AI-based solutions for legislative drafting in the EU

Summary report

September, 2024

Authors

Dr. Fotios Fitsilis, Hellenic Parliament
Prof. George Mikros, Hamad Bin Khalifa University

EUROPEAN COMMISSION

Directorate-General for Digital Services
Directorate A — Digital Transformation
Unit A.3 — HR & Legislative Solutions

Contact: Willy Van Puymbroeck

E-mail: Willy.Van-Puymbroeck@ext.ec.europa.eu

European Commission
B-1049 Brussels

EUROPEAN COMMISSION

AI-based solutions for legislative drafting in the EU

Summary report

Directorate-General for Digital Services

2024

Manuscript completed in September 2024

First edition

Luxembourg: Publications Office of the European Union, 2024

PDF ISBN 978-92-68-21348-3 doi:10.2799/6892931 NO-05-24-335-EN-N

Contents

Abbreviations

AI	Artificial Intelligence
AKN	Akoma Ntoso
API	Application Programming Interface
CBR-RAG	Case-Based Reasoning for Retrieval Augmented Generation
COCA	Corpus of Contemporary American English
CoT	Chain-of-Thought
DCEP	Digital Corpus of the European Parliament
DG	Directorate-General
EC	European Commission
EPIC	European Parliament Interpreting Corpus
EU	European Union
GPT	Generative Pre-trained Transformer
GUI	Graphical User Interface
I(C)T	Information (and Communication) Technologies
IE	Information Extraction
IRAC	Issue, Rule, Application, Conclusion
LEOS	Legislation Editing Open Software
LLM	Large Language Model
LOTM	Legal Ontology and Terminology Management
NER	Named Entity Recognition
NLG	Natural Language Generation
PoC	Proof(s)-of-Concept
R&D	Research and Development
RLLF	Reinforcement learning from logical feedback
SOULL	Sources of Language and Law
UAT	User Acceptance Testing
UNPC	United Nations Parallel Corpus
UX	User eXperience
XML	Extensible Markup Language

Executive summary

This publication provides an overview of the results of a European Union (EU) funded study entitled “Overview of smart functionalities in drafting legislation in LEOS”. The full study has been published on the European Commission's (EC) Joinup platform and centres on the concept of smart functionalities in law-making, i.e., advanced Information (and Communication) Technologies (I[C]T) services that assist legal drafters and policy developers in their daily work. The underlying research was conducted in view of the development of an “augmented LEOS”, an open-source solution developed by the EC for drafting legislation. The work draws on the results of a 2022 study on “Drafting legislation in the era of AI and digitisation”, referred to as the reference study.

The present study offers a thorough examination of various development steps of the "augmented LEOS" system. It confirms, updates, and expands upon the findings of the reference study. Moreover, it provides a detailed assessment of the business value associated with the proposed smart functionalities. The prioritisation of these functionalities is carried out based on their perceived business value. Furthermore, the study conducts an in-depth investigation into the implementation of these functionalities, addressing their deployment. Additionally, recognising the emergence of Large Language Models (LLMs), the study explores their utilisation in drafting legislation. In this context, potential implications and applications of LLMs in the legislative processes are analysed. Finally, the study suggests a high-level framework and roadmap for further work, outlining the necessary steps and milestones for the successful realisation of the augmented LEOS system.

The results presented herein draw on extensive desk research as well as a set of interviews with EC staff conducted in mid-2023. To the main results of the study belong the following ones:

- A revised categorisation of smart functionalities;
- An extended list of smart functionalities;
- A prioritisation for implementation of these smart functionalities;
- A detailed description of the business value of the priority smart functionalities;
- The identification of a set of core Artificial Intelligence (AI) technologies required for their implementation;
- A description of main attributes to consider when implementing and deploying smart functionalities;
- Details on an envisaged integration platform and datasets;
- Considerations on the use of LLM in law-making and policy development; and

- A high-level implementation and deployment framework and roadmap.

Moreover, the study indicates a series of considerations to be taken into account when integrating AI-based tools and services in legislative processes. Some of these considerations are privacy, security, legal interoperability, training requirements, and aligning with regulations like the EU AI Act. The study also notes that while AI technologies evolve rapidly, the key focus areas remain valid starting points. Very important, it reveals that most smart functionalities can be implemented without the need to resort to LLMs. It concludes by summarising its tangible approach and contributions toward developing an "augmented LEOS" capable of effectively supporting existing legal processes at EC level, while exploiting state-of-the-art technology.

1. Introduction

The use of innovative IT and AI has the potential to digitally transform legal drafting and decision making, among others by improving the quality, efficiency, and transparency of the legislative process. The premise of the underlying work is to bring about a paradigm shift in law-making by (i) embracing machine readable/processable law, (ii) leveraging advancements in I(C)T, (iii) using standards, and ultimately (iv) enabling a deepened understanding of law-making theory and practice. For materialising such a paradigm shift a list of smart functionalities are proposed. In time, these smart functionalities should be integrated within an “augmented LEOS” system, which represents the evolution of the existing LEOS platform ([1]) into a future-proof, complete, user-assistive drafting tool that can be used throughout the complete legislative life cycle. Eventually, this “augmented LEOS” could become the core of an extended digital ecosystem, aimed to ultimately benefit public administrations, citizens, society, and businesses.

The scope of this publication is to offer a concise and high-level representation of the main results of a comprehensive EU-funded study entitled "Overview of Smart Functionalities in Drafting Legislation in LEOS", which was implemented between June 2023 and February 2024 (Fitsilis et al., 2024). This work builds upon a reference study on “Drafting legislation in the era of AI and digitisation” that was presented in 2022 (Palmirani et al., 2021).

This reference study covered issues for representing legal knowledge and applying responsible hybrid AI to law, with an emphasis on the use of Akoma Ntoso (AKN) and LegalXML standards to provide for machine-readable legal data. In addition, it outlined use cases demonstrating the potential of using AI. Moreover, it identified an initial set of smart functionalities that could assist legal drafters and policy developers but also roadblocks and challenges that may inhibit their utilisation. Last but not least, initial considerations for implementation of an “augmented LEOS” architecture were shown along with a rudimentary implementation strategy.

Smart functionalities and their implementation are the core themes of this publication. The reference study introduced the term smart functionality, without however providing a tangible definition of the term “smart functionality”. For the follow-up study, a working definition is provided. According to it, a smart functionality is an advanced and intelligent software capability to enhance the performance of the LEOS system that uses both AI and legacy technologies. Smart functionalities may increase productivity, improve quality of legislation, automate tasks, improve user experiences, and make data-driven decisions, thus potentially transforming the law-making process.

