in application of the regulations
LE	Challenges to control change management (from revision to revision)
	To find out documents related to already relevant regulatory compliance documents
	To update a lot of documents in several languages at the same time
	Increase accuracy of document
LA	To find out documents related to already relevant regulatory compliance documents
	To increase the accuracy of the documents considered as relevant
O	The amount of documents is very high, the quality of those documents is bad (due to their historical nature), simple full-text search does not provide sufficient relevancy for user satisfaction
	Relevant accuracy, accessibility, in comparison from one country to another
PPA	The lack of easy to understand documents
	To find out valuable and understanding information

Table 9. Answers related to Q18.

Table 10 exposes the main difficulties identified by end-users relating to certain tasks such as accessing relevant documents or monitoring changes in regulatory compliance documents among others.

Pains	
Q19: Do your customers/entity face specific difficulties getting certain tasks done?	
SME	To monitor changes in regulatory compliance documents
	Optimizing costs and bureaucratic efforts for daily procedures (annual balance, change of share ownership)
LE	Monitor changes in regulatory compliance
	Translation and publication of contents
	Increase accuracy of document
LA	To monitor changes in regulatory compliance documents
O	Due to bad quality and historical nature of the documents, the quality of currently available NLP methods for analysing and deriving semantic structures in these documents is insufficient
	Access to relevant documents, storage of design data, impact analysis of design changes
PPA	To answer to the relevant documents

Table 10. Answers related to Q19.

2.3.3 Gains

Question 20 of the survey asks for the most preferred saving among time, money and effort. Figure 5 depicts the results obtained. In this context, all of the possible responses obtained equivalent results. In addition, question 21 asks for the preferred features that customers/entities would enjoy. Figure 6 shows that "Alert of changes in digital regulatory compliance documents" is the most common response.
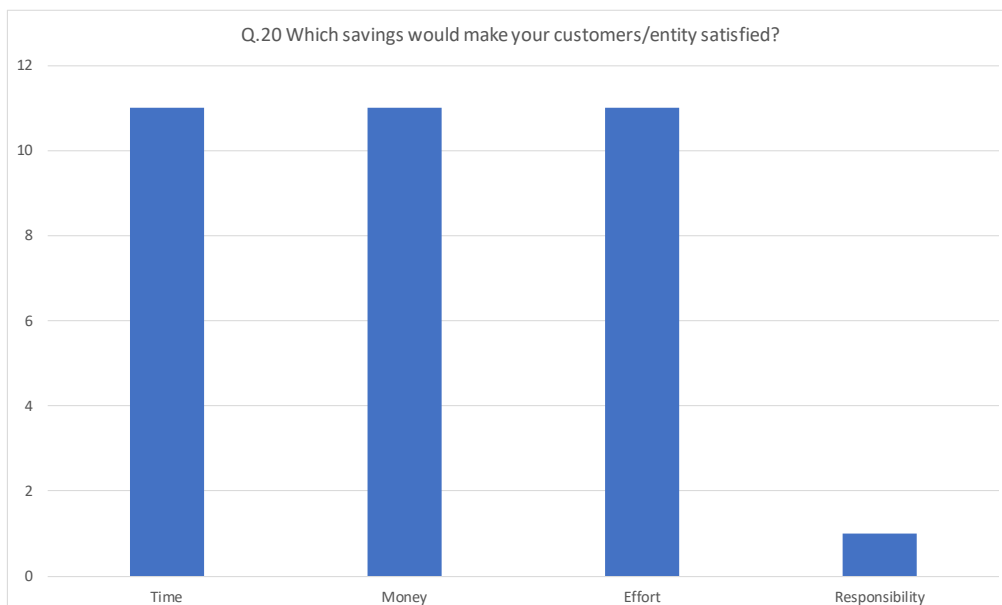
Figure 5. Answers related to question 20.
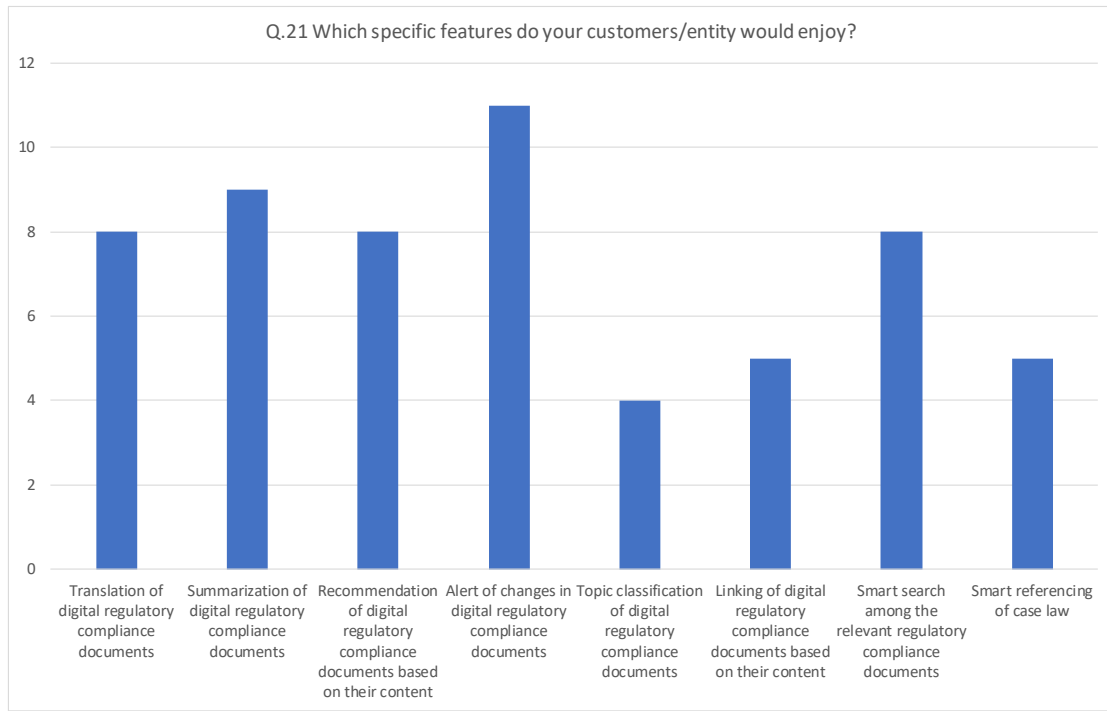
Figure 6. Answers related to question 21.

Figure 7 contains the most preferred options that could make customers'/entities' today workload easier.
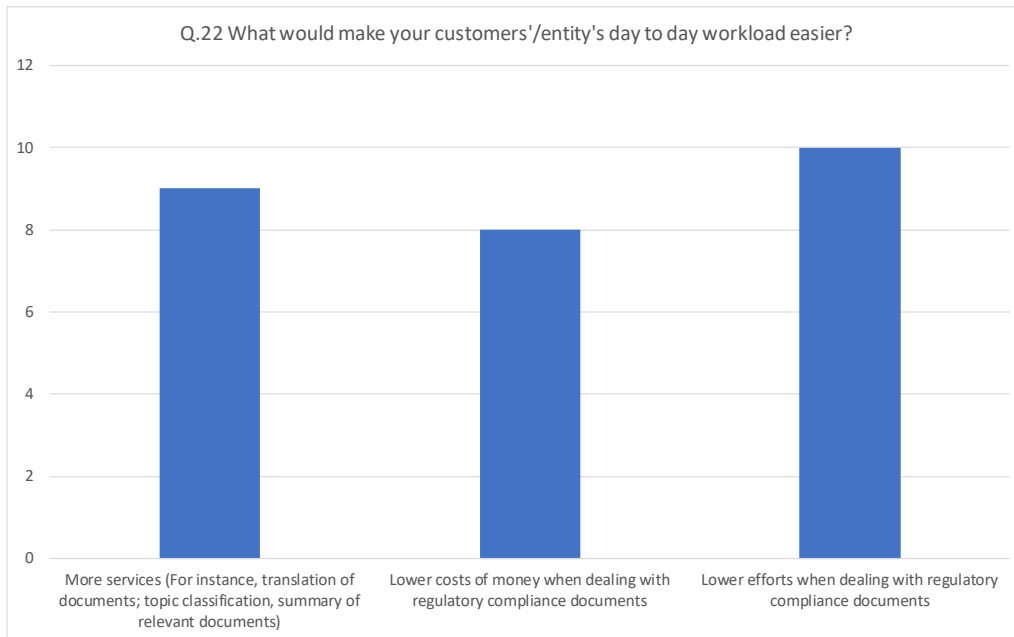
Figure 7. Answers related to Q22.

Table 11 contains suggestions provided by end-users in relation with services that could make their day-to-day workload easier.

Gains	
Q23: Which services would make your customers'/entity's day-to-day workload easier?	
SME	Precise translation of agreements for different countries and languages

LE	An up to date overview of all the applicable regulatory requirements with a link with their documentation to support compliance management (including revision control)
O	Services to perform semantic analysis and linking of content contained in the documents
	Accessibility, storage (archive), automated reference and selection of applicable standards

Table 11. Answers related to Q23.

Finally, it is remarkable the level of coincidence between most of the end-users when they were answered about the level of performance expected in dealing with regulatory compliance. In particular, they highlighted the need of: i) a high-speed update process when a change in regulatory compliance occurs; ii) and, 100% accuracy when setting relevant documents in their particular scenarios.

2.3.4   Analysis of the results from the survey

Regarding the economic sector of the respondents, the results obtained from the survey show that the most common sectors are: "Digital Economy", "Professional Services", "Raw materials", "Metals", "Minerals and forest-based industries", "Maritime industries" and "Automotive Industry". These economic sectors are ordered by importance in the results from the survey.

In addition, Table 12 shows the analysed results from the "Strategy for searching, analysis and processing of relevant digital regulatory compliance documents".

| Results related to the strategy for searching, analysis and processing relevant digital regulatory compliance documents by targeted end-users |
|---|
| In general terms the results reveal that different strategies are used depending on the typology of end-users and stakeholders |
| SME identified the need of adapting the strategy depending on the working needs related to a particular case. The exchange of documents with subsidiaries in other countries or the use of internal lawyers to perform the identification and implementation of latest changes are also highlighted as two different procedures to deal with digital regulatory compliance. |
| LE based the strategy on experts consulting. |
| LA mentioned the fact of paying attention to legislative innovations through consulting European and national legal databases (Official Journals, Gazettes, EurLex). |
| Answers related to which IT tool is used for storing and sharing relevant digital regulatory compliance documents indicates that shared folders are the most common tool used for the purpose mentioned above. However, LE and LA identified the use of databases, and particularly LA mention the use of document management systems on the cloud |
| The monitoring of changes related to regulatory compliance is carried out in a different way depending on the typology of the end-users and stakeholders. Therefore: (i) SME pointed out both the use of IT Tools such as LexisNexis and the procedures used to monitor changes. In this regard, they perform this task on manual basis by means of their own staff; (ii) LE put in practice different procedures to perform this task: some of them prefer externalising and to delegate it to expert teams, while others use internal legal teams. (iii) LA also reported the use of Official Journals Alerts or specific tools such as LexisNexis as well as the use of specialized teams that are in charge of providing legal contents and defining some alerts.
Answers related to Q13 show that the most frequent method to provide regulatory compliance used by end-users is a summary of selected relevant documents adding some comments. Nevertheless, LA mentioned also the use of specific automated subscription services, social media and blogs. |

Answers provided by end-users to Q14 suggest that no IT tools with semantic capabilities are frequently used in providing regulatory compliance to customers or entities. Only one of the responders, a LA, mentioned the use of an IT tool, which is Expert System Cogito.

Table 12. Results obtained from the survey related to the strategy for searching, analysis and processing relevant digital regulatory compliance documents by targeted end-users.

Regarding pains stated by respondents, Table 13 lists the analysis of the obtained answers.

Results related to pains
Too costly" are defined by the majority of the participants in terms of "cost too much money".
In relation with which features are missing, respondents provide the following three answers according to their ranking of priorities: i) summarization of digital regulatory compliance; ii) alert of changes in digital regulatory compliant documents; iii) smart search among relevant digital regulatory compliant documents.
The participants point out some malfunctions or performance issues such as (i) the collected documents are not related enough to their scenario, (ii) the big challenge that implies the task of monitoring control changes derived from regulatory changes (iii) the difficulty to understand collected documents without a specific legal support, (iv) the lack of summarized information or the inaccurate results of links between different relevant documents.
Participants highlighted the following main general challenges: (i) the huge amount of documents that would require analysis in a short period of time, (ii) the need of increasing the accuracy of documents considered as relevant ,and (iii) the accessibility to these relevant documents.
Participants identified monitoring changes in regulatory compliant documents as one of the most challenging task to deal with.

