

# Analysing Legal Information Requirements for Public Policy Making

Charalampos Alexopoulos<sup>1</sup>, Shefali Virkar<sup>2</sup>[0000-0003-0701-7684], Michalis Avgerinos Loutsaris<sup>1</sup>, Anna-Sophie Novak<sup>2</sup>, Euripidis Loukis<sup>1</sup>

<sup>1</sup> University of the Aegean, Greece

{alexop,mloutsaris,eloukis}@aegean.gr

<sup>2</sup> Danube University Krems, Austria

{shefali.virkar,anna-sophie.novak}@donau-uni.ac.at

**Abstract.** Most of the research that has been conducted in the area of legal informatics concerns its ‘supply side’, dealing with the development of effective systems for legal information provision. However, limited research has been conducted on the ‘demand side’ of legal information provision, though it is absolutely necessary to gain a good understanding of it, in order to design effective and useful systems for the provision of legal information; furthermore, this limited research is dealing with the legal information needs of the lawyers, and neglects the ones of other important groups. This paper contributes to filling this research gap. It analyses legal information requirements of a highly important for the society group: the designers of public policies. Initially we investigate current legal information sources and systems used by public policy makers, as well as their relevant search practices. Then we investigate their business needs for additional capabilities/functionalities for a better support of their policymaking activities using advanced legal analytics tools and services. Finally, we discuss the information, processing and technical requirements for the development of a legal information system providing the above advanced functionalities and services. For the above purposes, we have collected data through semi-structured interviews from 13 Greek and 7 Austrian public administrators dealing with the design of public policies, which lead to interesting and useful insights, as well as a novel set of additional advanced capabilities and functionalities that can give rise to a new generation of legal informatics.

**Keywords:** legal information, legal informatics, public policy, decision support systems, service provision.

## 1 Introduction

Recent trends in digitalization, open data, and social media have resulted in an exponential increase in the amount of data available for use by public policy makers in order to make sense of the socio-economic and political phenomena, and design relevant public policies [1]. Repositories of large quantities of novel types of information – including expert knowledge, sensor data, text, social media posts - have become available to

policy makers. An important part of this information, which is highly useful and important for policy makers, is legal information, concerning existing or previous relevant legislation, both of their country as well as other countries, and also European legislation. Advanced intelligent systems, together with sophisticated techniques of data harvesting, annotation, analysis and visualisation have enhanced our ability to understand and make sense of extensive and complex relevant information to policy makers.

Complex decision-making based on the profound analysis of societal problems and possible solutions using these large quantities of available data is a prominent aspect and target of evidence-based policymaking. However, policy makers are currently confronted with the challenge of accessing vast, hitherto untapped, sources of information in an efficient manner that provides all relevant information, separates ‘noise’ from ‘signal’, and assists and supports them substantially for designing effective public policies. Furthermore, policy makers are also not equipped with the skills and technical know-how necessary to integrate and process all relevant information, including the most current data, from various sources, and elicit meaning from it, in order to make informed policy decisions. The solution to the above problems is the development of advanced decision support systems that exploit available big data in order to facilitate the cognitive activity involved in the structuring of public policy decision situations, the design of policy options, the enumeration of alternative courses of action, and the evaluation of these alternatives leading to a policy decision. Until recently, these tools have been time-consuming to deploy, and also not user-friendly for decision makers, and frequently resulting in models that do not reflect ‘real-life’ policy realities [2].

This paper presents research conducted towards addressing the above challenges, for one of the most important and at the same time difficult to manage kinds of information needed for supporting policy making, the legal information, as part of the European ‘ManyLaws’ project [3]. The objective of this project is to develop advanced decision support tools and services for policy actors built on a robust foundation of legal information search retrieval. It aims to offer users a suite of targeted services to support policy making through the provision of advanced legal information, built upon semantic analysis techniques, text mining tools, and in general advanced processing technologies. For this purpose, it is necessary to examine critically how decision-making activities within the policymaking process might be supported through improved legal information search and retrieval capabilities.