([1]) See LEOS core at LEOS Community / LEOS · GitLab (europa.eu) and annotation utilities at LEOS Community / Annotate · GitLab (europa.eu)

Beyond the current introduction (Section 1), the publication discusses the categorisation and the prioritisation of smart functionalities (Section 2). The results reported on in this section draw on a series of interviews and a questionnaire done with staff in the EC in mid-2023. The analysis makes good use of (and thus serves as an illustration of what is possible with) advanced digital technologies, including AI. A techno-assessment of implementing prioritised smart functionalities is presented in Section 3. It details the core technologies needed to implement these priority smart functionalities, addresses key attributes, and considers main technical enablers. Substantial attention is vested in the discussion of the potential use of Generative AI and specifically of LLMs in the law-making and policy development process (Section 4). The section is based on extensive literature research within a rapidly evolving field. Section 5 presents a concrete high-level framework/roadmap and the publication closes with the conclusions and an outlook (Section 6).

2. Identifying categories and setting priorities

2.1. Methodology

For identifying distinct categories and setting priorities for the technical implementation of smart functionalities this study gathered and evaluated substantial empirical data through a combination of questionnaires and structured interviews. Eleven interviews with experts from various EC Directorates-General (DGs) were conducted. The interviews followed a predefined format and were limited to 60 minutes. They covered topics such as the experience with the LEOS/EdiT ([2]) system, the potential and challenges of using smart functionalities, and the use of LLMs. Following each interview, a list of smart functionalities was shared with the interviewees, who were asked to select and provide a rationale for five functionalities they would like to see implemented first.

The interviews were recorded and processed to extract accurate transcripts, using advanced AI tools and innovative audio analysis technologies. This comprehensive audio processing pipeline involved various stages, including isolating vocals, transcribing the audio, speaker diarisation, and refining the speech segment alignment, ultimately resulting in clear transcriptions alongside speaker identification. The approach employed purposeful sampling for data collection, emphasising qualitative interpretation. All research outputs were meticulously anonymised to ensure confidentiality and privacy. The following section presents the main findings of the interviews using eBriefing Lab – an internal LLM tool piloted in the EC.

2.2. Main findings of the interviews

After analysing the interviews, it becomes clear that AI and LLMs can indeed enhance legislative drafting. In particular, the interviews revealed a strong belief that AI and LLMs can improve the quality and efficiency of legislative drafting, e.g., by automating repetitive tasks, providing quality checks, and aiding in the extraction and analysis of information. In this regard, despite the potential of AI, the interviewees emphasised the importance of human oversight in the legislative drafting process. This is necessary to ensure the accuracy and appropriateness of the drafted legislation, as well as to maintain trust in the process.

However, the interviewees stressed that the adoption of AI and LLMs in legislative drafting requires a significant cultural shift among policy officers. This includes training in the use of these technologies and a change in mindset from working

([2]) EdiT is the LEOS instantiation used in the EC

independently to collaborating with AI systems. This again raises new challenges, such as the need for trust and transparency, in the sense that these technologies must be reliable to users, transparent in their operations, and capable of earning the trust of the policy officers who use them.

While the potential of smart functionalities to improve legislative drafting is significant, the approach to be taken toward their implementation needs to be determined and several interviewees emphasised the importance of investing in European, open-source technologies to ensure control over the data and systems used in legislative drafting. In the course of the interviews, additional smart functionalities were collected and analysed, thus extending the set from 34 to 61. All these smart functionalities were classified in categories (see *Section 2.3* and *Section 2.4*), prioritised based on the analysis of the questionnaires (see *Section 2.5*), and matched with specific AI technologies for implementation (see *Section 3.2*).

2.3. Categories of smart functionalities

The study suggested grouping smart functionalities in seven categories that each encompass a range of functionalities vital for the effective support of drafting legislation and developing policies. The seven categories are presented in *Table 1* along with a brief description.

Table 1 – Categories of smart functionalities

Category	Concise description
Verification	Smart functionalities that verify the legal correctness including among others the accurate usage of citations for validity and relevance, the proper referencing of relevant and up to date articles, the consistency of legal definitions, and the correct use of specific lexica and acronyms.
Change tracking	Smart functionalities to track and manage the changes made in a document, to allow tracing back history, and to assist comparing different versions of a document.
Linguistic support	Smart functionalities that help improve the quality and readability of legislative texts by checking grammar, consistency, clarity, and the use of correct linguistic formulations in accordance with relevant guides including the handling of multi-lingual aspects such as detecting divergences between linguistic translations.

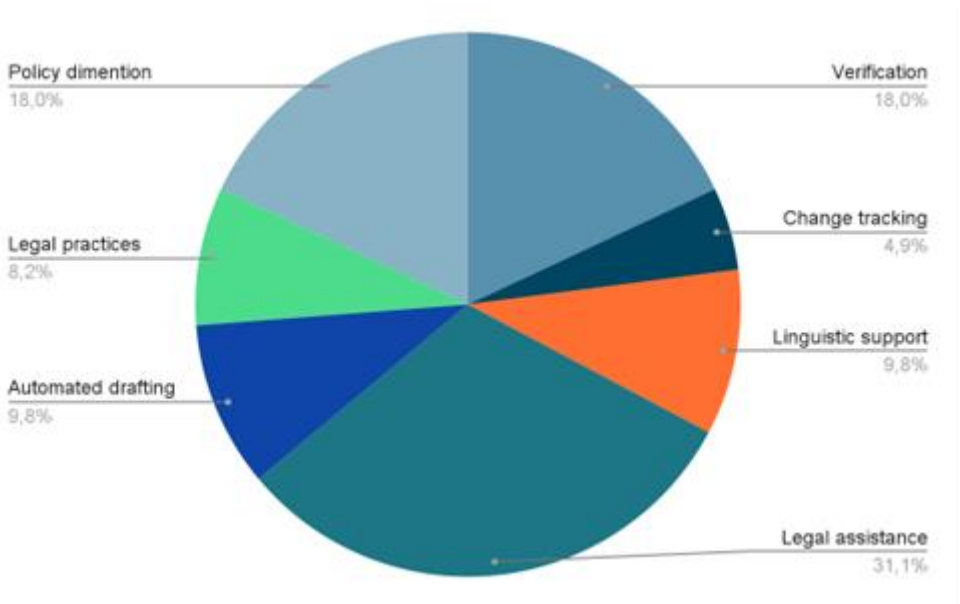
Category	Concise description
Legal assistance	Smart functionalities that support the legal consistency, coherence, and adherence to relevant legal drafting rules within an act and between acts including, among others, the detection and avoidance of structures that could create unintended ambiguities in legal interpretation, the correlation between recitals and enacting terms, the identification of linkages with preceding acts, and the detection of explicit or implied obligations, rights, permissions or penalties.
Automated drafting	Smart functionalities that use LLMs to support the drafting process, for instance by providing suggestions or guidance on terminology, structure, and conventions, and drafting or consolidating legislative texts.
Drafting practices	Smart functionalities to detect drafting patterns, good practices, and common errors thereby contributing to knowledge management and sharing
Policy dimension	Smart functionalities, e.g., to estimate the impact of a legislative act, measure its digital-readiness, verify interoperability aspects, or compliance with policy mandates such as gender-neutrality.