Table 13. Results obtained from the survey related to pains experienced by the respondents in their daily work.

Finally, Table 14 depicts the analysis of the answers collected that are related to the gains that respondents highlight. These gains affect the respondents daily work, contributing to make their life easier.

Results related to gains
There is no difference between time, money and effort in terms of being perceived by participants as savings that they would like to achieve.
About the specific features that they would welcome, responses offer the following ranking of preferences: (i) *alerts* about changes in digital regulatory compliant documents; (ii) *summarization* of digital regulatory compliance documents; (iii) *translation* of digital regulatory compliance documents; (iv) *recommendation* of digital regulatory compliance documents, based on their content and *smart search* among the relevant documents.
Lower efforts when dealing with regulatory compliant documents, more services in terms of topic classification, and summary of relevant documents are perceived by participants as different ways to make easier their day to day workload. To that end, they suggested some services such as: "(i) precise translation of agreements for different countries and languages;(ii) services to perform semantic analysis and linking of content contained within the documents, (iii) an updating overview of all the applicable regulatory requirements with a link to their documentation to support compliance management (including version control) and (iv) accessibility, storage, automated reference and selection of applicable standards".

Finally, there is a remarkable level of coincidence among most end-users in their answers about the level of performance expected when dealing with regulatory compliance. In particular, they highlighted the need of: (i) a high-speed updating process when a change in regulatory compliance occurs; and (ii) 100% accuracy when setting relevant documents in their particular scenarios.

Table 14. Results obtained from the survey related to gains wished by respondents in their daily work.

3 INTERVIEWS AND FOCUS GROUPS

3.1 METHODOLOGY

A "Qualitative Interview" is a method of collecting rich and detailed information about how individuals experience, understand and explain certain events or particular topics [Harvey2011] [Turner2010] [Flinders1997] [Rubin & Rubin2011]. Interviews are "semi-structured" because the interviewer has a list of questions or key points to be covered during the interview, and works through them in a methodical manner. Similar questions are asked to each interviewee, although supplementary questions could be asked as appropriate. In general, questions are worded so that responses are open-ended.

This open-endedness allows the participants to contribute with much detailed information as they desire; it also allows the interviewer to ask probing questions as a means of following-up. In other words, the interviewees could in principle respond how they like. This can make quite difficult for the interviewer to keep the interviewee on focus while interviewing, and then extract similar themes or codes from the interview transcripts. However, semi-structured interviews reduce individual biases within the study, particularly when the interviewing process involves many participants.

However, this perspective about the risks of qualitative research may lead to a reductionist view that we would like to avoid. Qualitative methods have been described at length in Knowledge Acquisition Processes (KAP) for modelling. Enrico Motta edited a special issue on 25 years of KAP in the Semantic Web area at *International Journal of Human-Computer Studies.* Elaborating on Gaines, Gruber and Bradshaw's contributions, he wrote [Motta2013, 132]:

" [...] much of the interesting action concerning knowledge technologies was actually taking place in the semi-secluded gatherings of this small community and that the real interesting issues were not the formal and abstract Knowledge Representation problems, tackled through ''dryerase whiteboard results'' (Gruber, this issue), but the ones *concerning the effective development of symbiotic intelligent systems* (Bradshaw, this issue; Gaines, this issue). *These issues could only be tackled effectively through an interdisciplinary approach, grounded as much into empirical investigations and cognitive science principles, as in formal knowledge representation and computational architectures*."

We could not agree more. A genuine non-eclectic interdisciplinarity orientation is key to tackle Lynx problems on building a legal graph, and to map legal and business requirements.

Hence, we adopted a two-fold strategy: (i) encompassing this empirical approach to properly *eliciting* modelling requirements across several business and legal fields (as a process); (ii) combining quantitative and qualitative methods in the structured fromal line advanced, e.g. by the Unified Modeling Language (UML) perspective (as an outcome). In this sense*, completeness, consistency, adequacy, unambiguity, measurability, pertinence, feasibility, comprehensibility, good structuring, modifiability, and traceability* will be deemed quality factors to define the goals of the Requirement Engineering process [van Lamsweerde, 35 and ff]. "The requirements emerging from the elicitation and evaluation phases of the RE process must be organised in a coherent structure and specified precisely to form the requirements document "(ibid. 174).

Qualitative research can specify and introduce useful nuances to the summary of preliminary survey results reported in Section 2.3.4. The interview and focus group techniques based on further elaboration of the previous questionnaire Section 3.2 leaded to interesting results, allowing end-users to refine some of the answers already obtained. First, revealing some internal organisational processes and strategies of government agencies, small / large companies, and law-firms which had not been detected by the survey. Second, providing illuminating expressions and language that summarise the end-user's conceptual perspective, concerns, and needs on compliance and regulatory problems.

3.2 ANALYSIS OF THE RESULTS FROM INTERVIEWS

Participation by profile: 5 interviews were conducted (1 Small Medium Enterprise from Spain; 1 Large Enterprise from Spain; 2 Public Agencies, one from Spain, one from Finland; 2 Legal Advisor from Spain). Total: 6 interviews.

3.2.1 Public Agencies

One PPA interviewed is a Catalan Internationalization Agency. It holds international premises with technical staff for helping companies in their internationalisation process. In fact, the international offices act as a consultancy agency. The degree of consultancy that they can provide depends directly on the expertise of its staff. If they cannot provide an accurate answer as a result of their lack of expertise, they follow the strategy of providing or recommending a set of external local consultancy firms. They reported that last year more than 1.200 consultancy projects were carried out by the Agency. In this regard, they pointed out that the main challenges that companies face when they decided to internationalise their activity within the European Single Market are:

"International Public contracting issues and workers mobility… rules governing the European Single Market are not followed…and in some European countries is mandatory to include a local company to display your activity… ".

One suggestion provided in the context of this interview related to the missing features identified according to their expertise:

" …Companies that get involved within an internationalisation process will welcome a software or application able to provide interpretable legal information. It is not a matter of language since they act with offices in the target country or with local consultancy firms, but a matter of the expert legal knowledge needed to understand regulatory documents instead…"

The other PPA interviewed is a European regulatory authority on the safe use of chemicals. Its main activity focus on helping companies to comply with chemical legislation for the benefit of human health and the environment, as well as for innovation and competitiveness within the chemical sector. Information related to production, labelling, biocides and import of chemicals in terms of regulatory compliance are the main areas in which they need to provide relevant information. The strategy followed by this public agency focus on developing an IT tool. In this regard:

" We are developing our own tool…available for everyone, which shall comprise all the relevant legislation. If a company wants to introduce a new product into the market…which regulations do I have to comply with?...We are considering developing automated tools…but to give precise, accurate answer is costly and we do not believe that this can be fully automated. In the chemical domain, there is much specific knowledge…"

Accuracy and *high-speed answer*s to queries are identified as the most relevant expectations in terms of evaluating the level of performance of an IT tool.

3.2.2 Large Enterprises

The Large Enterprise interviewed focus its activity within the Automotive Industry (Systems for railways vehicles and Systems for road vehicles). It is an internationalised corporation developing its business mainly in Europe, US, South America, Asia and Africa.

When dealing with the process of digital regulatory compliance the manager of the organisation highlighted as a general rule the need of externalising to another entity (a worldwide leader in testing, inspection and certification consultancy agency) this process. They have scheduled regular meetings as a strategy to discuss doubts concerning regulatory changes that might affect their activity, or to agree by

consensus practical ways of implementing those changes from a technical point of view. Nevertheless, they use to contract local legal firms in USA.

Moreover, in terms of monitoring the regulatory changes that affects regulatory compliance he pointed out that customers are usually the ones who alert them about these regulatory changes. Related to this, *he is asking for a tool allowing companies to identify in advance potential new scenarios as a result of regulatory changes, with the aim to provide a more professional response to their clients*.

3.2.3   Small and Medium Enterprises

We have conducted one interview to SME providing legal data solutions for law firms, corporations, universities, research centres, and governments through an intuitive-friendly interface platform that delivers the most relevant results in a fast manner. They pointed out that they are implementing services within their platform offering: (i) recommendations, (ii) smart referencing among related relevant documents, (iii) alerts, (iv) topic classification, (v) smart search, (vi) and temporal references in documents.

3.2.4   Legal Advisors

Two interviews have been conducted under this profile. One of them corresponds to a Legal Advisor of a Law Firm that is a Large Enterprise in Spain; and the other covers the views of a legal advisor who is a self-employed professional within a Spanish small law firm.

A. Input provided by a legal advisor developing his professional activity within a Large Enterprise

He highlighted the following issues:

1. Regarding the strategy for searching, analysis, processing, monitoring and handling digital regulatory compliance documents:

" *…For searching relevant regulatory compliance documents, he uses different commercial tools…when relevant, he downloads the document to his PC's desktop…for sharing and storing, the company uses a commercial document management system…"*

2. In relation with IT tools, he identified some commercial IT solutions for searching relevant documents provided by Aranzadi, Vlex and Wolters Kluwer. A management system (NetDocuments) is used for sharing documents among different members of a legal team.
3. As the firm is integrated into a Law firm European network, when an action has to take place abroad, they contact a local firm within this European network when an action has to take place abroad, transferring to a local firm the implementation of actions considered adequate to solve a particular need. English is the most common used language.
4. The law firm has a Knowledge Management team for monitoring changes in legal documents. A specific team is in charge for implementing the updating process related to these changes. They carry out this task on a manual basis. The aim is to create a monthly report covering changes on different topics. Nevertheless, if a change is sufficiently relevant or critical, they report it immediately with independence of the monthly report.
5. Concerning the communication process with their clients, he pointed out that they take care of the whole process in such a way that the customer never has to deal with documents; he only receives expert legal advice.
6. Concerning potential specific features that they would like to enjoy within the Lynx Platform, he suggested that the added value would be achieved implementing the following features: (i) *summarization of digital regulatory compliant* documents based on their content; (ii) alert *on changes*; (iii) *topic classification* of digital regulatory compliant documents; (iv) and, *linking* digital regulatory compliant documents on the basis of their content. Nevertheless, he considers that

legal experts are also required to perform a final check, to ensure that information is really relevant to the particular case.

B. Input provided by the legal advisor developing his professional activity within a Spanish small law firm

He highlighted the following issues:

1. First of all he identified two steps related to the assessing process of companies internationalization: i) externalization to another entity (Local Legal Advisor); ii) and, specific consultation on official aids, financing projects, services and benefits are addressed to the Spanish Institute for Foreign Trade (ICEX), Chambers of commerce or similar bodies. In relation with the first step he stated:

 "…It is always better to contract local legal advisors to avoid unnecessary over-bureaucratic proceedings…"

2. He always receives information through local legal advisors when a relevant change in regulatory compliance occurs. In that sense, he pointed out that:

"…I appreciate very much this information sharing strategy because it significantly reduces the volume of data eliminating irrelevant information…"

3. In the context of internationalization processes the most frequent queries are related to *Tax Law*, *Labour Law* (employment permits and work contracts), *required permits* or necessary authorizations, and operating *licenses*. So, key information should be provided on those topics.

4. The updating process when a relevant change occurs is carried out by the administrative department of the law firm, according to the information received by the local advisor, and previous consultation with the internal lawyer involved in a particular case.

5. Once they receive summaries of selected relevant documents by email from their local advisors, the next step consists on preparing a set of rules or protocols to be followed (understood as actions to be taken) after examination.

6. He identified the need of receiving summaries that are more accurate. In this regard, he stressed that:
" …Sometimes summaries provided by external consultants are very general. Each sector has its own specific features. So, content summaries should focus only on key information according to the query about what needs to be answered…"

7. The legal advisor has remarked two main malfunctions: (i) access to relevant documents is deemed too slow for everyday purposes; ii) and, (iii) the need of rigorous data analysis (in the sense that collected documents provided are not related enough to your specific sector).

8. With regards to specific features that he would enjoy within the Lynx platform, he mentioned the following ones: (i) *links to multiple sources*; (ii) *alerts*; iii) *smart search*; iv) *smart referencing across related documents*; (v) topic *classification*; vi) and, particularly, *he stressed that the Lynx Platform services should be customized according to the specific needs and professional profile* of end-users.

9. Finally, a *high-speed access* to relevant document and *100% accuracy* are the two features related to the expectation level of performance stressed during this interview.