However, these requirements have not been sufficiently researched. As explained in more detail in section 2.2 most of the research that has been conducted in the area of legal informatics concerns its ‘supply side’, dealing with the development of effective systems for legal information provision, with advanced search and processing capabilities, based on appropriate metadata as well as organization and annotation of large quantities of textual legal information. On the contrary, limited research has been conducted on the ‘demand side’ of legal information provision, though it is absolutely necessary to gain a good understanding of it, in order to design effective and useful systems for the provision of legal information. Furthermore, as concluded in section 2.2, this limited research is dealing with the legal information needs of the lawyers and neglects the ones of other important groups. Therefore, this paper contributes to filling this important research gap. It analyses legal information requirements of a highly important

for the society group: the designers of public policies. In particular, we investigate the following research questions:

- RQ1. What are the current legal information sources and systems used by public policy makers, as well as their relevant search practices?
- RQ2. What are the business needs of policy makers needs for additional capabilities/functionalities for having a better support of their policymaking activities using advanced legal analytics tools and services?
- RQ3. What are the information, processing and technical requirements for the development of an information system providing the above advanced functionalities and services?

Our study has been based on the collection of data through interviews with 13 Greek and 7 Austrian public administrators dealing with the design of public policies.

This paper is structured in six sections. The following section 2 critically discusses decision making and decision support tools within the context of public policy creation, as well as legal information provision and analytics. Next, section 3 outlines the methodological approach we adopted in order to collect primary data. Section 4 then presents the results of the abovementioned in-depth expert interviews with policymakers from Greece and Austria. Emerging themes and issues are analysed in section 5, and finally conclusions and recommendations for future research are outlined in section 6.

## **2 Background**

### **2.1 Decision Support Systems and the Public Policy Cycle**

Decision support systems (DSS) constitute a class of advanced computer systems comprising a collection of software applications and tools developed in order to facilitate managerial decision making and improve the quality of the decisions being taken - particularly under conditions of uncertainty, initially in the private sector, but later in the public sector as well [4]. These systems facilitate the co-ordination of data delivery and the development of data consistency, aid in data trend analysis as well as use for making forecasts, fulfilling users' data requirements, and supporting the quantification of uncertainty, as well as recommending courses of action [5]. [6] identifies seven different types of DSS applications, based on contemporary professional practice and actor-base: Personal Decision Support Systems, Group Support Systems, Negotiation Support Systems, Intelligent Decision Support Systems, Knowledge Management-based DSS, Data Warehousing, and Enterprise Reporting and Analysis Systems. Furthermore, they argue that among them the personal decision support systems, data warehousing, and enterprise reporting and analysis systems are the most widely available and used systems in day-to-day contexts. [5] identify a number of Artificial Intelligence paradigms that can be used in order to mimic complex human problem-solving behaviour that is otherwise difficult to describe mathematically using conventional programming methods (using symbolic logic, Artificial Neural Networks (ANNs), fuzzy systems, evolutionary computing, Intelligent Agents, and probabilistic reasoning models, etc.).

In order to fully understand the implications of developing a decision support system to facilitate the formulation of effective public policy, which is the main target of our

research, it is important to understand the nature of the policy-making process, and the main practical challenges that the creation of public policy presents. According to [7], the public policy process includes one or more cycles, each of them including the following stages: agenda setting, policy formulation, decision-making, policy implementation, policy evaluation and finally policy improvement/maintenance or termination. Implicit in this conceptualisation is the progression of the different stages in a distinct chronological order: to begin with, problems are defined and placed on the agenda; next policies are developed, adopted and implemented; and finally, select policies are evaluated and either terminated or pursued further [8]. The cyclical framing of the policy-making process underlines the manner in which a feedback loop is created between its various inputs and outputs, punctuated by decision points, resulting in a policy outcome [9]. The legislation-oriented DSS under development in this research project aims to support all the above stages of policy of policy making with respect to legal information provision, based on advanced techniques concerning how the legal information is handled, processes, semantically annotated, presented and accessed.

## **2.2 Legal Information Provision and Analytics**

Legal Informatics refer to the application of Information and Communication Technologies (ICT) within the context of legal environment [10], and is defined by [11] as the "...theory and practice of computable law, i.e. the cooperation between humans and machines in legal problem-solving". This area focuses on the opportunities and challenges that the exploitation of ICT in the legal system faces, and thus involves all related organizations and legal information users in the legal domain. One of these challenges lies in the supply of legal services, which are currently under-consumed by individuals and companies [12]. Therefore, the latest advancements in the legal informatics are targeted towards making services more open and promoting access to legal resources.