After identifying the various categories of smart functionalities in the law-making process, the study proceeded to classify them individually, also taking into consideration the findings from expert interviews and the analysis of the questionnaires.

2.4. Extended list of smart functionalities

The initial list of 34 smart functionalities, the starting point for this study, is given in *Annex 1*. In the course of the consultation process via interviews and questionnaires, additional 27 distinct ones were identified. The identification of the additional functionalities followed a thorough process of refinement, cross checking, and reformulation to arrive at specific proposals. The additional 27 smart functionalities are listed in *Annex 2.* The list that does not include the ones that were already present in the initial list. Furthermore, proposals that could not be immediately associated with an augmented LEOS were not taken on board.

Figure 1 – Distribution of smart functionalities



As a result, a list of 61 smart functionalities is identified. The distribution of these smart functionalities in their respective categories is depicted in *Figure 1*.

2.5. Prioritisation of smart functionalities

To prioritise the implementation of smart functionalities, the interviewees were asked to express their preferences. The results are shown in *Table 2*, which displays the 11 smart functionalities that were selected by at least four experts (parameter: No. of picks).

Table 2 – Priorityof Smart Functionalities (SF)

SF	Title	No. of Picks
#14	Correlation between recitals and the enacting terms	6
#20	Automatically identify existing legislation relevant for the act under development	
#3	Acronyms, organisations, and other abbreviations	4

SF	Title	No. of Picks
#9	Use correct linguistic formulations within the structure of the document	
#10	Correct formulation in accordance with the English Style Guide	
#11	Detect divergences between different linguistic translations	
#12	Suggest linguistic formulations in provisions	
#13	Detect and avoid structures that could create issues in legal interpretation	
#15	(Correlation) between previous acts and the new one	
#19	Detect obligations, rights, permissions, penalties	
#26	LLM based legal text generation	

3. Techno-evaluation

3.1. Approach

The main focus of the study targets the business value assessment of smart functionalities and their techno-business feasibility. In this regard, it needs to be mentioned that business value and business feasibility are two separate concepts in business analysis. In essence, while business value focuses on the benefits and outcomes of a certain smart functionality, business feasibility evaluates the practicality and likelihood of successfully implementing it.

The evaluation of the techno-business feasibility of smart functionalities was conducted across three tiers in a quasi-parallel fashion. The first tier involved assessing and documenting the business value. The overall value is determined by weighing various factors, such as quality, user effort, automation, efficiency gains, explainability of results, and complexity. The business value assessment was achieved through interviews with EC staff, the evaluation of responses to a questionnaire featuring the initial set of smart functionalities, and the detailed scrutiny of each of the priority smart functionalities.

The second tier focused on the technical feasibility of the priority functionalities, among others by identifying the necessary technologies and assessing their availability, considering the ease of integration into LEOS, bearing in mind data and knowledge requirements, and valuing the open-source dimension of relevant applications and tools. The third tier involves the overall business feasibility of deploying smart functionalities.

3.2. Mapping categories to attributes

As a starting step toward techno-evaluation, the seven categories of smart functionalities were mapped to a predefined set of attributes. The attributes that were considered fit for purpose for building and operating the resulting platform and applications are user experience (UX), potential business value, technology stack, aspects of related datasets, and performance considerations. The result is shown in *Annex 3*.

As can be noticed in this matrix, there are some key aspects that span multiple features and/or system architectural components. Specifically, features that require multiple visual items need to be implemented in a non-intrusive manner to avoid disruption and additional complexity, while remaining accessible when needed. Additionally, due to the processing needs of the algorithms and the multitude and

scale of the datasets, mechanisms must be incorporated within the implementation and the hosting infrastructure that can handle increased workloads.

3.3. Matching priority of smart functionalities to technologies

The core technologies for each of the priority smart functionalities were then identified and selected based on their relevance, efficacy, and considering the specific challenges in view of the unique demands of legal drafting. As a result, the prioritised smart functionalities were clustered into five distinct technology groups, i.e., I to V in Roman numerals, as per *Table 3*. It is important to note that the majority of these technologies are well-established and tested, with underlying algorithms that can be readily integrated into LEOS.

Table 3 - Grouping of priority smart functionalities in technology clusters.

SF Cluster	Technology	SF
I	Advanced Language Editing and Correction	#9-#10-#12-#13
II	Named Entity Recognition	#3
III	Semantic Similarity	#11-#14-#15-#20
IV	Natural Language Generation	#26
V	Information Extraction	#19

The five core technologies necessary for the implementation of the priority smart functionalities are as follows:

- **Semantic Similarity** - This technology was chosen for tasks that require a nuanced understanding of language and context, particularly for correlating different sections of legal documents and detecting latent semantic dimensions. It plays a crucial role in identifying and understanding the underlying meanings and relationships within complex legal texts;
- **Named Entity Recognition** - Named Entity Recognition (NER) is applied to functionalities involving the identification of specific legal entities, citations, and references. NER is known for its precision in extracting structured information from unstructured text, making it essential for accurately identifying and categorising entities within legal documents;

- **Information Extraction** - Information Extraction (IE) is designated for functionalities where extracting specific data, such as obligations, rights, and legal statuses, from complex legal texts is crucial. This technology is instrumental in automatically identifying and extracting relevant information from large volumes of legal documents;
- **Natural Language Generation** - Natural Language Generation (NLG) is assigned to tasks requiring the generation of new legal text, such as drafting amendments. NLG technology has the capability to produce coherent, contextually appropriate content, making it suitable for generating new legal text based on specific requirements;
- **Advanced Language Editing and Correction** - This technology was selected for functionalities that necessitate sophisticated linguistic and stylistic refinement of legal texts. It is essential for ensuring the accuracy, clarity, and coherence of legal documents, providing advanced editing and correction capabilities to enhance the quality of legal texts;

The final study report provides for each of these five technologies a detailed description, technology assessment, overview of the state of the art, and an indication of existing open-source solutions. In addition, it assesses that "'only" two additional technologies are necessary to implement the full initial set of 34 smart functionalities. These technologies are:

- **Legal Ontology and Terminology Management** - Legal Ontology and Terminology Management (LOTM) technology is applied to tasks involving the management of complex legal terminologies and definitions within legal documents. This technology is essential for ensuring consistency and accuracy in the interpretation and use of legal terms, which is critical for legal professionals and stakeholders. LOTM provides a structured framework for organising and managing legal concepts and their relationships, thereby facilitating precise and standardised communication within the legal domain. It enables the effective management of legal terminologies, definitions, and relationships, ultimately enhancing the clarity and coherence of legal documents;
- **Text Classification** - Text Classification technology is chosen for functionalities that require the categorisation and analysis of legal documents based on their content and structure. This technology plays a pivotal role in automatically categorising and organising legal texts based on predefined criteria, such as legal topic, document type, or relevant legal concepts. By leveraging machine learning algorithms and natural language processing techniques, text classification enables the efficient organisation and retrieval of legal documents, thereby facilitating effective information retrieval and analysis within the legal domain. It helps in automating the process of categorising and structuring legal documents, contributing to improved document management and accessibility.