3.3 ANALYSIS OF THE RESULTS FROM FOCUS GROUP

One focus group has been conducted in the premises of a Law Firm in Spain. Participants: one Tax Law legal advisor; two Labour Law legal advisors; three ICT lawyers. Coordinator: IDT-UAB

Results are provided according to the following topic classification, and based on the field notes taken by the researches during the development of the focus group:

- Topic 1: How legal advisors are searching for relevant information

- Topic 2: How legal advisors prepare relevant information for their lawyers. Identification of the most challenging task of the process.
- Topic 3: Accuracy of the information provided
- Topic 4: Information provided to the lawyer
- Topic 5: The need of creating a subsidiary in another Member State
- Topic 6: Suggestions provided by the participants related to the Lynx Platform functionalities

TOPIC 1: How legal advisors search relevant information

1. Legal databases such as Aranzadi are too generalist tools for searching.

2. Classification of the information retrieved is carried out taking into account the speciality of the content, not the lawyer expertise.

3. Some difficulties are identified related to this process: i) they deliver an English summary instead of translating the document from a legal expert point of view. Other languages such as Spanish, Catalan and French are also used to provide summaries; ii) however, although these summaries might be useful, they are not detailed and precise enough to be interpreted as an in-depth legal analysis of a particular case.

TOPIC 2: How legal advisors prepare relevant information to their lawyers. Identification of the most difficult tasks of the process.

1. The task of performing the summaries are time-consuming and implies too many costly efforts.

2. A huge amount of information is received, and in order to perform a summary some tasks need to be done: first, all the information received must be read and taken into account; second, a legal analysis is required according to a given client and domain.

3. The summary should fulfil the purpose of providing key information to a lawyer in such a way that once read, the lawyer could be empowered to make a reasonable decision about whether he is interested or not

4. Identifying judgements that involve significant or radical changes in relation to previous legal framework would be useful. Identification of key issues removed by the new legal framework. "It would be great!"

5. Lawyers expertise should concentrate on identifying which is the best legal strategy based on the information provided.

TOPIC 3: Accuracy of the information provided
1. The problem is the content.
2. Providing an accurate classification of documents is relevant. They do not have statistics on how much time a lawyer is consuming to consult a particular document.
3. They are in possession of many information sources, but it turns out that they are too disperse to work them out in an efficient way.

TOPIC 4: Information provided to the lawyer
1. Legal advisors provide a summary. They offer arguments about key issues to make easier for the lawyer to choose one strategy or another, taking into account the client's needs.

2. " Our lawyers need to know that they know everything. *We are like a radar system*. In this regard we should have a lead on the way the market is developing from a technical or legal perspective."

3. They pointed out two different support strategies addressed to lawyers, depending on the type of clients they are dealing with. In this regard, they differentiate between a "business man", who always will raise business problems, or a "legal advisor". The later "comes to you looking for arguments to support a decision that he has previously made".

4. Concerns related to *transforming expert legal language into natural language.*

TOPIC 5: The need of creating a subsidiary in another Member State

1. They indicate that team-setting tasks are carried out according to the case. Sometimes they resort to subcontracting local lawyers. Normally they start by resorting to those offices with which they have agreements.

2. Usually, they use English language within this process, but they pointed out the need of having access to databases from different countries. In this regard, an accurate translation in English or Spanish would be a big progress.

TOPIC 6: Suggestions provided by the participants related to the Lynx Platform functionalities

1. They suggested the possibility of setting up an important number of agreements at company level related to different sectors and in different languages such as Spanish, Catalan, Basque and Galician with the aim of testing the Lynx Platform.

2. When a regulatory change occurs, a superficial analysis is always detrimental. An urgent deeper analysis is required because "if not you are late". They stress that they need to be careful with inaccuracies or disinformation. They compare this situation with journalists covering courts and legal related information.

   As a result, and related to the last suggestion provided, researchers identified the workflow as in Figure 8.
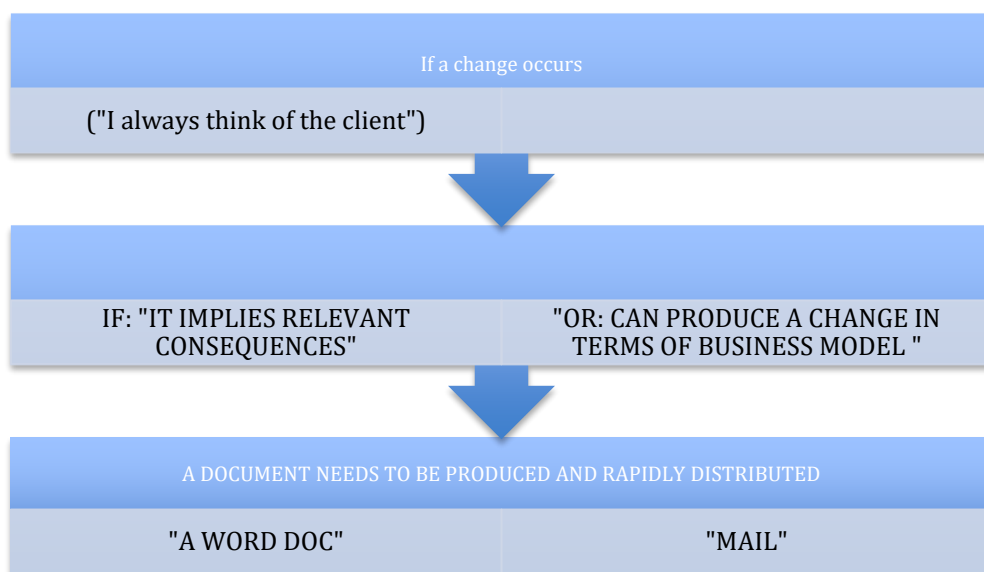
Figure 8. Workflow on regulatory changes

3. A final suggestion is provided related to the summary: comments or technical clarifications on the summaries cannot go against law firm criteria in a particular topic. They insisted on the fact that they do not perform interpretations of the legal content provided through summaries. Their task is to provide neutral information about what this would mean: the implications of significant changes regarding regulatory compliance.

4 CONCLUSIONS

4.1 LEGAL COMPLIANCE

Legal and business requirements for compliance (especially for compliance by design) have attracted much attention [Casanovas, Palmirani, Peroni, van Engers, Vitali2016]. Previous EU projects — especially COMPAS, OPENLAWS, EU Cases, MIREL, and BO-ECLI — have developed conceptual toolkits. The Workshop on Requirements Engineering and Law (RELAW)[2], has been running from ten years now, led by specialised researchers such as Sepideh Ghanavati.

In a recent LYNX Workshop on Legal and Regulatory Compliance [Rodríguez-Doncel, Casanovas, González-Conejero2017], we presented some preliminary results from the survey we are carrying out [Casanovas, Gonzalez-Conejero and de Koker2018], after examining 280 works on compliance by design in the past fifteen years.

After examination of the state of the art, we suggested the concept of *Legal Compliance through Design* (LCtD) to complement LCbD by recognizing the role of social, political, and economic conditions (as pre-conditions) and governance and ethical requirements (as constraints) when designing legal compliance, encompassing norms and principles that require a balancing of competing rights, obligations or policies. Conditions for legal compliance are broader and more entangled than for regulatory compliance, as legal conditions can be described by means of rules, but rules alone do not play out the stakeholders' rights, duties, and legal effects of their behaviour.

We focused on the definition of legal (not only documentary) sources to select and define requirements. *Compliance through Design* (CtD) explicitly encompass the social and institutional aspects that are not explicitly included by the regular way of approaching this subject (i.e. legal interpretation processes — beyond the conversations between experts and computer scientists—, institutionalisation, the interface between modelling and coordination, and the relation between citizens, consumers, and the law).

This is coherent with Motta's assertion about the interdisciplinarity of descriptive empirical approaches (2013), and with the need to consider software requirements as *prescriptive* statements.[3]

Thus, the results summarised in 5.2 (Table 15 and 16) could be reframed into a general classification of legal sources, properties, and entity relations, respecting the autonomy and decision-making capacity of lawyers, rulers, administrators, companies, business-holders, and lay-people. This is compatible with the LYNX approach as well.

4.2 FUNCTIONAL REQUIREMENTS AND GENERAL EXPECTATIONS

Table 15 plots the functional requirements, as extracted from the surveys:

Functional Requirement regarding the Lynx Platform	Number of Mentions	Mentioned by Stakeholder Group
Monitor law, jurisdictions, regulatory compliance and alert users in case of changes, innovations, modifications	12	SME (5); LE (3); LA (4)
Provide access to (at least) the following content areas: tax law, labour law, required	2	LA (2)

---

[2] Cfr. http://gaius.isri.cmu.edu/relaw/2017/

[3] "A software requirement is a prescriptive statement to be enforced solely by the software-to-be and formulated only in terms of phenomena shared between the software and the environment". [van Lamsweerde2009, 19].

permits or necessary authorisations, and operating licenses		
Provide smart search services among relevant digital regulatory compliance documents that produce highly relevant results	12	SME (4); LE (1); LA (1); PPA (3); O (2)
Provide summaries of relevant documents	14	SME (3); LE (2); LA (4); PPA (3); O (2)
Provide translations of relevant documents	8	SME (2); LE (2); LA (1); PPA (3)
Provide smart references and links among the retrieved documents and any other potentially relevant documents	18	SME (7); LE (1); LA (4); PPA (3); O (3)
Provide topic classification within the documents	10	SME (3); LE (2); LA (3); PPA (1); O (1)
Provide temporal references in the documents	1	SME (1)
Provide recommendations of documents that may also be potentially relevant	7	SME (1); LE (1); LA (2); PPA (1); O (2)
Include relevant background information and add explanatory information to legal documents so that laypersons are able to understand them	3	LA (1); PPA (2)
System should exhibit high performance and be able to cope with a very large number of documents	15	SME (2); LE (2); LA (4): PPA (4), O (3)

Table 15. Functional requirements[4]

Table 16 summarises the expectations of potential end users of the Lynx Platform, provided through knowledge acquisition techniques (both, quantitative and qualitative) to achieve Task 1.1. The expectations have been extracted from gains and pains highlighted by the participants in relation to the specific functionalities that the Lynx Platform should provide. Its final goal is to enrich and facilitate the alignment with pilot user's requirements provided in D4.1.

General Requirements related to specific features of the Lynx Platform (Expectations)	
R1	Platform services should be customized according to the professional profile of the end users
R2	Summarization of digital regulatory compliance documents should be provided according to the professional profile: - SME, LE, needs to receive specific recommendations related to the relevant regulatory changes occurred within their respective business activity sector. - Consultancy and legal firms need to receive key information related to changes in regulatory compliance with the aim of empowering reasonable and optimal decisions. - LE Smart search among relevant regulatory documents would be welcome.

---

[4]We thank Georg Rehm (DFKI), Stefanie Hegele (DFKI) and Julian Moreno (DFKI) for his careful review: this table is one example of their work.

	- Identifying judgements that involve significant or radical changes in relation to previous legal framework would be useful. - Identification of key issues removed by the new legal framework with the aim of providing implications of significant changes regarding regulatory compliance. - Services to perform semantic analysis and linking of content contained within the documents. PAs need to provide interpretable legal information
R3	Alerts about changes in digital regulatory compliant documents should be provided
R4	Precise translation of digital regulatory documents should be provided
R5	Updating overview of all the applicable regulatory requirements with a link to their documents to support compliance management needs to be provided
R6	100% accuracy when setting relevant documents in particular scenarios: Providing an accurate classification of documents is really relevant
R7	High-speed updating process is demanded

Table 16. General requirements.

4.3   Interpretation: the Lynx "radar"

The information provided by qualitative interviews and the narratives of the focus group converges with the requirements summarized in Table 15 (smart search, updatings, links, cross references etc.).

This is also consistent with the findings of D4.1 (pilot use cases general needs). It is usual to search for plain legal information to satisfy e.g. a client's demand about how to start up a company in a foreign country. But in controversial cases going before the bench (and labor law is particularly adversarial), what is expected from the queries is not only information, but some guidelines or structure in order to build up the case.