Furthermore, accurate and timely legal information is an essential component of effective decision-making by several different societal actors. However, the ability of citizens, businesses, public servants as well as politicians and their advisers to easily access, fully comprehend and apply complex legal information to their everyday contexts often hinges on an advanced understanding of governmental procedures, legal language, and the law itself. Unfortunately, usually this does not happen, and most people struggle and have difficulties to locate the legal artefacts they need. On a practical level, two immediate problems can be identified. The first pertains to the quantity of legal information currently available online, as a direct consequence of the increasing complexity of the European legal system, coupled with advances in digital technologies, cloud storage capabilities and the Open Data movement. The second problem concerns users' ability to comprehend legal information, as well as the singular nature of legal jargon, wherein most individuals do not possess the specialist legal education and practical knowledge required to grasp complex legal terminology or follow developments in legislative processes. Both these phenomena have resulted in the burgeoning popularity of legal analytics, or the application of big data analysis methods within the field of the law [13], which has, within a very short span of time, moved from the margins of the legal profession into the mainstream.

Recently, an important trend in the area of legal informatics is the increasing exploitation of artificial intelligence technologies, as new legal applications based on text and natural language processing and machine learning have recently emerged, which seem to influence significantly and gradually transform the practice of the law [14]. These tools and services, support conceptual legal information retrieval and predictive legal analysis by connecting computational models of legal reasoning directly with legal text. The legal field, however, remains one of the most difficult domains for the application of automated text retrieval, as legal text retrieval is based primarily upon concepts, and not on the explicit wording in the document texts [15]. [16] argue that this is because legal concepts are not discrete, but instead are situated along "...a dynamic continuum between common sense terms, specific technical use, and professional knowledge, in an evolving institutional reality." Moreover, legal text retrieval has traditionally relied upon external knowledge sources such as thesauri and classification schemes, and the manual indexing of documents [17].

However, the research that has been conducted in the area of legal informatics concerns mainly its 'supply side', dealing with the development of effective systems for legal information provision, with advanced search and processing capabilities, based on appropriate metadata as well as organization and annotation of large quantities of textual legal information. On the contrary, limited research has been conducted on the 'demand side' of legal information provision, though it is absolutely necessary to gain a good understanding of it, in order to design effective and useful systems for the provision of legal information; furthermore, this limited research is dealing with the legal information needs of the lawyers, and to a much lower extent of the citizens (general public), however it neglects the ones of the policy makers, though they rely heavily on a wide range of legal information, and they are quite important for the society. Some early papers have focused on the demand for paper-based legal information sources. [18] investigates the legal information needs of the general public (citizens who are not lawyers), as well as the sources (paper-based ones) they use in order to fulfil these needs, and the role of the general and law libraries, and also of legal aid centres, on this. In [19] are investigated the legal information needs of lawyers, the main purposes they are searching legal information for, the types of information required, and also sources and ways (paper-based ones) for meeting these needs; they conclude that only large law firms have extensive legal libraries and therefore sufficient access to legal information, while this does not hold for smaller law firms: for them the only practical solution, due to the inherently large volume of legal information required is co-operation among such firms, or use of courts' legal libraries. Furthermore, there is some subsequent research focusing on the demand for electronic legal information sources. [20] and [21] examine existing online sources of legal information from users' perspectives, however focusing mainly on their usability, and secondarily on the types of legal content they offer, aiming to identify problems that reduce accessibility and effective use by the users. In [22] are investigated the legal information needs of law faculty for their teaching and research work, as well as the paper-based and online sources used by them, their relevant perceptions and also their computing skills. Recently [23] examines the perceptions of academic and practicing lawyers about existing online legal information resources, as well as the degree of satisfaction with them, the barriers to their effective use, and also

relevant requirements for the design of better legal information systems. Therefore, since the existing previous literature concerning the ‘demand side’ (users) of legal informatics is limited, and focusing mainly on lawyers, further research is required for understanding the usage and search behaviour of current users, as well as the needs of existing and potential users, investigating different user groups; furthermore, it is necessary to conduct ‘innovation-oriented’ research in order to identify additional novel advanced capabilities/functionalities that can provide higher levels of relevant decision support and lead to a new generation of legal informatics. Our paper makes a contribution in this direction.

### 3 Methodological Approach

Fundamental to the successful construction of an effective decision support system is the elicitation of user requirements, and the subsequent accurate definition of hardware and software specifications [24, 25]. [26] describes a framework that enables the integration of agile design methods with user-focused design approaches, showing that agile user-centric design focuses on an iterative and rigorous collaborative development process [27]. Important techniques for this phase of requirements engineering include the questionnaire, interviews, use cases/scenarios, observation and social analysis, focus groups, and brainstorming [28, 29].