Within the current scope, however, a detailed examination of the final two technologies was not conducted. Instead, the study deliberated on and assessed the aforementioned attributes for all prioritised smart functionalities.

3.4. Attributes

In particular, two main attributes, notably *UX* and *performance*, deserve special mention in view of the implementation of smart functionalities. UX aims to proactively identify potential user satisfaction issues, establish benchmarks, and guide development decisions to optimise user experience. The ex-ante evaluation for UX specifications is based on conducted interviews and collected questionnaires.

The analysis highlights the necessity for a high degree of customisation to meet diverse user needs, inter alia considering that identical smart functionalities may serve different purposes within a single or multiple DGs. For example, customisable UX templates could accommodate tailored experiences in certain DGs, additional DG-level data sources, or specific processes that need to be incorporated. EC experts have identified specific UX features, akin to those found in standard commercial applications.

Performance is paramount when implementing software tools and features, particularly in scenarios involving large, distributed datasets requiring access and processing. Also, addressing non-functional requirements is central in ensuring that the overall system performance is not compromised by the substantial workload imposed by such implementations. Critical considerations for mitigating and managing this risk are necessary, guaranteeing the availability of features without compromising the integrity of the system.

Performance analysis extends beyond technical aspects and encompasses user engagement, quality of suggestions, legal consistency and coherence, session duration, and feature usage patterns to evaluate overall satisfaction. To ensure reliable tracking and evaluation by system developers, performance indicators must be specific, measurable, and relevant to the specific smart functionalities to which they apply.

3.5. Technical enablers

Inevitably, the focus is placed on the implementation of smart functionalities. For this, two technical enablers, integration technology and datasets, merit special handling. The discussion on integration tackles the technology stack and architectural components. These are necessary to facilitate the seamless integration of smart functionalities into LEOS. Such an integration platform, or simply “platform”, would

necessarily include several key features, the most significant of which are appearing below.

One aspect of integration centres on the development process and involves the technical implementation of connecting and integrating external systems with LEOS, ensuring interoperability and seamless communication. This involves managing the Application Programming Interfaces (APIs) to enable the smooth exchange of data and functionalities between LEOS and external systems.

In addition, work on exposing and accessing the business logic and features of LEOS via API seems to be necessary. The platform would also address the introduction of different document types into LEOS. This includes creating document collections, such as written question and answer documents, to support a comprehensive document ecosystem with correlations and hierarchies between documents.

The platform would explore the inclusion of logic for integrating business process automation use cases in LEOS. This involves incorporating and adapting business processes into the system to streamline and automate law making and policy development workflows. Moreover, it seems necessary to consider the integration of new, advanced features into the Graphical User Interface (GUI) of LEOS.

In developing an integration platform for law making purposes, such as the one envisaged around the augmented LEOS, it is important to consider several issues around datasets. This involves for instance effective structuring, which is necessary for accessing and searching through multiple repositories housing legal texts and documents, often based on various criteria. In this regard, the platform's capability to access historical data and facilitate cross-referencing between different cases and court rulings is considered important. Additionally, jurisdictional data might be needed when dealing with EU member state specific legal orders. Hence, securing access to repositories from third parties and organisations is deemed critical.

However, when developing smart functionalities, it is advisable to initially prioritise attention on EU law and jurisprudence, given the accessibility and quality of relevant datasets. Expanding beyond this scope, for instance to encompass member state or international law, introduces considerable challenges, predominantly stemming from issues related to the quality and homogeneity of decentralised, non-standardised datasets.

4. On the use of LLM in legal drafting in the public sector

4.1. Context

This section explores the use of AI, particularly LLMs, in the legal domain and the public sector. AI encompasses a range of methods aimed at replicating and approximating human behaviour when solving complex problems. As AI tools and services become more sophisticated, their use in the public sector is increasing, with the potential to transform governance in institutions and change how public services are developed and delivered.

The section is organised as follows: *Section 4.2* discusses the importance of legal corpora and various resources for studying and applying legal language. *Section 4.3* examines the capabilities and limitations of LLMs in handling basic legal text, highlighting the need for fine-tuning and collaboration between AI specialists and legal professionals. *Section 4.4* explores the concept of legal prompting and its potential to teach LLMs to think more like legal professionals by incorporating legal reasoning frameworks and relevant context. *Section 4.5* compares and contrasts the use of hybrid AI and LLMs in legal drafting, discussing their strengths, limitations, and potential synergies. Finally, *Section 4.6* proposes piloting hybrid AI and LLMs in legislative drafting as a responsible approach to innovation, emphasising the importance of human control, data capitalisation, building early experience, and the establishment of effective guardrails.

4.2. Legal corpora

Legal language demands a dedicated set of resources for study and application due to its specialised nature. These resources include corpora, lexical databases, grammatical, and stylistic guidelines, as well as references for acronyms, organisations, and abbreviations.

Corpora, which are collections of texts processed using specialist software, are instrumental for examining words and phrases in their contexts and comparing them across different periods. Notable legal corpora include:

- The EU institutions have put together considerable resources on EU law (3) and related vocabularies; (4)
- The United Nations Parallel Corpus (UNPC), (5) which is a large multilingual corpus consisting of manually translated UN documents from 1990 to 2014 in the six official UN languages. The UNPC allows for studying legal language usage across different languages and time periods;
- The Digital Corpus of the European Parliament (DCEP), containing 1.37 billion words in 23 languages from various document types produced between 2001-2012. DCEP enables analysis of legal terminology and phrasing within the context of European Parliament proceedings (Hajlaoui et al., 2014);
- The SOULL (Sources of Language and Law) platform, (6) which provides information about existing data collections and corpora of legal language in various sizes, languages, and text types;
- The Corpus of Contemporary American English (COCA), (7) the first large and diverse corpus of American English from 1990-2008, balanced across spoken, fiction, magazines, newspapers and academic journals. While not exclusively legal, COCA provides a representative sample of modern American English usage, including in legal domains (Davies, 2009);
- The European Parliament Interpreting Corpus (EPIC), which is used to research the simultaneous interpretation of European Parliament proceedings. EPIC supports the study of interpreting strategies and challenges in rendering legal language (Sandrelli & Bendazzoli, 2006).