The notion of "customization" of the service — i.e. adaptation to the needs of different end-users— and the metaphor of "radar", as used in the legal focus group, suggest an *intended meaning* which is implicit in this kind of narratives. E.g. Section 4.3. (Topic 4):

1. Legal advisors provide a 'summary': *arguments about key issues to make easier for the lawyer to choose one strategy or another, taking into account the client's needs*.
2. "Our lawyers need to know that they know everything. *We are like a radar system*. In this regard we should have a lead on the way the market is developing from a technical or legal perspective."
3. Lawyers expressed some concerns related to *transforming expert legal language into natural language.*

From our point of view, these expressions —'customization', 'radar', 'natural language'…— are referring to the legal argumentative process of interpretation and case-building. Actually, this is what a documentarist (para-legal) or internal legal advisor is supposed to do in benefit of a lawyer: selecting, summarising, repairing, commenting, and pointing out what is important, i.e. constructing criteria of relevance according to the sub-field, the lawyer, the jurisdiction, and the legal system as it may be conceived from a particular discipline.

Hence, we could differentiate at least between (i) *systemic* requirements (affecting the whole legal system as conceived by the stakeholders) (ii) *functional* requirements. The later ones can be prioritized and may eventually lead to building functionalities on the platform. The former ones are more generic, and denote the properties of the legal "ecosystem" they intend to construct or deal with.

Some of these systemic properties are *formal* properties, coincident with the requirements for Rule Interchange Languages in the Legal Domain identified by [Gordon, Governatori & Rotolo2009]: (i) 'isomorphism between norms and rules' (normative legal content can be expressed by rules), (ii)

reification (jurisdiction, authority, temporal properties), (iii) defeasibility (rule semantics: solution to normative conflicts), (iv) validity, (v) legal procedure, (vi) normative effects (including those related to new norms and derogation of the old one) etc.

Other properties refer to the information that end-users deem *relevant* to take action in setting their professional niche or "regulatory ecosystem". They are *substantive* properties, and could be identified through the particular narrative of each stakeholder, i.e. through the identification of the intended meaning of the narrative elicited in the knowledge acquisition process.

Is worth noting that there is a distance between (i) the entities captured by semantic formal languages from different regulatory sources, (ii) and the capacity of human agents of creating, interpreting, applying, implementing and eventually enforcing legal instruments and norms. Such a distance is a creative one: lawyers (or any other end-user or stakeholder) should be provided with the ability to adapt the functionalities of the platform to their needs.

Moreover, the legal knowledge graph —law and regulatory open data, interlinked and offered through a set of cross-sectorial, cross-lingual services— should respect the internal legal connection between regulatory instruments. I.e. the different legal value of hard law (legislation, case-based law, contracts, and civil obligations), policies (regulations and policies implemented by state and international agencies), soft law (recommendations, standards, best practices, and protocols), and ethics (evaluations by ethical committees, e.g. in bio-ethics and privacy and data-protection committees (according to the recent GDPR).

ANNEX I – ELECTRONIC PARTICIPANT CONSENT FORM

The participant consent form is placed at the beginning of the survey and participants must read and voluntarily agree to it before allowed to proceed. In order to ease the readability of the text, the electronic participant consent form is divided in two parts: the first one contains a description of the project, information about personal data protection issues and the explicit consent to participate in the survey; the second part is accessible through a clickable link, in this link the complete note is shown, facilitating to print a copy for the participants records.

Description of the project

The main objective of the H2020 Lynx project (http://lynx-project.eu) is to provide more effective ways of accessing a huge amount of digital regulatory compliance documents, including legislation, case law, standards, industry norms and best practices. In particular, the Lynx solution envisages an ecosystem of smart cloud services to better manage compliance. This solution integrates processing, analysis, handling and linking of digital regulatory compliance documents. In particular, the Lynx ecosystem will enable smart search, smart assistance and smart referencing of case law as well as Artificial Intelligence technologies and automatic translation of regulatory compliance documents. In summary, the Lynx solution will provide a legal -- knowledge and information-- services one-stop shop for SMEs and companies operating internationally.

This questionnaire is the first activity of the H2020 Lynx project and aims at collecting functional requirements regarding the Lynx project solution to be deployed. Thus, respondents have now the chance of providing important and valuable input that will shape the future features provided by the Lynx solution. Specifically, the elicitation of functional requirements in this questionnaire addresses your current strategy for search, analysis, processing, monitoring and managing of regulatory compliance documents and, your pains and gains when dealing with digital regulatory compliance documents.

Thank you very much for your collaboration.

Personal Data Protection

The Lynx project is committed to user privacy. The specific policy for the protection of your privacy has been designed on the basis of Regulation 2016/679 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).

Electronic consent

Please read carefully the complete legal notice at http://idt.uab.cat/index.php/lynx-survey-legal-notice and select your choice below. You may print a copy of this consent form for your records. Clicking on the "Agree" button indicates that:

- You have read and understand the above information
- You voluntarily agree to participate
- You are 18 years of age or older

[ ] Agree   [ ] Disagree

Table 17. First part of the Electronic participant consent.

Description of the project

You are about to enter the Online survey designed in the context of the Lynx Project. The main objective of the H2020 Lynx project is to provide more effective ways of accessing a huge amount of digital regulatory compliance documents, including legislation, case law, standards, industry norms and best

practices. In particular, the Lynx solution envisages an ecosystem of smart cloud services to better manage compliance. This solution integrates processing, analysis, handling and linking of digital regulatory compliance documents. In particular, the Lynx ecosystem will enable smart search, smart assistance and smart referencing of case law as well as Artificial Intelligence technologies and automatic translation of regulatory compliance documents. In summary, the Lynx solution will provide a legal -- knowledge and information-- services one-stop- shop for SMEs and companies operating internationally.

As an example, enterprises processing personal data in the EU are affected by existing and upcoming data protection legislation at European, national and regional levels. Besides, enterprises are also affected by industry norms and standards on the basis of their specific activity. As a result, legal advisors and/or legal experts have to face a large amount of digital documents even in multiple languages. The new General Data Protection Regulation (GDPR) foresees administrative fines up to 20 million Euro or 4% of the enterprise annual revenues. Therefore, legal compliance in the area of data protection and privacy becomes a key issue for enterprises.

In another example it is possible to see that, the international expansion of enterprises may include new premises and new staff in a foreign country, thus making local labour law or tax obligations, among others, a crucial issue. Moreover, legal information, legal data and legal documents are usually only available in the local language. Therefore, in this scenario legal experts in addition to having to go through hundreds or thousands of digital documents, encounter language as a mayor constraint.

Thank you very much for your collaboration.

Disclaimer

The questions contained in the survey are:

- Of a general nature only and not intended to address the specific circumstances of any particular individual or entity
- Not necessarily comprehensive, complete, accurate or up to date
- Not professional or legal advice (if you need specific advice, you should always consult a suitably qualified professional)

The Lynx project and the Consortium are not responsible for the opinions provided by the participants and for any misuse of this questionnaire.

However, this disclaimer is not intended to limit the liability of the Lynx Consortium in contravention of any requirements laid down in applicable European or national law.

Personal Data Protection

The LYNX project is committed to user privacy. The specific policy for the protection of your privacy has been designed on the basis of Regulation 2016/679 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).

Information collected

This questionnaire is the first activity of the H2020 Lynx project and aims at collecting functional requirements regarding the Lynx project solution to be deployed. Thus, respondents have now the chance of providing important and valuable input that will shape the future features provided by the Lynx solution. Specifically, the elicitation of functional requirements in this questionnaire addresses your current strategy for search, analysis, processing, monitoring and managing of regulatory compliance documents and, your pains and gains when dealing with digital regulatory compliance documents.

Purpose of the collection

The results of this study will be used for scientific and scholarly purposes only. In particular this online survey intends to all the functional requirements provided by final users and stakeholders (SMEs, BAR associations, etc.).

Recipients of the information

Your replies will be shared with the Lynx Consortium

Retention period

The results of this survey will be stored for a 5-year period in order to comply with the European Union requirements for possible audits of the results of the project.

Confidentiality and data security measures

Your survey answers will be sent to Universitat Autonoma de Barcelona where data will be stored in CSV format.

Rights

Your participation in this research study is voluntary. You may choose not to participate. If you decide to participate in this research survey, you may withdraw at any time. If you decide not to participate in this study or if you withdraw from participating at any time, you will not be asked the reasons why. You may access, rectify or erase any data collected at any time during the retention period.

Contact

If you have any questions about the research study, or you want to exercise your rights please contact: Jorge González-Conejero at jorge.gonzalez.conejero@uab.cat

Table 18. Complete Electronic participant consent form.

ANNEX II – PARTICIPANT CONSENT FORM

The informed consent for the participants in the interviews and focus groups conducted within the Task 1.1 is designed considering the ethical guidelines discussed in Annex I. Therefore, Table 19 shows the Part I of the informed consent. The first part consists of an information sheet about the Lynx project that includes: the purpose of the data collection, the contact person for the activity, some details about the activity, the possible risks and the types of data to be collected. Some of the fields—highlighted in yellow—must be filled in by the organizer of the interview or focus group.

Purpose of the research and of data collection

Project description: The main objective of the H2020 Lynx project is to provide more effective ways of accessing a huge amount of digital regulatory compliance documents, including legislation, case law, standards industry norms and best practices. The Lynx solution envisages an ecosystem of smart cloud services to better manage compliance. This solution integrates processing, analysis, handling and linking of digital regulatory compliance documents. In particular, the Lynx ecosystem will enable smart search, smart assistance and smart referencing of case law as well as Artificial Intelligence technologies and automatic translation of regulatory compliance documents. In summary, the Lynx solution will provide a legal --knowledge and information-- services one-stop-shop for SMEs and companies operating internationally.

Questions are aimed at obtaining valuable input that will shape the future features provided by the Lynx solution. Specifically, the elicitation of functional requirements in this questionnaire addresses your current strategy for search, analysis, processing, monitoring and managing of regulatory compliance documents and, your pains and gains when dealing with digital regulatory compliance documents.

Contact person responsible for the activity

Name and surname [to be filled out by the involved partner]
Name and surname [to be filled out by the involved partner]

Activity details

Exercise Plan Form:

You are being invited to take part in this interview because of your experience as an expert in [add respondent expertise]. If you choose to take part to the interview, you will be asked to provide your professional view on a series of topics concerning strategies for the collection, analysis, processing and management of regulatory compliance.

In particular, the interview consists of 12 questions aimed at investigating/discussing the requirements, challenges, opportunities and obstacles regarding the collection and management of regulatory compliance in SMEs and Large Enterprises internationalization and in [concerned country].
Depending on your preference, the interview can be held:
   a. in person (about 45 minutes/1 hour) -;
   b. by telephone or via Skype (about 45 minutes/1hour);
   c. in writing, i.e. by filling out a written questionnaire to be returned to me.

Representatives of the Lynx project will carry out the interview. It can be carried out in English or in [please insert your native language, if needed]
If you agree and provide your written consent, the interview could be audio/video recorded (see below). Recordings will be used for the purpose of transcribing the content of the interview only, and

be destroyed after transcript. They will not be disseminated without your explicit, freely given and written consent. Please, mind that you can legitimately choose not to be recorded. In case you deny giving us your consent, this will not negatively affect your participation to the interview.

The person conducting the interview will process data provided by you through the interview. This data will be anonymised and then shared with the Lynx consortium. Information will input project Deliverable D1.1. The Deliverable will describe the functional requirements for the Lynx solution provided by final users and stakeholders.

Possible risks

There could be a risk that you may share some confidential information by chance, or that you may feel uncomfortable talking about some of the topics. However, we do not wish for this to happen. You do not have to answer any question or take part in the discussion if you feel the question(s)/topic are too personal or if talking about them makes you uncomfortable. If you say anything that you then realise you do not want to be reported you can inform us at any moment during the interview or when it is ended.

Incentives

You will not be provided any incentive to take part in the interview. If you like and consent, your participation to the interview – and your contribution to the project – can be explicitly acknowledged in the Project deliverable.

Types of data to be collected

Your personal data (name, surname, professional contact details and affiliation) will be collected and processed by [please add your institution name]'s team working on Lynx. Data will be processed for the purpose of running the interview within Lynx and carrying out project activities only. Your data will be gathered on paper and on computer files, stored in our offices at [please add your institution name] premises and accessed only by us or other selected personnel who might be authorised to work on Lynx.

If you opt for being recorded, your voice and/or image will be processed too. Recordings will be used for the purpose of transcribing the content of the interview only, and be destroyed after transcript. They will not be disseminated without your explicit, freely given and written consent.

Any personal or professional views or opinions you might express during the interview will be then processed and rendered anonymous before dissemination within the Lynx consortium. This implies that by no means the opinions or views you expressed will be related directly to your personal data in the final project deliverable. In other words, your opinions or views will be processed in a way that inhibits tracing them back to you.