One of the most commonly applied techniques for requirements elicitation, the expert in-depth interview is defined by [30] as “...a one-to-one method of data collection that involves an interviewer and interviewee discussing specific topics in depth.” Typically, in-depth interviews are used in research to seek information on individual personal experiences pertaining to a specific issue or topic [31]. In the context of user requirements elicitation, in-depth interviews may be thought of as ‘conversations with a purpose’ during which stakeholders and/or domain experts are questioned to elicit information about their current attitudes, patterns of behaviour, modes of practice, and needs or requirements in relation to the new system or application. In-depth interviews are considered as an efficient way to collect large amounts of uniform data quickly [32, 28]. Following [28], this technique was selected as a data collection instrument that enabled researchers to collect rich qualitative data from interview subjects that reflected their experiences with legal information search and retrieval, as well as their expectations, ideas, and opinions.

Within the context of this study, the authors conducted a series of *in-depth, semi-structured* interviews in Greece and Austria in early 2019. The purpose of the interviews was to investigate patterns of legal information access and use amongst public administrators in the two pilot countries, as well as to gain deeper insight into current policymaking and legislative methods and practices identifying specific user needs and business requirements. In total 20 semi-structured interviews were conducted. Thirteen of them (four male and nine female subjects) were individuals representing different functions and levels of Greek public administration between 28 March and 30 April 2019. Seven of them (three male and four female participants) were members of Austrian public administration and/or from private legal practice representing different

functions between 15 March and 15 April 2019. Ages ranged between 22 years and 64 years old, with further details withheld from being reported to preserve anonymity. Interviews were facilitated remotely, via a digital communications platform, as well as in person, and the working language was English. All interview transcripts have subsequently been codified and analysed.

**Table 1.** Expert Interview Guide

<b>Section / Main Topic</b>	<b>Purpose</b>
<i>Introduction</i>	Provision of global context - introduction to interviewee by local co-ordinator, establishment of background to proposed project, information about the purpose of the interview.
A. Demographic Questions	Basic questions to establish demographic context of interviewee, including familiarity with information technology.
B. Policy Creation - Best Practices, Methods, and Systems	Questions about current policy making and/or legislative processes within national context, step-by-step procedure from interviewee perspective including current role, nature of collaboration between different actors, current information technology systems used.
C. Policy Creation – Requirement for Legal Information, Search Strategies Employed	Questions about interviewee requirement for and use of legal information.
D. Policy Creation - Sources of Data	Further questions on legal information sources used currently used by interviewee, especially concerning preferred sources of legal data.
E. Business Needs Identification	Questions to identify business needs for potential users, including about features missing from current sources of legal information and reasons for use/non-use of databases or platforms.
<i>Summary</i>	Closing statement(s) by interviewer summarising key points arising during interview, time given for participants to clarify or add to previous input.

The interview schedule was derived based on the one used to collect data for the five initial user stories, and comprised of the following top-level topics: *demographic questions; policy creation - best practices, methods and systems; policy creation – requirement for legal information, search strategies employed; policy creation - sources of data; and the identification of business needs*. The results of the in-depth expert interviews with policymakers affiliated to the Hellenic and Austrian parliaments are reported in this research paper and further elucidated upon in the interview guide (see Table 1, *above*).

## **4 Results**

The study reported in this research paper sought to identify and determine the exact nature of the legal information requirements of public policymakers with a view to designing an advanced legal information system that fulfilled these needs. Through a detailed analysis of interview data, researchers sought to assess the current legal information data sources and systems available to their sample population, together with the practices, strategies and tools used to obtain legal information. From this enquiry, it was expected that the business needs of policy makers and legal administrators to support their day-to-day activities could be extracted, and that a better understanding of the underlying technical requirements gained. A summary of the salient points raised during these interviews may be found in the Table 2 below.

### **4.1 Best Practices, Methods and Systems**

In both Greece and Austria, interview participants reported differences in the manner in which policymakers collaborated with legal experts, and also a significant degree of uncertainty as to the extent of these interactions. Both sets of participants emphasised the importance of information – in particular, legal information – as a vital input in policymaking. However, they also highlighted the importance of expert knowledge and judgement in selecting information as an input in the policymaking process. Even in the case of Austria where a very sophisticated system exists, it is managed by hand-written information (data entry by the administrators). The legal systems in both cases do not fully support the administrative processes of handling parliamentary data.