Lexical resources such as Black's Law Dictionary ([8]) and Garner's Dictionary of Legal Usage ([9]) are crucial for understanding the specific terminology used in legal language. Furthermore, grammatical and stylistic guides like the Legal English Grammar Guide ([10]) and Adobe Legal Department Style Guide ([11]) offer rules and recommendations for writing in legal language. These help ensure clarity, consistency, and accuracy in legal drafting. Additionally, resources for acronyms,

([3]) See https://op.europa.eu/en/law

([4]) See https://op.europa.eu/en/web/eu-vocabularies/home

([5]) See https://conferences.unite.un.org/uncorpus

([6]) See https://legal-linguistics.net/data-collections/

([7]) See https://www.english-corpora.org/coca/

([8]) See https://thelawdictionary.org/

([9]) See https://global.oup.com/academic/product/garners-dictionary-of-legal-usage-9780195384208

([10]) See https://www.amazon.com/Legal-English-Grammar-Guide/dp/B087SCDQH3

([11]) See https://www.adobe.com/content/dam/cc/en/legal/documents/ADOBE-LEGAL-STYLE-GUIDE.pdf

organisations, and abbreviations like Eurostat's List of Abbreviations and Acronyms, ([12]) YourDictionary's List of EU Abbreviations, ([13]) and fi-compass Acronyms ([14]) provide comprehensive listings for reference, catering to the diverse linguistic needs within the legal domain.

These resources collectively form a vital foundation for the study and application of legal language, offering essential support for linguistic analysis, comprehension, and effective communication within the legal field. Parallel corpora enable comparisons of legal language across languages, while monitor corpora like COCA allow tracking linguistic changes over time. Hierarchical document encoding models show promise for mining parallel legal data. Overall, the breadth of multilingual legal corpora available provides ample opportunity for in-depth research into this specialised domain.

4.3. Can LLMs handle basic legal text?

The capability of LLMs to handle basic legal text is a critical consideration in the integration of artificial intelligence into legal practice. Recent studies have revealed significant limitations in the performance of leading LLMs such as GPT-4, Claude, and PaLM 2 when tasked with basic legal text handling (Blair-Stanek et al., 2023). Blair-Stanek et al. (2023) developed a comprehensive benchmark (BLT Benchmark) to assess LLMs' ability to handle tasks that lawyers and paralegals would reasonably expect them to perform, such as retrieving specific information from a witness deposition or contract subsection. However, the results of this benchmark indicate that LLMs currently struggle to perform adequately on these tasks, raising doubts about their reliability for legal practice without further enhancements.

Interestingly, fine-tuning LLMs for these specific legal tasks has shown promising results. Even smaller models, when fine-tuned, achieve near-perfect performance on the test set, demonstrating the potential for improvement with targeted training (Li et al., 2023). Moreover, fine-tuning for basic legal text-handling tasks also enhances performance on related legal tasks, highlighting the transferability of improvements across domains (Douka et al., 2021).

These findings suggest that while foundational LLMs may currently lack the necessary capabilities for basic legal text handling out-of-the-box, additional engagement from subject matter experts and fine-tuning processes can significantly

([12]) See https://ec.europa.eu/eurostat/esa2010/chapter/view/27/

([13]) See https://www.yourdictionary.com/articles/eu-abbreviations

([14]) See https://www.fi-compass.eu/info/acronyms

enhance their performance. Collaboration between AI specialists and legal professionals is crucial in identifying and addressing the specific requirements of legal practice, ensuring that LLMs can effectively support legal professionals in their work.

It is important to note that the performance of LLMs in legal text handling may vary depending on the specific domain and complexity of the legal documents. For example, LLMs have shown promising results in the semantic annotation of legal texts in zero-shot learning settings, particularly with the emergence of more capable models like GPT-4 (Savelka & Ashley, 2023). On the other hand, LLMs may struggle with more complex tasks, such as multi-label classification of longer legal documents, highlighting the limitations of lightweight approaches (Clavié et al., 2021).

In this developing field, continued research and development efforts are essential to optimise LLMs for legal applications and improve their overall reliability and utility in legal practice. This includes exploring domain-specific pre-training strategies (Douka et al., 2021), leveraging structural information in legal documents (Sun, 2023), and developing efficient fine-tuning techniques (Clavié et al., 2021). By addressing these challenges and opportunities, LLMs have the potential to revolutionise the way legal professionals access, analyse, and utilise legal information, ultimately enhancing the efficiency and effectiveness of legal practice.

4.4. Legal prompting: teaching an LLM to think like a lawyer

The emergence of LLMs capable of zero or few-shot prompting approaches has sparked a new research area known as prompt engineering. These innovative approaches leverage LLMs' capacity for understanding prompts to improve performance across various tasks. Notably, recent advances, such as the Chain-of-Thought (CoT) prompts, have demonstrated significant enhancements in tasks related to arithmetic and common-sense reasoning (Wei et al., 2022).

In the legal domain, researchers are exploring how these prompting techniques can be applied to improve LLMs' legal reasoning capabilities. A study by Yu et al. (2022) investigated the effectiveness of various prompting approaches on the COLIEE entailment task, which is based on the Japanese Bar exam. Their findings showed that while CoT prompting and fine-tuning with explanations led to improvements, the best results were obtained using prompts derived from specific legal reasoning techniques, such as IRAC (Issue, Rule, Application, Conclusion).

The IRAC framework is a widely used method for structuring legal analysis and arguments. By incorporating this framework into the prompts, researchers were able to guide the LLM to break down legal problems into their constituent parts, identify

the relevant legal rules, apply those rules to the facts of the case, and draw conclusions based on the analysis (Yu et al., 2022). This structured approach to legal reasoning aligns with the way lawyers are trained to think and argue, making it a promising avenue for enhancing LLMs' legal capabilities.

Another important aspect of legal prompting is the inclusion of relevant legal authorities and context. As demonstrated by the CBR-RAG (Case-Based Reasoning for Retrieval Augmented Generation) approach, providing LLMs with contextually relevant cases as part of the prompting process can lead to significant improvements in the quality of generated answers for legal question-answering tasks (Wiratunga et al., 2024; Mamalis et al., 2024). By augmenting the original LLM query with retrieved cases, the model is provided with a richer prompt that includes the necessary legal context to reason more effectively.