However, if you voluntary consent in writing, we can make evident the association of the expressed opinions or views with your identity whereas relevant and appropriate. This means that you can agree to have your name and affiliation together with your expressed opinion reported in the final deliverable that documents the interview process. You are free to decide whether providing us with your specific consent or not (see below). In case you deny giving us your consent, this will not negatively affect your participation to the interview.

Table 19. Lynx Informed consent for the interviews and focus groups. Part I – Lynx information sheet.

Table20 shows the Part II of the informed consent, which includes: the voluntary participant data, the right to withdraw, confidentiality details, the applicable laws and Directives, date and place of the

interview or focus group and the declaration. The organizer of the interview or focus group must fill some of the fields such as participant data and data.

Voluntary participant data

Name and surname of the participant

Voluntary participation and Right to withdraw

Your participation in the Lynx project is voluntary.

You are free to withdraw from the project, without giving a reason for your withdrawal and without any consequences to your future treatment by the researcher.

You retain all rights provided by the applicable data protection legislation and, in any case:

- Information
- Rectification
- Erasure
- To be forgotten
- Access
- Restriction of processing

If you decide to withdraw from the project, please contact the Lynx contact person(s).

You should know that you may be withdrawn from the project for any of the following reasons:

- If you don't follow the Consortium instructions.
- If you don't attend the scheduled data collection sessions.
- If the whole project is stopped, for reasons not known now.

Confidentiality

The Lynx representatives who see/access this information will keep it confidential.

Applicable Laws/Directives

- European legislation: Regulation (EU) 2016/679.
- European guidelines: opinions and recommendations by the European Data Protection Supervisor, the recently appointed Ethics Advisory Group and the Article 29 Working Party.
- National legislation: relevant national rules.

Date and place

Date and place of the interview

Declaration

I have read the foregoing information; I have had the opportunity to ask questions about it and questions have been answered to my satisfaction. By signing the Form, I acknowledge that I have understood and agreed to the above terms.

Signature

_____

CONSENT FOR SPECIFIC ACTIVITIES
Video/Audio Recording
I consent voluntary to be video or audio recorded during the interview.
□ I DO consent voluntarily to be recorded;
□ I DO NOT consent voluntarily to be recorded.

Signature

_____

Explicit Acknowledgment

I consent voluntary that any personal or professional views or opinion that I provided during the interview will be associated with my name, surname and affiliation and reported – where relevant and useful - in the final report that will be accessible in print and/or on-line.

☐ I DO consent voluntarily to my personal data to be processed;
☐ I DO NOT consent voluntarily to my personal data to be processed.

Signature

_____

Table 20. Lynx informed consent for the interviews and focus groups. Part II – Certificate of consent.

ANNEX III – SURVEY RESPONSES

Q.	SME.S.1	SME.S.2	SME.S.3	SME.S.4	SME.S.5
Country	Austria	Germany	Germany	Latvia	Germany
Strategy					
Q5	This process is carried out in-house	This process is externalized to another entity (Consultancy agency or Legal Advisor)	This process is carried out in-house	This process is carried out in-house	This process is externalized to another entity (Consultancy agency or Legal Advisor)
Q.8	In individual case; according to work-needs	--	Exchange of documents between subsideries in Germany, Australia, India.	Internal laweyrs follow up the latest changes in digital regulatory and apply changes as required.	--
Q.9	LexisNexis; provides topic classification	--	no, mainly via Mail	--	--
Q.10	a shared folder	--	shared folder (e.g. dropbox)	Shared folders, data bases	--
Q.11	IT Tool (LexisNexis)	--	manually, via Mail	Monitor the governmental activities, participate in discussions	--
Q.12	I provide a summary of these changes every month	--	Anglo saxon rights and BGB rights are very different and need humand communication	Project based consultations, blogs and articles on the topic in media, socila networks etc.	--
Q.13	I send a summary of selected relevant documents	--	sending documents together with calls and meetings	Summary, presentation based consutlation, relevant	--

				documents. Depends on client and situation.	
Q.14	No	--	no	No	--

Pains

Q.15	Costs too much money	Costs too much money	Takes a lot of time Costs too much money Requires substantial efforts need always lawyers from all involved parties	Costs too much money	Costs too much money
Q.16	Translation of digital regulatory compliance documents Summarization of digital regulatory compliance documents Alert of changes in digital regulatory compliance documents Linking of digital regulatory compliance documents based on their content	Alert of changes in digital regulatory compliance documents	Translation of digital regulatory compliance documents Summarization of digital regulatory compliance documents Recommendation of digital regulatory compliance documents based on their content Alert of changes in digital regulatory compliance documents Topic classification of digital regulatory compliance documents	Topic classification of digital regulatory compliance documents Smart search among relevant digital regulatory compliance documents Smart referencing of case law	Summarization of digital regulatory compliance documents Alert of changes in digital regulatory compliance documents Smart search among relevant digital regulatory compliance documents Smart referencing of case law

| | | | Linking of digital regulatory compliance documents based on their content

Smart search among relevant digital regulatory compliance documents | | |
|---|---|---|---|---|---|
| Q.17 | the collected documents are not related enough to their scenario | no | difficult handling of different laws for companies and financial accounting | Not clear the how to introduce the regulation. | -- |
| Q.18 | to analyse a huge amount of documents in a short period of time | not aware of any | translation and transfer of texts for state and financial authorities of different countries | hudge amount of documents, accuracy in aplication of the regulations | -- |
| Q.19 | to monitor changes in regulatory compliance documents | no | optimizing costs and bureaucratic efforts for daily procedures (annual balance, change of share ownership) | All mentioned in example | -- |

Gains

Q.20	Time	Time Money Effort	Time Money Effort	Time Money Effort	Time Money Effort
Q.21	Translation of digital regulatory compliance documents	Alert of changes in digital regulatory compliance documents	Translation of digital regulatory compliance documents

Summarization of digital regulatory compliance documents | Translation of digital regulatory compliance documents

Summarization of digital regulatory | Summarization of digital regulatory compliance documents

Recommendation of digital regulatory compliance |

			Recommendation of digital regulatory compliance documents based on their content Alert of changes in digital regulatory compliance documents	compliance documents Recommendation of digital regulatory compliance documents based on their content Alert of changes in digital regulatory compliance documents Topic classification of digital regulatory compliance documents Linking of digital regulatory compliance documents based on their content Smart search among the relevant regulatory compliance documents Smart referencing of case law	documents based on their content Alert of changes in digital regulatory compliance documents Smart search among the relevant regulatory compliance documents Smart referencing of case law
Q.22	More services	Lower costs of money when dealing with regulatory compliance documents	More services Lower costs of money when dealing with regulatory	More services Lower costs of money when dealing with regulatory	More services Lower costs of money when dealing with regulatory

Q.					
		Lower efforts when dealing with regulatory compliance documents	compliance documents Lower efforts when dealing with regulatory compliance documents	compliance documents Lower efforts when dealing with regulatory compliance documents	compliance documents Lower efforts when dealing with regulatory compliance documents
Q.23	--	not sure	precise translation of agreements for different countries and languages	--	--
Q.24	a high-speed update process when a change in regulatory compliance occurs	not sure	90	All mentioned	--

Table 21. Answers by SMEs

Q.	LE.S.1	LE.S.2	LE.S.3		
Country	Norway	Italy	Italy		
Strategy					
Q5	This process is carried out in-house	This process is externalized to another entity (Consultancy agency or Legal Advisor)	This process is carried out in-house		
Q.8	Partly by expert knowledge, partly by defining (public available) standards and requirement documents	--	Internal approval after external partner suggestion		
Q.9	Yes; inclusion of metadata in an inhouse production tool	--	no IT tool		

Q.10	Database	--	Shared folder		
Q.11	Delegation to expert (teams) who establish shared overviews	--	Our internal legal team		
Q.12	Summary of changes by mail, a requirement overview for specific applications	--	With email		
Q.13	We define a set of rules to be follow	--	Our legal team send a summary		
Q.14	No	--	No		

<div align="center">Pains</div>

Q.15	Takes a lot of time Overlapping requirement standards giving way to overspecificati-on	Takes a lot of time	Takes a lot of time		
Q.16	Summarization of digital regulatory compliance documents Alert of changes in digital regulatory compliance documents Topic classification of digital regulatory compliance documents Smart search among relevant	Translation of digital regulatory compliance documents Summarization of digital regulatory compliance documents Recommendation of digital regulatory compliance documents based on their content	Translation of digital regulatory compliance documents		

	digital regulatory compliance documents	Alert of changes in digital regulatory compliance documents			
Q.17	Challenges to control change management (from revision to revision)	no	No		
Q.18	to find out documents related to already relevant regulatory compliance documents	to update a lot of documents in serveral languages at the same time	Increase accurary of document		
Q.19	monitor changes in regulatory compliance	translation and publication of contents	Monitor changes		
Gains					
Q.20	Time Effort	Time Money Effort	Time		
Q.21	Summarization of digital regulatory compliance documents Alert of changes in digital regulatory compliance documents Smart search among the relevant regulatory compliance documents	Translation of digital regulatory compliance documents Summarization of digital regulatory compliance documents Recommendation of digital regulatory compliance documents based on their content	Alert of changes in digital regulatory compliance documents		

Q.22	Lower efforts when dealing with regulatory compliance documents	More services	Lower costs of money when dealing with regulatory compliance documents		
	Alert of changes in digital regulatory compliance documents				
Q.23	An up to date overview of all the applicable regulatory requirements with a link with their documentation to support compliance management (up to date, i.e. including revision control)	--	--		
Q.24	a high-speed update process when a change in regulatory compliance occurs	100% accuracy	a high-speed update process when a change in regulatory compliance occurs		

Table 22. Answers by LE.

Q.	LA.S.1	LA.S.2			
Country	Spain	Italy			
Strategy					
Q.8	difficult to answer (talk about strategy is not a clear/concrete thing)	we pay attention to legislative innovations through official bulletins such as Gazzetta Ufficiale for			

		domestic law and EurLex for European law			
Q.9	Expert System Cogito Discover	online database			
Q.10	NetDocuments (Document Management System on the cloud)	legal database			
Q.11	We have an specialized KM team, but we normally use BOE alerts and specific tools like Lexis Nexis, VLEX ... that provide us legal contents and it has the possibility to define some alerts	throug ufficial buletins			
Q.12	I don't understand what do you mean as "updating process"	preparing reports and dossier			
Q.13	we have created specific automated subscription services to inform to our clients, but also we use some other channels, like blogs, social media and website and extranets	I send a summary of selected relevant documents with a minimum of comment			
Q.14	Yes, Expert System Cogito	NO			
Pains					

Q.15	Takes a lot of time; Costs too much money	Requires substantial efforts				
Q.16	Recommendation of digital regulatory compliance documents based on their content; Alert of changes in digital regulatory compliance documents; Topic classification of digital regulatory compliance documents; Linking of digital regulatory compliance documents based on their content; Smart referencing of case law	Recommendation of digital regulatory compliance documents based on their content				50
Q.17	I don't know	the collected documents are too complicate to be clearly understood without a specific legal support				
Q.18	to find out documents related to already relevant regulatory compliance documents	to increase the accuracy of the documents considered as relevant				
Q.19	difficult to answer	to monitor changes in regulatory				

		compliance documents			
Gains					
Q.20	Time;Money;Eff ort	Money			
Q.21	Translation of digital regulatory compliance documents; Recommendatio n of digital regulatory compliance documents based on their content; Alert of changes in digital regulatory compliance documents; Topic classification of digital regulatory compliance documents; Linking of digital regulatory compliance documents based on their content; Smart referencing of case law	Alert of changes in digital regulatory compliance documents			
Q.22	More services (For instance, translation of documents; topic classification, summary of relevant documents);Lo wer costs of money when dealing with	Lower costs of money when dealing with regulatory compliance documents			

Q.		
Q.23	--	--
Q.24	--	a high-speed update process when a change in regulatory compliance occurs

(continued) regulatory compliance documents; Lower efforts when dealing with regulatory compliance documents

Table 23. Answers by Legal Advisor (Law firm or Lawyer).