### **4.2 Requirement for Legal Information, Search Strategies Employed**

Interview participants from both countries mentioned their requirement for unfettered access to a wide range of legal information sources, with particular emphasis on digital resources. They discussed the search strategies currently used to locate and retrieve information from these databases and expressed the desire to possess advanced search and retrieval capabilities, including complex string searches and translation features, in order to fully optimise their use. The currently offered functionality is restricted to services of searching by keywords which are not semantically annotated and/or extracted with a way that could illustrate correlations or even show the history of a law and its changes in time. The absence of semantically annotated does not allow to provide useful advanced services, such as the estimation of the degree of transposition of an EU Directive into the national legal system.

### **4.3 Sources of Legal Data**

When questioned about the sources of legal data currently used, and the reasons for their preference, interview respondents from both Greece and Austria were able to name popular sources of this sort of information, and justify their use within individual



professional contexts. Participants from Greece differed from their counterparts in Austria in saying that one reason for their use of particular data sources was organisational endorsement. All participants emphasised the importance of open legal data sources.

#### 4.4 Business Needs Identification

Interview participants were asked to identify those features, tools and services either missing from current legal information retrieval resources, or whose addition or removal would facilitate their professional activities. Both sets of participants noted an absence of collaborative software available to them. They also identified an absence of customisable content – personalised dashboard, individual history, saved or book-marked content – as key functionalities missing from current online legal resources. Participants from Austria believed that they would be able to overcome professional challenges through better access to legal information, and that the quality of policy-making would be thereby enhanced. They differed from their counterparts in Greece through expressing a general ambivalence to subscription fees, even if those were paid for by the organisation.

Even more, the interviewees reported on general capabilities they would like to have from such a DSS: parallel search in multiple EU member-state legal frameworks using simple keywords; the capability to assess the degree of transposition of an EU directive in national legislation; an indication of the national legislation relevant to each directive and the capacity to monitor the status of transpositions; tools to analyse references made to the European legislation within national laws; the ability to make comparative analyses of equivalent or relevant laws from different EU member states and between connected laws from the same member state; the functionality to monitor the progress and/or current status of a specific piece of national or European legislation, including preparatory acts and agreements, over time; a visual timeline analysis tool for all legal elements; the provision of geo-visualisations, text-related visualisations, and other common visual decision aids; visualisations of correlations, dependencies and conflicts between different laws; and other dedicated decision support services, such as impact assessment, for expert users. It is envisioned that these services, when used alone or in combination through the proposed project portal, would enable policymakers to construct accurate models of the legal environment circumscribing the problem under consideration, identify the various possible outcomes based on an in-depth understanding of relevant legal matters, and design policy outcomes that conform with the current legal framework.

**Table 2.** Summary of expert interview results

Topics	Greece	Austria
<b>Best Practices, Methods and Systems</b>	<ul style="list-style-type: none"> <li>• Variance in the Degree of Collaboration between Policymakers and Legal Experts. Great degree of disperse among organisational units.</li> </ul>	<ul style="list-style-type: none"> <li>• List of Most Popular Platforms and Sources of Legal Information Online.</li> </ul>