As LLMs continue to advance, their ability to engage in legal reasoning tasks is expected to improve. However, it is crucial to recognise that legal reasoning is a complex and detailed process that requires logical thinking and understanding of the underlying legal principles, precedents, and societal context. While legal prompting techniques can help guide LLMs towards more lawyer-like thinking, there is still a significant gap between current models' capabilities and legal professionals' expertise.

To further enhance LLMs' legal reasoning abilities, researchers are exploring various approaches, such as reinforcement learning from logical feedback (RLLF) (Nguyen et al., 2023) and the development of specialised legal benchmarks like LegalBench (Guha et al., 2023). These efforts aim to refine LLMs' reasoning capacities and comprehensively evaluate their legal capabilities.

As research in this area progresses, the potential for LLMs to support and augment legal professionals in their work is expected to grow. However, it is essential to recognise the limitations of current models and the need for continued research and development to bridge the gap between machine and human legal reasoning capabilities.

4.5. Hybrid AI versus (or in addition to) LLMs

While the current study concludes that essentially hybrid AI suffices to implement priority smart functionalities, exploring synergies/complementarities between the use of hybrid AI and LLM in legal drafting is useful. The rapid advancements in both hybrid AI and LLMs have opened up new possibilities for enhancing legal drafting processes. Hybrid AI, which combines symbolic and sub-symbolic AI techniques, has been proposed as a promising approach for implementing smart functionalities in legal systems (Palmirani et al., 2021). On the other hand, LLMs have demonstrated remarkable capabilities in natural language understanding and

generation, making them well-suited for legal text analysis and generation tasks (Xu & Ashley, 2023; Chalkidis et al., 2019).

Despite the initial focus on hybrid AI in the reference study conducted in 2022 (Palmirani et al., 2022), the rapid progress in LLMs has led to a growing interest in exploring their potential applications in the legal domain. LLMs, such as GPT-4, have shown impressive performance in legal question answering, legal text classification, and even legal reasoning tasks (Li et al., 2023; Bommarito & Katz, 2021; Branting et al., 2021). These advancements suggest that LLMs could play a significant role in augmenting and enhancing legal drafting processes.

However, it is important to recognise that hybrid AI and LLMs have strengths and limitations. Hybrid AI approaches excel in incorporating domain knowledge and explicit reasoning capabilities, which are crucial for ensuring the accuracy and interpretability of legal decisions (Palmirani et al., 2021). On the other hand, LLMs are highly effective in handling unstructured text data and generating human-like responses, making them valuable for tasks such as legal document summarisation and drafting assistance (Steenhuis et al., 2023). A detailed comparison of both approaches is shown in *Table 4*.

Table 4 – A comparison of the advantages and the disadvantages of LLMs and the Hybrid AI systems

	LLMs	Hybrid AI
Advantages	**Implicit knowledge**: LLMs can leverage vast amounts of text data to learn implicit patterns and relationships, which can aid in legal research and analysis.	**Completeness**: Hybrid AI systems can incorporate both rule-based logic and machine learning capabilities, potentially providing a more comprehensive approach to legal analysis.
	Efficiency: LLMs can process large volumes of legal documents quickly, potentially saving time and resources for legal professionals.	**Language understanding**: Hybrid AI systems can combine natural language processing with structured data analysis, enabling better comprehension of legal texts and contexts.
	Versatility: LLMs can handle various legal tasks, including document review, contract analysis, and legal research, with relatively little customisation.	**Incorporation of unseen facts**: Hybrid AI systems can integrate new or unseen information into their analysis, enhancing adaptability and responsiveness to evolving legal scenarios.

	LLMs	Hybrid AI
	Automation: LLMs can automate repetitive tasks, freeing up human lawyers to focus on more complex or strategic aspects of legal work.	**Interpretability**: Hybrid AI systems can offer more transparent decision-making processes, allowing users to understand and validate their outputs more easily.
Disadvantages	**Hallucinations**: LLMs may generate erroneous or misleading information, particularly when faced with ambiguous or incomplete input.	**Complexity**: Hybrid AI systems may require more extensive development and maintenance efforts compared to LLMs, due to the need to integrate multiple components and data sources.
Disadvantages	**Indecisiveness**: LLMs may struggle to provide definitive answers or legal advice due to uncertainties or conflicting information in legal texts.	**Lack of standardisation**: Hybrid AI systems may lack standardised methodologies or frameworks, making it challenging to ensure consistency and reliability across different implementations.
	Black-box approach: LLMs' decision-making processes can be opaque, making it difficult to understand how they arrive at their conclusions or predictions.	**Resource intensiveness**: Developing and deploying hybrid AI systems may require significant computational resources and expertise, particularly for training and fine-tuning machine learning models.
	Lacking domain knowledge: LLMs may lack a deep understanding of legal concepts and nuances, leading to errors or oversights in legal analysis.	**Potential for bias**: Hybrid AI systems, like any AI model, can inherit biases from the data they are trained on, potentially leading to unfair or discriminatory outcomes if not properly addressed.

Given the complementary strengths of hybrid AI and LLMs, it is worth exploring the potential synergies between these two approaches. For instance, hybrid AI systems could be used to encode legal rules, regulations, and domain-specific knowledge, while LLMs could be leveraged for natural language understanding and generation tasks (Palmirani et al., 2021; Steenhuis et al., 2023). By combining the strengths of both approaches, it could be possible to develop more robust and comprehensive legal drafting solutions that can handle a wide range of tasks and requirements.

Moreover, the choice between hybrid AI and LLMs for legal drafting may depend on the specific smart functionality being implemented. For tasks that require explicit reasoning and adherence to strict legal rules, hybrid AI approaches may be more suitable (Palmirani et al., 2021). On the other hand, for tasks that involve the analysis and generation of large volumes of legal text, LLMs may offer significant advantages in terms of efficiency and scalability (Xu & Ashley, 2023).

As both hybrid AI and LLM technologies continue to evolve rapidly, it is crucial for legal professionals and researchers to stay updated on the latest developments and assess the suitability of each approach for their specific use cases. Regular evaluations and benchmarking studies, such as LegalBench (Guha et al., 2023) and the use of question-answering approaches for evaluating legal summaries (Xu & Ashley, 2023), can provide valuable insights into the strengths and limitations of different AI approaches in the legal domain.

Furthermore, the ethical implications of using AI tools like LLMs in legal drafting and scholarly writing must be carefully considered (Hosseini et al., 2023). While LLMs can be useful in writing, reviewing, and editing text, they cannot be held morally or legally responsible for their actions due to their lack of free will. Therefore, it is essential to establish guidelines for the responsible use and disclosure of AI tools in legal and academic contexts.