Q.	PPA.S.1	PPA.S.2	PPA.S.3		
Country	Spain	Spain	Spain		
Pains					
Q.15	Takes a lot of time; Costs too much money; Requires substantial efforts; Any of them defines it as too costly.	Takes a lot of time; Costs too much money; Requires substantial efforts	Takes a lot of time; Costs too much money; Requires substantial efforts		
Q.16	--	Summarization of digital regulatory compliance documents; Smart search among relevant digital regulatory compliance documents	Summarization of digital regulatory compliance documents; Smart search among relevant digital regulatory compliance documents		
Q.17	Yes. Norms and standards ignorance	Lack of summarized infirmation	Lack of summarized infirmation		

Q.18	The lack of easy-to-understand documents.	To find out valuable and understanding information	To find out valuable and understanding information		
Q.19	I am not sure.	To access to the relevant documents	To access to the relevant documents		
			Gains		
Q.20	Time;Money;res ponsability.	Money;Effort	Money;Effort		
Q.21	Translation of digital regulatory compliance documents; Summarization of digital regulatory compliance documents; Recommendatio n of digital regulatory compliance documents based on their content; Alert of changes in digital regulatory compliance documents; Topic classification of digital regulatory compliance documents; Linking of digital regulatory compliance documents based on their content; Smart search among the relevant regulatory compliance	Translation of digital regulatory compliance documents; Summarization of digital regulatory compliance documents; Smart search among the relevant regulatory compliance documents	Translation of digital regulatory compliance documents; Summarization of digital regulatory compliance documents; Smart search among the relevant regulatory compliance documents		

Q.			
documents; Smart referencing of case law; Easy-to-understand documents for non specialized people.			
Q.22	Lower costs of money when dealing with regulatory compliance documents; Lower efforts when dealing with regulatory compliance documents	More services (For instance, translation of documents; topic classification, summary of relevant documents);Lower efforts when dealing with regulatory compliance documents	More services (For instance, translation of documents; topic classification, summary of relevant documents);Lower efforts when dealing with regulatory compliance documents
Q.23	--	--	--
Q.24	Documents easy-to-understand.	Accuracy and updated information	Accuracy and updated information

Table 24. Answers by Public/Private Agencies.

Q.	O.S.1	O.S.2		
Country	Germany	Netherlands		
Description	Public Foundation for Cultural Heritage	Third party assessment		
Strategy				
Q.8	Workflow for document digitisation, text recognition (OCR), indexing and online publication	--		

Q.9	Kitodo (http://www.kitodo.org/)	--			
Q.10	Fedora (http://fedora-commons.org/)	--			
Q.11	Not applicable	--			
Q.12	Not applicable	--			
Q.13	Not applicable	--			
Q.14	Currently not as a "standard" feature though realized via plugin modules (typically research projects in the NLP domain)	--			

<div align="center">Pains</div>

Q.15	Requires substantial efforts	Takes a lot of time; Costs too much money; Requires substantial efforts			
Q.16	Summarization of digital regulatory compliance documents; Recommendation of digital regulatory compliance documents based on their content; Linking of digital regulatory compliance documents based on their content; Smart search among relevant digital regulatory	Translation of digital regulatory compliance documents; Summarization of digital regulatory compliance documents; Recommendation of digital regulatory compliance documents based on their content; Alert of changes in digital regulatory compliance			

	compliance documents	documents; Topic classification of digital regulatory compliance documents; Linking of digital regulatory compliance documents based on their content; Smart search among relevant digital regulatory compliance documents			
Q.17	The quality of links that are derived between documents and other external knowledge bases is often too unspecific	all of the above mentioned			
Q.18	The amount of documents is very high, the quality of those documents is bad (due to their historical nature), simple full-text search does not provide sufficient relevancy for user satisfaction	relevant accuracy, accessibility, comparing one country with another			
Q.19	Due to bad quality and historical nature of the documents, the quality of currently	access to relevant documents, storage of design data, impact analysis			

Q.22	available NLP methods for analysing and deriving semantic structures in these documents is insufficient	of design changes			

Gains

Q.20	Effort	Time;Money;Effort			
Q.21	Summarization of digital regulatory compliance documents; Recommendation of digital regulatory compliance documents based on their content; Linking of digital regulatory compliance documents based on their content; Smart search among the relevant regulatory compliance documents	Recommendation of digital regulatory compliance documents based on their content; Alert of changes in digital regulatory compliance documents; Topic classification of digital regulatory compliance documents; Linking of digital regulatory compliance documents based on their content; Smart search among the relevant regulatory compliance documents; Smart referencing of case law			
Q.22	More services (For instance, translation of documents;	Lower efforts when dealing with regulatory			

	topic classification, summary of relevant documents);Reduction of manual effort required through the assistance of (semi-)automated document processing and analysis	compliance documents				58
Q.23	Services to perform semantic analysis and linking of content contained in the documents	accessibility, storage (archive), automated reference to and selection of applicable standards				
Q.24	Highest possible level of accuracy with least manual effort	100%				

Table 25. Answers by respondents in "Other" category.

ANNEX IV – INTERVIEW RESPONSES

Interview participant data
IP.SME.1 (vLex)
SME; Spain
Coordinator:
UPM

Question 1: Strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents

- They are suscribed to every major legal information provider.

Question 2: When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, How do you monitor this change?

- After an analysis, customers are alerted (balloons in Windows desktops, notification in the app)

Question 3: When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, how do you carry out the updating process?

- Automatically

Question 4: How do you provide your customers/entity's responsible with the relevant regulatory compliance in their scenario? (For instance, I send the whole relevant documents; I send a summary of selected relevant documents; I manually/automatically define a set of simplified rules to follow; etc.)

- Whole documents provided by search.

Question 5: How do your customers/entity define "too costly"?

- Not applicable

Question 6: Which features are your customers/entity missing?

- They are working on solutions for services: recommendation, smart referencing among related documents, alerts, topic classification, smart search, temporal references in documents.

Question 7: Are there performance issues or malfunctions your customers/entity cite? (For instance, the collecting of regulatory compliance documents takes a lot of time/resources; the collected documents are not related enough to their scenario, the access to the collected documents is slow, etc.)

- No

Question 8: What are the main general challenges your customers/entity encounter? (For instance, to analyse a vast number of documents in a short period of time; to increase the accuracy of the documents considered as relevant; to find out documents related to already relevant regulatory compliance documents; etc.)

- Apply artificial intelligence algorithms and data analytics

Question 9: Do your customers/entity face specific difficulties getting certain tasks done? (For instance, to access and share relevant documents; to monitor changes in regulatory compliance documents; to determine time references within documents; etc.)

	• Not answered

Question 10: Which specific features would your customers/entity enjoy?

	• They are working on improving the services mentioned in question 6 for their customers within the same jurisdiction and are interested on adapting them to perform the services across jurisdictions.

Question 11: What would make your customers'/entity's day to day workload easier?

	• They don't know

Question 12: What level of performance do your customers/entity expect? (For instance, a 100% accuracy when setting relevant documents in their scenario; a high-speed access to relevant documents; a high-speed update process when a change in regulatory compliance occurs; etc.)

	• High

Table 26. vLex (SME) interview.

Interview participant data
IP.PPA.1 (Internationalization Agency)
Public Agency; Spain
Coordinator:
UAB

Question 1: Strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents

	• Not applicable

Question 2: When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, How do you monitor this change?

	• Not applicable

Question 3: When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, how do you carry out the updating process?

	• Not applicable

Question 4: How do you provide your customers/entity's responsible with the relevant regulatory compliance in their scenario? (For instance, I send the whole relevant documents; I send a summary of selected relevant documents; I manually/automatically define a set of simplified rules to follow; etc.)

	• IP.PPA.1 has international premises with technical staff that help companies in their internationalization process • The international offices act as a consultancy agency. The degree of consultancy depends on the office staff expertise • When no expert staff is in the specific office, they recommend a set of local consultancy firms • IP.PPA.1's office focused on internationalization barriers for companies is recently created

Question 5: How do your customers/entity define "too costly"?

	• Consultancy services charge a subsidized fee to the company • They have carried out more than 1.200 consultancy projects during the last year

Question 6: Which features are your customers/entity missing?

	• Companies that are seeking internationalization will enjoy a software or application able to provide interpretable legal information. It is not a matter of language since they act with offices in the target country or with local consultancy firms, it is a matter of the technicity of legal and regulatory documents

Question 7: Are there performance issues or malfunctions your customers/entity cite? (For instance, the collecting of regulatory compliance documents takes a lot of time/resources; the collected documents are not related enough to their scenario, the access to the collected documents is slow, etc.)

Question 8: What are the main general challenges your customers/entity encounter? (For instance, to analyse a vast number of documents in a short period of time; to increase the accuracy of the documents considered as relevant; to find out documents related to already relevant regulatory compliance documents; etc.)

	• They identify two main problems in the internationalization in the European Single Market: 1) International public contracting (rules from the Single Market are not followed); 2) Workers mobility • In International public contracting in some European countries is mandatory to include a local company. It is against the Single Market rules and it is a major concern in the European Institutions (DG-Growth)

Question 9: Do your customers/entity face specific difficulties getting certain tasks done? (For instance, to access and share relevant documents; to monitor changes in regulatory compliance documents; to determine time references within documents; etc.)

	• Not applicable

Question 10: Which specific features do your customers/entity would enjoy?

	• Not applicable

Question 11: What would make your customers'/entity's day to day workload easier?

	• Not applicable

Question 12: What level of performance do your customers/entity expect? (For instance, a 100% accuracy when setting relevant documents in their scenario; a high-speed access to relevant documents; a high-speed update process when a change in regulatory compliance occurs; etc.)

	• Not applicable

Table 27. Public Agency interview.

ANNEX V – FOCUS GROUPS RESPONSES

Focus group participant data
FGP.LA.1 (Tax Law), FGP.LA.2 (Labour Law), FGP.LA.3 (ICT Lawyer)
Legal Advisor (Law Firm and Large Enterprise); Spain
Coordinator:
UAB-IDT

Question 1: Strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents

Strategy #1: CIJ	- The first strategy is focused on the compile and enrich of legal documents (legislation, case law, legal doctrine (books and journals). - This strategy is carried out by the "Centro de Información Jurídica" (CIJ), Center for the Legal Information. FGP.LA.3 is part of this team. - They elaborate a summary of each legal document - Then, the document is tagged from a list of static tags in order to classify the document. The main source for the document classification is the summary. - Finally, documents are stored in an internal database - The whole process is carried out manually by experts in the ACI team - They have external commercial solutions for the search and analysis of the legislation. They have "all of them" - However, they prefer to consult directly to the source of the information (for instance, a particular court decision. Commercial tools are too slow to take into account the most recent sentences - The relevant documents are stored in the internal database. Thus, from 60% to 70% of the documents are duplicated within the internal database. The ones searched through the external commercial tools and the ones classified manually by the experts within the group - The internal database is accessible through a knowledge tool self-developed by CC - The search in the internal database is full-text. It also implements prioritization by word distance - If the content of the summary is not accurate, thus the classification is not accurate.
Strategy #2	- The second strategy is focused on the updating, standardization, support information and legal standard of the law firm (criterion) - Strategy #2 is more focused on the customers side - In this second strategy the search of the documents is carried out completely manually - They check for the information provided by courts, specialized blogs from reputable and experienced lawyers - There is a list of resources that every member of the team check, but there is no standardized protocol.

	• They download the documents and then they print these documents in order to analyse in deep the information. • The printed document is underlined with relevant facts. • A word document is generated with this analysis • Finally, the information is sent to a law firm lawyer or to a customer by e-mail • The process should be quick, otherwise, other law firms could contact the customer first

Question 2: When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, How do you monitor this change?

Digital documents	• Changes in documents (legislation, case law, etc.) is manually monitored by experts in Strategy #1 • They check the sources mentioned in the previous question
Standardization and criterion	• Changes in documents (legislation, case law, etc) is manually monitored by experts in Strategy #2 • They check the sources mentioned in the previous question

Question 3: When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, how do you carry out the updating process?

Digital documents	• When a change is detected, the procedure is the same to Strategy #1. They provide a summary and the corresponding tags. • The system contains all the documents and summaries tagged and also the subscription preferences of lawyers (subscription rules) • Every day, the system sends an e-mail containing the changes and novelties
Standardization and criterion	• When a change is detected, the procedure is the same to Strategy #2. They process and analyse the changes from their usual sources specified in Question 1. • They inform of the novelties that could affect the legal standard (legal criterion) of the law firm. • An e-mail is sent weekly • They also carry out internal sessions bi-weekly or monthly

Question 4: How do you provide your customers/entity's responsible with the relevant regulatory compliance in their scenario? (For instance, I send the whole relevant documents; I send a summary of selected relevant documents; I manually/automatically define a set of simplified rules to follow; etc.)

By e-mail	• Daily e-mail with novelties from Strategy#1 • Weekly e-mail with novelties and DOC document with customer side analysis
Internal sessions	• From Strategy #2 outputs bi-weekly and monthly meetings are scheduled

Question 5: How do your customers/entity define "too costly"?