	<ul style="list-style-type: none"> <li>• Basic Informational Input for Policymaking is both General and Specialised.</li> <li>• Uniform Criteria for Determining the Relevance or Suitability of Information: trusted sources and easiness of understanding.</li> <li>• Experience and Expertise is Key When Assessing, Comparing or Evaluating Legal Information.</li> </ul>	<ul style="list-style-type: none"> <li>• Extent of Collaboration between Legal Experts and Policymakers in Austria is Uncertain.</li> <li>• General Emphasis on Accurate, Timely and Robust Legal Information.</li> <li>• General Characteristics of 'Useful' Legal Information Identified.</li> </ul>
<b>Requirement for Legal Information, Search Strategies Employed</b>	<ul style="list-style-type: none"> <li>• Types Searched-For Legal Information.</li> <li>• Search Strategies Adopted to Obtain Relevant Results.</li> <li>• Further Capability Desired While Searching for Legal Information.</li> <li>• Correlation of legal artefacts.</li> <li>• Comparative analysis of laws.</li> <li>• Transposition of EU directives in the national legal system.</li> </ul>	<ul style="list-style-type: none"> <li>• Wide Range in the Types of Searched-For Legal Information.</li> <li>• Emphasis on Digital Legal Information Sources.</li> <li>• Few Desired Abilities and Wished-for System Functionalities.</li> <li>• Timeline analysis.</li> </ul>
<b>Sources of Data</b>	<ul style="list-style-type: none"> <li>• Current Sources of Legal Information and Reasons for Preference.</li> <li>• Organisational Endorsement of Legal Data Sources.</li> <li>• Mainly Legal Database applies fees, but users prefer to use a legal database based on open data sources.</li> </ul>	<ul style="list-style-type: none"> <li>• Popular Online Sources of Legal Information in Austria.</li> <li>• Reasons for Preference of Currently Used Sources.</li> <li>• No Particular Organisational Endorsement of Legal Information Sources.</li> <li>• General Absence of Subscription Fees.</li> <li>• Importance of Open Legal Data Sources.</li> </ul>
<b>Business Needs Identification</b>	<ul style="list-style-type: none"> <li>• No Collaborative Software Available.</li> <li>• Missing Features and Functionalities from Current Online Sources of Legal Information.</li> <li>• Added-value Features Looked-for in Current Sources of Legal Information.</li> <li>• Willingness to pay a Subscription Fee for a Comprehensive, All-Inclusive Legal Information Resource.</li> <li>• Noteworthy Thoughts or Issues Arising.</li> </ul>	<ul style="list-style-type: none"> <li>• Desired Features Missing from Current Legal Data Sources.</li> <li>• General Ambivalence Towards Subscription Fees.</li> <li>• Professional Challenges May Be Overcome Through Access to Better Legal Data.</li> <li>• Facilitating Policymaking Through Improved Access to Legal Information.</li> <li>• Monitor progress</li> <li>• Visualisation of accumulative results</li> </ul>

## 5 Discussion

In order to create an effective legal information retrieval system which functions like a DSS that assists the policy making process, it is important to apply the design principles of systems integration. As a first step, the key factors involved in the policy process and central to the construction of appropriate tools and services were identified through a combination of desk-based research and in-depth expert interviews, which are:

1. The principle model of the policy decision making process, based on a critical evaluation of the policy cycle framework.
2. The methods adopted for the implementation of its components in each pilot country, from the perspective of the policy actors interviewed.
3. The central decision makers and other primary stakeholders, and the relationships among them.

Furthermore, based on the results presented in section 4 of the research paper, the primary requirements for the proposed decision support system were extracted, and they have been categorised in three groups: Information Requirements, Processing Requirements and Technical Requirements. In the following paragraphs we are discussing a possible solution towards covering the business needs of the participants.

**Information Requirements:** The complexity of the policy making process, captured through the lens of different policy roles and job functions, warrants the regular making of decisions based on accurate and complete legal information. In practice, it was seen that policymakers use both legislative information (information on the current national legal framework, the exact status of specific pieces of legislation, or comparative analyses of two or more pieces of legislation), policy information (general background information, policy reports, data and expert analysis), and political information (public opinion through social media and dedicated e-consultation and e-participation platforms, or of support for a policy among legislators) to support or oppose policy alternatives, or to make evidence-based judgements. The actual routes of information acquisition may include individual policymaker research, policy staff collaborative activity, committee hearings, oversight activities, or interdepartmental transactions.

**Processing Requirements:** In order to cover the information requirements, the proposed solution will structure the legal documents based on legal ontologies such as ELI and AKN. These ontologies are focused on representing information of two different kinds of data: (a) the description of the basic information of a law when it is published, and (b) information for better support the parliamentary procedure. For this purpose, the back-end may utilise various text-mining techniques, such as information extraction by using Regular Expressions, Tokenization, Word clustering, Word stemming, Results filtering, Data cleansing, Word Vector (Term Occurrences), just to name the most important ones [33]. All these techniques will extract all the necessary information that will be used for annotating the data and building the necessary files. Particularly, Regular expressions will separate the body of the legal document in components (Sections, Parts, Chapters, Articles, Paragraphs) and they extract all the correlations that included in each component. The general background information will be represented by keywords that will be extracted by using Tokenization of the Legal Document Body, Filtering the Stopwords, Stemming or Clustering the words (depends on the language of

the legal document) and by creating a word vector of the terms that will be presented in a legal document. The first words, usually, are the most important words of a specific legal document and gives in the policy maker and legislator the opportunity to understand easily the content of this legal document [33]. In addition, the above serialization of the techniques will be used for generating the n-grams of the legal document towards the support of finding the similar legal documents. Finally, social media and governments portals that are used for collaborating of every new legal document among governments and citizens will be used in order to retrieve all related comments. These comments will be processed and semantically analysed to help legislators and policy makers to better design citizen-centric policies and propose policy alternatives.