4.6. Piloting hybrid AI and LLMs in legislative drafting

Drafting legislation presents a unique opportunity to explore the combined potential of hybrid AI and LLMs in the legal domain while prioritising human control and involvement throughout the process. This use case allows for the early adoption of these technologies, providing valuable insights and opportunities to assess risks and establish effective guardrails. By piloting hybrid AI and LLMs in legislative drafting, lawmakers can gain hands-on experience and develop a deep understanding of the capabilities and limitations of these tools, enabling them to make informed decisions about their future implementation and regulation.

Acknowledging that humans will be the primary enabling factor in the successful integration of hybrid AI and LLMs in legislative drafting, it becomes essential to prioritise human-centric software development, legal and ethical legitimation, and inclusive training of stakeholders involved in the law-making process. Human-centric software development ensures that the tools are designed with the needs and requirements of lawmakers and legal experts in mind, promoting usability, transparency, and trust.

Legal and ethical legitimation involves the establishment of clear guidelines and principles that govern the use of these technologies, ensuring that their application aligns with fundamental legal and ethical values. Inclusive training of stakeholders, including lawmakers, legal experts, and AI developers, fosters a shared understanding of the capabilities, limitations, and potential risks associated with hybrid AI and LLMs, enabling effective collaboration and responsible use of these tools.

One of the key advantages of using hybrid AI and LLMs in legislative drafting is the ability to capitalise on the vast amount of data available to lawmakers. These technologies can process and analyse large volumes of legal texts, case law, and other relevant information, empowering lawmakers to make data-driven decisions and enhance the overall quality of the legislative process. By putting this wealth of data at the fingertips of lawmakers, hybrid AI and LLMs can help identify patterns, trends, and best practices in legislative drafting, ultimately leading to more consistent, coherent, and effective legislation. However, it is crucial to emphasise that lawmakers retain control throughout the entire process, with the final responsibility for the drafted legislation lying with human legislators.

The early experiences gained through piloting these technologies in legislative drafting will allow lawmakers to assess the risks and establish effective guardrails to ensure the responsible and transparent use of AI tools in this critical area of law. These guardrails may include the development of clear guidelines and standards for the use of hybrid AI and LLMs, as well as the implementation of robust human oversight and validation mechanisms. By establishing these safeguards early on, lawmakers can mitigate potential risks, such as the generation of erroneous or biased content, and ensure that these technologies align with legal and ethical principles.

Moreover, piloting hybrid AI and LLMs in legislative drafting allows lawmakers to gain "skin in the game" and actively shape the development and application of these technologies in the legal domain. By being at the forefront of this innovation, lawmakers can work closely with AI developers, legal experts, and other stakeholders to ensure that these tools are designed and implemented to uphold the integrity and credibility of the legislative process. The insights and experiences gained through early pilots will be invaluable in guiding the future development and regulation of hybrid AI and LLMs in the legal sector, ultimately leading to more efficient, transparent, and reliable legislative processes.

5. Roadmap

5.1. Description

This section provides a high-level roadmap for the implementation and the deployment of smart functionalities. The iterative approach extensively uses PoC solutions, captures lessons learned, and runs projects for the full-scale development of an augmented LEOS platform adopting an agile software development approach. In this context, the interaction with users/stakeholders to continuously obtain and act on feedback throughout the process is considered important.

The roadmap consists of four components as follows: (i) setup the framework, (ii) explore technologies and functionalities, (iii) implement projects, and (iv) deploy in operations. Each of these components is detailed in the subsequent sections.

5.2. Setup the framework

It is imperative to define a well-established framework for the project, set expectations right, not overpromise, build trust, and aim for concrete added value in a relatively short time. The proposed framework consists of two main elements, *enable* and *govern*.

The actions to enable the projects should (i) ensure compliance with the EU AI act and any relevant guidance including contributing to the latter, (ii) facilitate the organisational setup including stakeholder engagement, and (iii) examine the technical preparedness. The following actions are proposed:

- Examine privacy and security considerations, and their alignment with the EU AI Act, as part of the preparatory phase for the implementation projects for smart functionalities;
- Adopt, customise existing or develop new guidelines, and address legal and ethical considerations to align with the specific context of the EU, and lay the groundwork for the use of the LEOS smart functionalities;
- Assess organisational readiness, develop a change management plan and trainings to support the use of the LEOS smart functionalities;
- Establish a project team composed of legal experts, developers, AI specialists, UX designers, and project managers, and define roles;

- Conduct a comprehensive evaluation of the current technology framework and ensure alignment and consistency with the proposed “5 + 2 technologies”.

To implement an augmented LEOS platform, it is necessary to launch a properly resourced, well-governed initiative. The following actions are proposed:

- Extend the current governance structure of LEOS to additional relevant stakeholders, develop transparent effective procedures and keep them under constant review, ensure discussion and agreement on priorities, planning and developments;
- Promote community-building leveraging the existing community within the EU’s JoinUp platform, engage with stakeholders and continue to identify and on-board key stakeholders within the public administration, disseminate project updates and outcomes, establish feedback loops with legal professionals, administrators, and end-users.

Considering the specific public sector context, three main aspects need to be catered for:

Enhance multi-disciplinarity and collaborative efforts

The collaboration across diverse domains such as law, technology, government, and linguistics, and among practitioners and scholars from the private and public sectors across the EU is essential for a holistic approach to legislative drafting. To this end, a dedicated task force could be formed comprising experts from relevant disciplines to drive forward initiatives and coordinate collaborative efforts.

Highlight the importance of broad and timely consultation

It is necessary to engage with users effectively, ensuring their buy-in, and gain thorough understanding of concerns and opportunities. Such a collaborative approach may have practical implications for the effective implementation and roll-out. Guidance documents to assist practitioners in navigating the complexities of legislative drafting in the digital age, addressing ethical considerations and other relevant aspects should be developed.

Exploit existing and emerging technologies

Harnessing the potential of readily available AI to improve legislative drafting processes should be a priority. This involves using existing open-source components to gain efficiencies and improve quality. Pilot initiatives can be used to test innovative approaches, with regular evaluation to refine strategies and adapt to changing

needs. This should include the integration of LLMs, drawing attention to their emerging applications and rapid evolution.

5.3. Explore technologies and functionalities

Two main tracks are proposed to explore the technological landscape further: (i) conduct studies to prepare the ground for implementation, and (ii) carry out PoC, among other things, to experiment with new technologies and critically evaluate their potential use in circumstances close to reality. It is posited that while agile methodologies have their merits in software development, projects involving legal frameworks, regulatory compliance, and complex data processing can benefit from a more adaptive and experimental approach, leveraging trials and errors, piloting and scaling to achieve positive results and stakeholder satisfaction.