Key costly aspects	• Time • Effort

Question 6: Which features are your customers/entity missing?

Change alerts	• Warnings for legislation changes from external platforms are too slow
Languages	• There are no relevant issues with languages • They work with Spanish, different premises in another languages work in their language • Apart from Spanish, they mainly receive documents in English • Legal concepts are not translated in summaries for Strategy#1. The semantic differences among both languages is difficult to translate

Question 7: Are there performance issues or malfunctions your customers/entity cite? (For instance, the collecting of regulatory compliance documents takes a lot of time/resources; the collected documents are not related enough to their scenario, the access to the collected documents is slow, etc.)

	• No specific malfunctions are reported

Question 8: What are the main general challenges your customers/entity encounter? (For instance, to analyse a vast number of documents in a short period of time; to increase the accuracy of the documents considered as relevant; to find out documents related to already relevant regulatory compliance documents; etc.)

Processing	• Analyse a document in a short period. Strategy #2 has the constraint of the competence since they are also in the customer side • The analysis and the summary of a document in a short period in Strategy #1. They cannot afford to spend hours to provide a summary for tagging.

Question 9: Do your customers/entity face specific difficulties getting certain tasks done? (For instance, to access and share relevant documents; to monitor changes in regulatory compliance documents; to determine time references within documents; etc.)

	• Problems are mentioned during the rest of the questions

Question 10: Which specific features would your customers/entity enjoy?

Translation	• Accurate translation of legal documents (for instance, the Brazilian civil code)
Criterion change	• Detect that a specific sentence means a change in previous case law for a specific scenario

Question 11: What would make your customers'/entity's day to day workload easier?

Translation	• Accurate translation of legislation pieces
Summary	• Accurate creation of summaries for the later enrichment of the documents
Criterion change	• Detect that a specific sentence means a change in previous case law for a specific scenario

Question 12: What level of performance do your customers/entity expect? (For instance, a 100% accuracy when setting relevant documents in their scenario; a high-speed access to relevant documents; a high-speed update process when a change in regulatory compliance occurs; etc.)

Translation	• An accurate translation is always required

Time	• Strategy #1 has to provide and tag the summary in a short period of time with the maximum accuracy • Strategy #2 has to provide an analysis of legislation, case law, etc. in a short period of time for lawyers and customers

Table 28. Legal Advisor focus group summary.

Interview participant data
IP.LA.1 (Commercial Law)
Legal Advisor (Law Firm and Large Enterprise); Spain
Coordinator:
UAB

Question 1: Strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents

Strategy	• For the search of relevant regulatory compliance documents, the respondent uses different commercial tools • When relevant, the respondent downloads the documents to his/her PC's desktop • For sharing and storing, the company uses a commercial document management system • No specific strategy for the analysis and processing of documents is mentioned
IT Tools	• For searching relevant documents, the commercial soluations are: Aranzadi, Vlex and Wolters Kluwer. For these commercial solutions, services are independent and have different fees. So, they have contracted just some of them. • The document management system for sharing among team members is NetDocuments
Law firms network	• The company is integrated in a Law firm network across Europe. Therefore, when an action has to take place abroad, they contact a local law firm within this network and the actions are completely transferred to the local law firm
IT Tools performance	• They do not know how these IT tools work, but for instance, Aranzadi, provides the searched document within the first two or three results.

Question 2: When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, How do you monitor this change?

Knowledge Management Team	• The company has a Knowledge Management Team (KMT) that monitors changes in legal documents

Question 3: When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, how do you carry out the updating process?

KMT Strategy	• The KMT is subscribed to different services. The analysis and processing of the information is processed with no assistance from any IT tool (manual).
	• The team is completely dedicated to this task
	• This team creates a report every month with the changes occurred for every topic
	• Critical changes are reported as soon as possible (independently from the month report), however, it is unlikely

Question 4: How do you provide your customers/entity's responsible with the relevant regulatory compliance in their scenario? (For instance, I send the whole relevant documents; I send a summary of selected relevant documents; I manually/automatically define a set of simplified rules to follow; etc.)

Output	• The law firm takes care of the whole process. Thus, the customer receives no documents.

Question 5: How do your customers/entity define "too costly"?

Overall	• There scenarios for every point in the list (time, money, efforts).
	• Different countries have different fees for the same services
	• English is the common language when communicating with the network legal firms in Europe

Question 6: Which features are your customers/entity missing?

Translation of legal documents	• When they work with documents in a language different from Spanish, a disclaimer is included. This disclaimer declines accountability for further misunderstandings with language and interpretation of different legal systems than the Spanish one.
List of services	• The services listed in the survey are interesting to offer an added value service to their potential customers.
	• However, at the end of the process, they think that the customer must go to a legal firm

Question 7: Are there performance issues or malfunctions your customers/entity cite? (For instance, the collecting of regulatory compliance documents takes a lot of time/resources; the collected documents are not related enough to their scenario, the access to the collected documents is slow, etc.)

IT Tools performance	• They do not know how these IT tools work, but for instance, Aranzadi, provides the searched document within the first two or three results.

Question 8: What are the main general challenges your customers/entity encounter? (For instance, to analyse a vast number of documents in a short period of time; to increase the accuracy of the documents considered as relevant; to find out documents related to already relevant regulatory compliance documents; etc.)

	• Not applicable

Question 9: Do your customers/entity face specific difficulties getting certain tasks done? (For instance, to access and share relevant documents; to monitor changes in regulatory compliance documents; to determine time references within documents; etc.)

	• No relevant answer

Question 10: Which specific features do your customers/entity would enjoy?

Added value service	• Summarization of digital regulatory compliance documents based on their content • Alert of changes in digital regulatory compliance documents • Topic classification of digital regulatory compliance documents • Linking of digital regulatory compliance documents based on their content

Question 11: What would make your customers'/entity's day to day workload easier?

IT Tools performance	• For the search of specific documents, the IT tools provides a lot of useless documents or none. For instance: "agency contracts + word + word"

Question 12: What level of performance do your customers/entity expect? (For instance, a 100% accuracy when setting relevant documents in their scenario; a high-speed access to relevant documents; a high-speed update process when a change in regulatory compliance occurs; etc.)

	• Not applicable

Table 29. Commercial Law expert interview.

ANNEX VI – ENGLISH SUMMARY TEMPLATE

This annex shows the English version of the summary template used for the focus groups or interviews.

Focus groups or interviews participants data
Coordinator

Question 1: Strategy for the search, analysis, processing, monitoring and handling of digital regulatory compliance documents

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 2: When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, How do you monitor this change?

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 3: When a relevant change in regulatory compliance for your customers'/entity's scenario occurs, how do you carry out the updating process?

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 4: How do you provide your customers/entity's responsible with the relevant regulatory compliance in their scenario? (For instance, I send the whole relevant documents; I send a summary of selected relevant documents; I manually/automatically define a set of simplified rules to follow; etc.)

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 5: How do your customers/entity define "too costly"?

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 6: Which features are your customers/entity missing?

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 7: Are there performance issues or malfunctions your customers/entity cite? (For instance, the collecting of regulatory compliance documents takes a lot of time/resources; the collected documents are not related enough to their scenario, the access to the collected documents is slow, etc.)

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 8: What are the main general challenges your customers/entity encounter? (For instance, to analyse a vast number of documents in a short period of time; to increase the accuracy of the documents considered as relevant; to find out documents related to already relevant regulatory compliance documents; etc.)

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 9: Do your customers/entity face specific difficulties getting certain tasks done? (For instance, to access and share relevant documents; to monitor changes in regulatory compliance documents; to determine time references within documents; etc.)

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 10: Which specific features do your customers/entity would enjoy?

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 11: What would make your customers'/entity's day to day workload easier?

Keyword #1	• Argument 1 • Argument …
Keyword #n	• Argument 1 • Argument …

Question 12: What level of performance do your customers/entity expect? (For instance, a 100% accuracy when setting relevant documents in their scenario; a high-speed access to relevant documents; a high-speed update process when a change in regulatory compliance occurs; etc.)

Keyword #1	• Argument 1 • Argument …

Keyword #n	- Argument 1 - Argument …

Table 30. English summary template for the interviews and focus groups.

ANNEX VII – ETHICAL GUIDELINES FOR THE PROCESSING OF DATA IN THE CONTEXT OF THE KNOWLEDGE ACQUISITION ACTIVITIES CARRIED OUT OUTSIDE THE CONSORTIUM

Inform Consent in the context of the knowledge acquisition activities

In the context of the knowledge acquisition activities of the Lynx Project it is important to differentiate between two dimensions of informed consent: i) the consent of the subject to participate in the research; ii) and, the consent in terms of collecting personal data from a data subject. In the former dimension informed consent can be defined as "meant to guarantee the voluntary participation in research" [European Commission2013, p. 15] while, in the context of personal data, informed consent acts as the key element for lawful processing, as per article 6 of Regulation 2016/679.[5]

However, the two dimensions of informed consent share a common aspect, i.e the importance of providing the research participants with all the information needed to make a truly informed decision, prior to the performance of the research activity. There is not a *numerus clausus* list on the elements that such information should contain. However, common agreement has been reached as per minimum standard that includes "any significant risks, the purpose of the research, any financial interests (e.g. do they receive a fee for each person recruited?), and the source of any external research funding (because people might, for example, object to helping certain companies or governments) [European Commission 2010, p. 3].

Within the Research Ethics domain, informed consent has become associated with the concept of confidentiality. Particularly, it has been conceptualized as a strategy to preserve confidentiality together with the concept of anonymization [European Commission2010]. As such, some authors understand anonymization only as a strategy to achieve confidentiality [Traianou2014. This defines confidentiality as a fundamental ethical principle that operates in a preventive way in relation to data within the research context (Hammetrsley & Traianou2012]. Therefore, confidentiality and anonymization are strongly related to privacy and data protection rights. Although confidentiality stems directly from the respect for privacy, in its conceptualization in the legal technical domain, and from a contemporary perspective, it implies: (i) on the one hand, preventing others from gathering information about ourselves that we do not want to share; (ii) and on the other, maintaining control over the processing of this information related to ourselves [European Commission, 2010].

In this regard, it is worth noting in Social Sciences research that the concept and procedures to obtain the informed consent of participants is a matter of good academic practices, a mechanism to guarantee confidentiality and voluntary participation [Lie & Witteveen2015]. As a result, formal matters related to informed consent have monopolized the debate, instead of focusing on other key elements, such as the right to be informed, or transparency in obtaining the consent. We deem this formal approach detrimental to the understanding of informed consent as a key element in achieving confidentiality and the protection of the right to privacy. In fact, the many different approaches coming from quite diverse ethical frameworks[6] alone do not seem able to provide practical solutions to researchers in order to cope with quantitative and, particularly, qualitative research demands in Social Sciences. These approaches do not solve the main problems related to informed consent, which can be summarized as follows: (i) full understanding cannot be taken for granted (giving formal information to research participants does not

---

[5] Regulation 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).

[6]This refers to the most common approaches to ethical decision-making such as Consequentialism, Duty-Based Ethics, Virtue Ethics, Ethics of Care, Discourse Ethics, Principlism or "The Four Principles Approach", Liberalism, and Communitarianism. It should not be interpreted as a dismissal of ethics. On the contrary, ethical frameworks are essential to design-driven regulations.

necessarily mean that they understand it) ; (ii) hence, there is a the need to rethink formal requirements and the management of informed consent in digital environments[7] [Miller & Boulton2007].

The former statement should not be interpreted as a dismissal of ethics. On the contrary, we deem ethical frameworks essential to design-driven regulations. Nevertheless, its is our contention that the implementation of values and principle requires the interactive communication between participants and researchers during the whole research lifecycle to properly address their content and qualify the rules and procedures in place. This means a *personalization* of the research objectives and goals.As for the procedure for obtaining consent, in terms of collecting personal data, Regulation 2016/679 defines consent of the data subject in article 4(11): "any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her." [8] Moreover, article 13 of the Regulation requires that controller provides the data subject with information on: (i) the identity and contact details of the controller and relevant authorities,(ii) the purpose and legal basis for the processing, (iii) the recipients of personal data, (iv) the retention period, (v) the rights to access, rectify or erase the data, (vi) the right to withdraw,(vii) the right to complain before a relevant authority and (viii) if applicable, personal data will be submitted to automated decision-making.