**Technical Requirements:** The above processing requirements need the usage of a High-Performance Computing (HPC) [34], since the need is to produce n-grams at the degree of 10 (10-grams) as well as to analyse and semantically annotate the legal basis of two countries and the EU along with their associate parliamentary information. Parallelization of resources should be also applied by using computational clusters. This kind of clusters provide the ability to analyse huge amounts of unstructured data in a distributed computing environment. As a result, we will receive the legal documents in a structured and annotated way. Finally, a dedicated translation component, like the e-Translation DSI will translate all these structured data in different languages allowing the comparison of legal texts between different languages.

## 6 Conclusions

The policy making environment in today's digital world is characterised by high complexity, which necessitates access to and processing of large amounts of many different kinds of information, in order to gain a deep understanding of the big problems and challenges of modern societies, and design effective policies for addressing them. One of the most important kinds of required information for modern policy making is definitely the legal information. The comprehension and extraction of meaning from numerous specialised domain-specific legal texts and laws is quite difficult for policy makers who are not legal experts. So, it is quite important to develop advanced legal information provision systems, which can offer substantial assistance and support in the above policy making legislation-related tasks. However, this presupposes a deep understanding of legal information requirements of public policy making.

In this direction in the previous sections of this paper initially we investigated current legal information sources and systems used by public policy makers, as well as their relevant search practices. Then we examined their business needs for additional more advanced capabilities/functionalities that provide a better support of their policymaking activities using advanced legal analytics tools and services. Based on the findings we defined the information, processing and technical requirements for the development of a legal information system providing the above advanced functionalities and services. Our analysis has led to interesting and useful insights, as well as a novel set of additional advanced information provision capabilities and functionalities that can give rise to important innovations in the area of legal informatics, and finally to the emergence

of a new generation of it. They include parallel search in multiple EU member-state legal frameworks using simple keywords, in order to identify and compare national legislations concerning a policy-related topic of interest, and capabilities for finding interconnections as well as conflicts among laws, as well as assessing the degree of transposition of an EU directive into national legislation, and visualizing results.

A limitation of our study is that it is based on interviews with a small number of policy makers, so it is necessary to proceed with a quantitative survey of a larger number of individuals using a questionnaire in order to develop a more holistic picture of the legal information environment within the two pilot countries, and then in other countries with different legal systems and traditions. Furthermore, it is necessary to implement the novel legal information services identified in our study, and then evaluate them, and possibly improve and extend them.

## References

1. Lugo-Gil, J., Jean-Baptiste, D., Jaramillo, L. F.: *Use of Evidence to Drive Decision-Making in Government*. Mathematica Policy Research. (2019).
2. Kamateri, E., Panopoulou, E., Tambouris, E., Tarabanis, K., Ojo, A., Lee, D., Price, D.: A comparative analysis of tools and technologies for policy making. In *Policy Practice and Digital Science*, pp. 125-156. Springer, Cham. (2015).
3. ManyLaws Consortium: Deliverable 1.1: Report on the Assessment of Existing Tools, DSIs, and Data Infrastructures.(2019)
4. Arnott, D., Pervan, G., O'Donnell, P., Dodson, G.: An analysis of decision support systems research: Preliminary results. In *Proceedings of the 2004 IFIP International Conference on Decision Support Systems*, pp. 25-37. (2004) [full citation unavailable].
5. Phillips-Wren, G., Jain, L.: Artificial intelligence for decision making. In: *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, pp. 531-536. Springer, Berlin, Heidelberg (2006).
6. Arnott, D., Pervan, G.: Eight key issues for the decision support systems discipline. *Decision Support Systems*, 44(3), 657-672 (2008).
7. Sabatier, P. *Theories of the policy process*. Routledge, New York, N.Y. (2019)
8. Jann, W., Wegrich, K.: Theories of the policy cycle. In Fischer, F., Miller, G.J. (eds.) *Handbook of Public Policy Analysis*, pp. 69-88. Routledge, New York, N.Y. (2017)
9. Power, D. J. *Decision support systems: concepts and resources for managers*. Greenwood Publishing Group, Westport C.T. (2002).
10. Erdelez, S, O'Hare, S: Legal informatics: application of information technology in law. *Annual Review of Information Science and Technology*, 32: 367-402. (1997)
11. Sartor G., Francesconi, E.: Legal informatics and legal concepts. In: *Eurovoc Conference – 18-19 November 2010, Luxembourg*. (2010).
12. Hornsby, W.: *Improving the Delivery of Affordable Legal Services Through the Internet: A Blueprint for the Shift to a Digital Paradigm* (1999).
13. Byrd, O: *Legal Analytics vs. Legal Research: What's the Difference?*. *Law Technology Today*, 12 June 2017. <https://www.lawtechnologytoday.org/2017/06/legal-analytics-vs-legal-research/>
14. Ashley, K. D.: *Artificial intelligence and legal analytics: new tools for law practice in the digital age*. Cambridge University Press, Cambridge UK (2017).
15. Moens, M.-F.: Innovative techniques for legal text retrieval. *Artificial Intelligence and Law*,

- 9(1), 29–57 (2001).
16. Casanovas, P., Palmirani, M., Peroni, S., Van Engers, T., Vitali, F.: Semantic web for the legal domain: the next step. *Semantic Web*, 7(3), 213-227 (2016).
  17. Schweighofer, E.: Semantic indexing of legal documents. In: Francesconi, E., Montemagni, S., Peters, W., Tiscornia, D. (Eds.) *Semantic Processing of Legal Texts: Where the Language of Law Meets the Law of Language*, pp. 157-169. Springer, Berlin-Heidelberg (2010).
  18. Otiike, J.: Legal information needs of the general public. *Library Review*, 46(1), 28-33 (1997).
  19. Otiike, J., Matthews, G.: Legal information needs of lawyers in Kenya: a case study. *Library Management*, 21(5), 241-251 (2000).
  20. Peoples, L.F. Testing the limits of Westlaw next. *Legal Reference Services Quarterly*, 31(2), 125-149 (2012).
  21. Bhardwaj, R. K., Margam, M.: Online legal information system for Indian environment: a user's perspectives. *Library Review*, 65 (8/9), 593-624 (2016).
  22. Bhatt, A. A.: Information needs, perceptions and quests of law faculty in the digital era. *The Electronic Library*, 32(5), 659-669 (2014).
  23. Bhardwaj, R. J.: Development of Online Legal Information System: Lawyers' Perceptions. *Journal of Library & Information Technology*, 39(2), 131-139 (2019).
  24. Van Lamsweerde, A.: Goal-oriented requirements engineering: A guided tour. In *Proceedings - Fifth IEEE International Symposium on Requirements Engineering*. pp. 249-262, IEEE Press (2001).
  25. Chamberlain, S., Sharp, H., Maiden, N.: Towards a Framework for Integrating Agile Development and User-Centred Design. *Extreme Programming and Agile Processes in Software Engineering*. In *XP 2006: Extreme Programming and Agile Processes in Software Engineering*, pp 143-153, Springer, Berlin, Heidelberg (2006).
  26. Collier, K.: *Agile analytics: A value-driven approach to business intelligence and data warehousing*. Addison-Wesley, Boston, M.A. (2012)
  27. Wiegers, K., Beatty, J.: *Software requirements*. Pearson Education, Harlow, UK. (2013)
  28. Zowghi, D., Coulin, C.: Requirements Elicitation: A Survey of Techniques, Approaches, and Tools. In: *Engineering and managing software requirements*, pp. 19-46. Springer, Berlin, Heidelberg. (2005)
  29. Paetsch, F., Eberlein, A., Maurer, F.: Requirements engineering and agile software development. In *Proceedings - Twelfth IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*, pp. 308-313, WET ICE 2003. IEEE Press. (2003).
  30. Hennik, M., Hutter, I., Bailey, A. *Qualitative Research Methods*. SAGE Publications, London (2011).
  31. Yin, R.K.: *Qualitative Research from Start to Finish*. The Guilford Press: New York, N.Y. (2011).
  32. Kumar, R.: *Research Methodology. A Step-by-Step Guide for Beginners*. Fourth edn. SAGE Publications, London, UK. (2014).
  33. Lachana, Z., Loutsaris, M. A., Alexopoulos, C., Charalabidis, Y.: Automated Analysis and Interrelation of Legal Elements Based on Text Mining. *International Journal of E-Services and Mobile Applications (IJESMA)*, 12(2), 79-96 (2020).
  34. Loutsaris, M., Androutopoulou, A., Charalabidis, Y.: EU-wide legal text mining using big data processing infrastructures. In *12th Mediterranean Conference on Information Systems (MCIS 2018)* (2018).