From a practical perspective in a given research context, applying this technical and legal conceptualization of privacy and data protection rights entails that: (i) researchers cannot take actions that may affect privacy; (ii) awareness that research interventions could affect privacy at any time of the research process; (iii) issues related to privacy and data protection, and those strongly linked with it such as confidentiality and anonymization, cannot be reduced to the achievement of technical and legal compliance with the legal and technical requirements that might be at stake [Punch2013] [Casanovas 2015].

In this regard and particularly *within the context of the Lynx Project, informed consent should then be understood as a complex process in which researchers need to focus on providing, sharing and managing questions and concerns that may arise during the development of the research instead of focusing solely on gathering a written consent form signed by the research participants.*

Specifically, and in line with contemporary approaches to informed consent, *it will be considered as an ongoing decision-making process that entails two different elements: first, informed consent as a document and secondly, informed consent as a process.* As a document, informed consent should guarantee legal provision, according to the legal requirements stated by Regulation 2016/679. However, informed consent understood as an ongoing process has to do with the action of providing information to research participants by the researchers, at any time, and in any step of the project lifecycle to guarantee informed decisions related to the research. Actually, this dynamic and flexible conceptualization of informed consent needs to be put in place with the aim of tackling the ethical concerns raised by qualitative research to preserve privacy and enhancing personal data protection. In fact, following this approach may help researchers to decouple informed consent not only from procedural and formal issues in terms of good research practices, but to understand consent as a complex

---

[7]In that sense, research proposals regarding inform consent in the Visual Research Methods domain (Lie & Witteveen, 2015) provide alternative, innovative and insightful ways of dealing with consent in a digital environment. The concept of Visual Inform Concept is presented as a way of replacing the paper-based informed consent procedure in research circumstances where data are collected and used visually. Procedures to obtaining a filming consent focus on the right to be informed and transparency in terms of gathering the consent instead of the understanding of informed consent as a formal procedure. This kind of proposal entails a paradigm shift regarding consentt. An informed consent paradigm of respect of autonomy and self-determination grounded solely on paper-based procedures needs to be replaced by another paradigm including alternative and complementary values such as reciprocity, universality, and solidarity, among others.

[8] Article 29 Data Protection Working Party has adopted on 10 April 2018 a set of guidelines for providing practical guidelines to ensure compliance with the GDPR on consent: Article 29 Working Party Guidelines on consent under Regulation 2016/679, adopted on 28 November 2017 as last revised and adopted on 10 April 2018. Available at: http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051

decision-making process in which researchers should guarantee that any research participant is in the best position to make informed decisions.

Processing personal data for scientific purposes

The processing of personal data for scientific purposes is specifically addressed in Article 89 (1) of the General Data Protection Regulation. This article, in line with Recital 156, states that:

"Processing for archiving purposes … scientific or historical research purposes…, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organizational measures are in place in particular in order to ensure respect for the principle of data minimization. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner…"

The data minimization principle referred to in this Article 89 is defined in Article 5 (1c) with the following wording:

"Personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed".

So, as it will be explained in detail in the following sections, this principle will work as a minimum legal constraint within the research framework of the Lynx Project for both quantitative and qualitative techniques conducted during the knowledge acquisition process. For instance, and due to the nature of the research that the Lynx Project entails, special categories of personal data may emerge,[9] especially when carrying out particularly qualitative interviews or focus groups. Even though the umbrella of the scientific purpose covers the processing of such data, the data minimization principle is going to be used as a precautionary principle.

The following sections explain in detail the concrete measures to be put in place-, for each of the knowledge acquisition activities- by any researcher within the Lynx project when conducting such activities.

Ethical guidelines for conducting online surveys outside of the Consortium

The Lynx project includes the deployment of an online survey according to Task 1.1. The purpose of this survey is that of gathering information on industry requirements (business requirements). This questionnaire will be implemented using the survey tool Google Forms and will be anonymous. In this regard, no personal data from the participants will be collected. Against this background, Article 26 of the Regulation 2016/679 becomes relevant, as it states that data protection legal requirement does not apply to anonymous information, and defines this concept as "information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes."

Although the online survey does not foresee the collection of personal data, in order to conduct an ethical research in terms of conducting knowledge acquisition techniques, the participants must be provided by relevant information, as stated in the introduction to these guidelines. Bearing in mind the electronic

---

[9]According to Regulation 2016/679, processing of personal data revealing racial or ethnic origin, political opinions, religious and philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural's person sex life or sexual orientation shall be prohibited. The Regulation includes exceptions to this general prohibition when: (i) the data subject has given explicit consent; (ii) processing relates to personal data made manifestly public by the data subject; (iii) for reasons of public interest; (iv) and, when processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.

nature of the survey a legal notice, with a specific section for gathering the consent of the participants, will be included in the survey (See Annex X). This legal notice includes:

- Information on the project, the Consortium, and the source of the funding.
- Privacy policy: information collected, purpose, confidentiality clause, data security measures, and retention period.
- Rights of the participant: voluntary nature of participation, right to withdraw, right to access, and right to rectify or erase.
- Disclaimer: limited scope of the questions, not necessarily addressing all particular circumstances.

In order to verify that the online survey is conducted in compliance with the above-mentioned requirements for ethical research some procedures need to be implemented by all partners of the Lynx Consortium. In particular:

1. UPM, as coordinator of the Lynx project, will act as controller of the information gathered.
2. Those partners conducting knowledge acquisition tasks in the project are responsible for applying these ethical guidelines.
3. The UAB-IDT team will validate the final version of the legal notice, and the questionnaire for the online survey, in ethical terms, prior to its transmission to any potential participant.
4. The UAB-IDT team will address any query or petition received from any participant in the online survey in relation to the access and withdrawal rights.
5. UAB-IDT team will address any query or doubt on how to conduct ethical research.

Ethical guidelines for conducting interviews, focus groups and workshops outside of the Consortium

The Lynx Project foresees the development of a set of interviews, focus groups and workshops outside of the Consortium within the Task 1.1. Due to the nature of these activities the double nature of consent appears again as both personal data, and potential sensitive information will be collected. Therefore, two issues become crucial from the ethical perspective: the confidentiality of the information, and the anonymisation of personal data. The Code of Ethics of the International Sociological Association reminds to researchers: "The security, anonymity and privacy of research subjects and informants should be respected rigorously... The sources of personal information obtained by researchers should be kept confidential, unless the informants have asked or agreed to be cited. Should informants be easily identifiable, researchers should remind them explicitly of the consequences that may follow from the publication of the research data and outcomes." (International Sociological Association, 2001, p. 2.3). From this article we can extract some general rules that researchers must apply when designing and conducting their research:

- The information gathered from the participants should be kept confidential, unless specific consent to be cited is given by the participant.
- The information gathered should be anonymised, and used only for the purpose for which it was collected.
- Participants must be informed when the researcher believes that some of the information shared may make them identifiable, and the potential consequences.
- Participants must be given, in a clear and transparent manner, the opportunity to withdraw at any time, and especially after being informed of their potential identification and consequences.

Since the conduction of knowledge acquisition techniques entails also the processing of personal data, data protection principles and legal requirements extracted from Regulation 2016/679 must be taken into consideration. In particular the controller needs to put in practice organizational and technical measures addressed to "minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data

subject to monitor the data processing" (Regulation 2016/679, Recital 78). Such measures, within the Lynx Project are exposed below:

1. Information collected from the participants will be **anonymised**. Each of the partners of the Consortium that conducts interviews, focus groups and workshops carried out outside of the Consortium will prepare a summary, in English, of the results of the knowledge acquisition activities. The raw information will be kept in local resources by the partners under their own responsibility, and according to the data protection policies of their own organisations. Partners should pay special attention to the respect of the minimisation principle following article 89 (1) of Regulation 2016/679.

2. Each task leader will collect the **English summaries** and prior to share them within the Consortium a check review process should be conducted with the aim of ensuring that no personal or sensitive information is contained in the summary, unless the participant has given specific consent. Once this point is verified the summary can be shared within the Consortium.

3. Researchers must obtain **specific consent** from all the participants prior to their involvement in the different activities.

4. Consent must be specific for each activity. The UAB-IDT team has created a template **Electronic Participant Consent Form**. (Annex I).

5. **Informed consent** must be obtained, as a general rule, in a **writing** form.

6. Informed consent might be obtained through an **electronic Informed Consent Form**. This requires the implementation of some technical measures that ensure the identity of the participant, such as electronic signature. The check bottom system cannot be considered as signature of the consent for personal data processing.

7. **Oral informed consent** is highly discouraged. Although oral consent is legally valid, the controller of the data must be able to "demonstrate that the data subject has consented to processing of his or her personal data" (Regulation 2016/679, article 7.1). Therefore, researchers should only use this procedure when there is no other possibility and after having consulted with UAB-IDT team. They will oversee evaluating the situation, bearing in mind the potential value of the information that could be obtained from the participant. Annex III contains a model for a suggested script with the minimum content that the researcher must register (audio or video) in case oral consent is authorised by the IDT-UAB team.

REFERENCES

[Casanovas2015] Pompeu Casanovas. Semantic Web Regulatory Models: Why Ethics Matter. Philosophy & Technology, 28(1), pp. 33-55.

[Casanovas, Palmirani, Peroni, van Engers, Vitali2016] Pompeu Casanovas, Monica Palmirani, Silvio Peroni, Tom van Engers, and Fabio Vitali. Semantic web for the legal domain: the next step. Semantic Web 7, no. 3 (2016): 213-227.

[Casanovas, Gonzalez-Conejero & de Koker2018] Pompeu Casanovas, Jorge Gonzalez-Conejero, Louis de Koker. Legal Compliance by Design (LCbD) and through Design (LCtD): Preliminary Survey. TERECOM@JURIX 2017: 33-49.

[European Commission2010] European Textbook on Ethics in Research. Luxembourg: Publications Office of the European Union, 2010

[European Commission2013] Ethics for Researchers. Publications Office of the European Union, Luxembourg, 2013.

[Flinders1997] David J. Flinders. InterViews: An introduction to qualitative research interviewing: Steinar Kvale. Thousand Oaks, CA: Sage Publications, 1997, pp. 287-288.

[Gordon, Governatori & Rotolo2009]. Thomas Gordon, Guido Governatori, and Antonino Rotolo. Rules and norms: Requirements for rule interchange languages in the legal domain. In International Workshop on Rules and Rule Markup Languages for the Semantic Web, pp. 282-296. Springer, Berlin, Heidelberg, 2009.

[Hammersley & Trianou2012]. Martyn Hammersley and Anna Traianou. Ethics in Qualitative Research: Controversies and Context. Sage, London, 2012.

[Harvey2011] William S. Harvey. Strategies for Conducting Elite Interviews. Qualitative Research 11(4), pp. 431-441.

[Lie & Witteveen2015] Rico Lie and Loes Witteveen. Visual informed consent: Informed consent without forms. *International Journal of Social Research Methodology* 20, no. 1 (2017): 63-75.

[Miller & Boulton2007] Tina Miller and Mary Boulton. Changing constructions of informed consent: Qualitative research and complex social worlds. Social science & Medicine 65, no. 11 (2007): 2199-2211.

[Motta2013] Enrico Motta. 25 years of knowledge acquisition. International Journal of Human-Computer Studies 71 (2): 131-134.

[Punch2013]. Keith F. Punch. Introduction to Social Research: Quantitative and Qualitative Approaches. Sage, London, 2013.

[Regulation 2016/679] Regulation 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).

[Rodriguez-Doncel, Casanovas, Gonzalez-Conejero2017] Víctor Rodríguez-Doncel, Pompeu Casanovas, Jorge González-Conejero: Proceedings of the 1st Workshop on Technologies for Regulatory Compliance co-located with the 30th International Conference on Legal Knowledge and

Information Systems (JURIX 2017), Luxembourg, December 13, 2017. CEUR Workshop Proceedings 2049, CEUR-WS.org 2018 http://ceur-ws.org/Vol-2049/

[Rubin & Rubin2011] Herbert J. Rubin and Irene S. Rubin. Qualitative interviewing: The art of hearing data. Sage, 2011.

[Traianou2014] Anna Traianou. The Centrality of Ethics in Qualitative Research. In P. Leavy, The Oxford Handbook of Qualitative Research Methods, Oxford University Press, Oxford, 2014, pp. 62-79.

[Turner2010] Daniel W. Turner III. Qualitative Interview Design: A Practical Guide for Novice Investigators. The Qualitative Report, 15(3) pp. 754-760.

[Van Lamsweerde2009] Axel van Lamsweerde. Requirements Engineering. From Systems Goals to UML Models to Software Specifications. Wiley & Sons, Chichester, 2009.