Given the state of progress of NLG and its likely evolution, it is logical that the exploration of the technology strand initially focuses on the selection of a fundamental model. This model should be at the forefront of AI developments, capable of understanding and generating complex legal language. Evaluation criteria will include architectural considerations, performance metrics, and adaptability to various language contexts and tasks. The model must support on-premises deployment to ensure that data remains under EC jurisdiction. Given the sensitive nature of data, compliance with data protection and intellectual property regulations is the top priority.

When it comes to piloting, it is proposed to build on the experience gained and results obtained in an on-going study on the “Proof of Concept of Context-Aware Legal Verification,” and to:

- Develop multiple PoC solutions for an augmented LEOS, with each focusing on a distinct smart functionality cluster;
- Develop and ensure seamless access to various relevant data stacks through an augmented LEOS system, with provisions for processing the underlying datasets in accordance with legal document standards;
- Establish and validate solutions within interoperability regulatory sandboxes to enhance data and system security during the development of augmented LEOS.

5.4. Implement projects

Within the above framework, several software projects could be launched in parallel when conditions “are right”, i.e., when, for example, sufficient experience/understanding of the functionality/technology has been acquired, appropriate resources are identified and set aside, the right set of skills are available, and a clear need for the project has been formulated. These software projects can be (i) implemented in-house, (ii) procured, or (iii) implemented within the LEOS Community. In any case, when making decisions regarding new digital solutions, the dual pillar approach ([15]) should be followed, i.e., consider first reuse, then buy, and, as a last option, build. In the same spirit, an open-source software approach should be favoured. This will facilitate the reuse of solutions, co-creation, and sharing of results between departments and with other European public administrations.

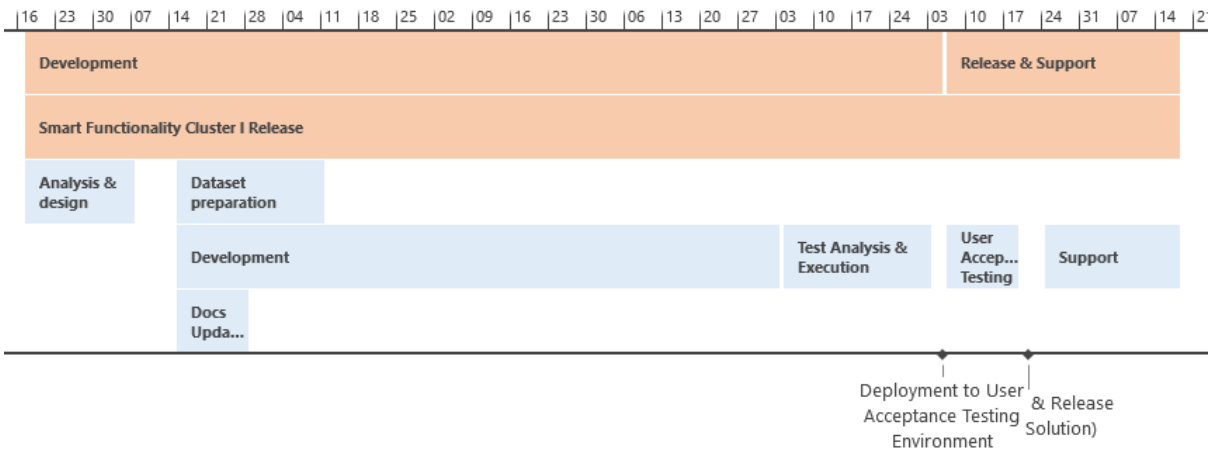
Specifically for software projects, development will need to follow a best practice approach ([16]) as shown in *Figure 3*. At this stage, it is not possible to assign precise timelines, as these will depend on factors such as budgetary constraints and system complexity. Given the innovative and evolving nature of the task, it is advisable to treat developments as separate contracts and involve multiple internal and external stakeholders. This approach would further ensure the exploitation of specific expertise tailored to the unique requirements of each project.

Development and preparation of datasets can begin as soon as the project is confirmed. Regarding the preparation of datasets, it is crucial to assess the required data sources and allocate significant resources to develop structured, validated, and open datasets, preferably in a standardised AKN-based format.

([15]) See 70703206-2592-4175-b10d-12f97382094a_en (europa.eu)

([16]) Using, for instance, PM2 project management methodology. See https://pm2.europa.eu/index_en

Figure 2 – Indicative approach for the implementation of software projects.



Software projects will follow an agile approach, with project management and active governance to oversee the strategic direction of the project, manage risks and resource allocation, and to ensure compliance with deadlines, specifications, and regulatory requirements, while optimising project outcomes and facilitating effective decision-making. Testing, including User Acceptance Testing (UAT), and documentation processes should be an integral and important part of the software project. Considering the potential use of augmented LEOS and its components by various actors, attributes such as scale, configurability, and customisation will have to be “designed-in” in order to adapt to different legal contexts and document types or to meet specific needs of different public administration departments.

5.5. Deploy in operations

Smart functionalities’ roll-out is closely linked to business feasibility and requires a staged deployment strategy. This will include organising training sessions for end-users and administrators, providing monitoring tools to track the performance of the system, establishing channels for continuous feedback from users and for informing them, amongst others, on updates, and creating comprehensive documentation for end-users, administrators, and developers.

Emphasis should be placed on cultivating trust among stakeholders in AI-based systems of this nature. Yet the notion of trust is not sufficiently defined in the current context of AI-supported lawmaking. Trust depends, among others, on providing high-quality, diverse, representative, and free from bias legal datasets. There are probably no objective criteria for deciding on a sufficient level of trust. The following activities can help developing trust:

- Build expertise and share experiences in the use of generative AI technologies;
- Develop and continually review practical guidelines to ensure responsible and ethical use;
- Implement measures to verify the provenance, transparency, and reliability of data; and
- Involve senior management in decision-making on the actual use of AI and LLMs in particular.

Ultimately, the readiness of the organisation, encompassing its stakeholders and staff, must be assessed. Such measures are closely linked to considerable investments in human resources to develop the necessary skills, while creating a culture of openness to the adoption of AI technologies. In any case, as already indicated, compliance with relevant regulations, such as the EU AI Act, has to be ensured to mitigate legal risks and ensure deployment of ethical AI systems.

6. Conclusions and outlook

The current publication reports and builds on the main outcome of an eight-month technical study entitled “Overview of Smart Functionalities in Drafting Legislation in LEOS”. It is centred around a prioritised list of 11 smart functionalities, selected based on the preferences of interviewed EC experts. Efforts concentrated on the discussion of five main AI technologies essential for enabling and implementing these smart functionalities, including the exploration of the use of LLMs in legal drafting.

The study's approach and insights represent a significant step towards an "augmented LEOS". The study amongst others stresses the importance of legal interoperability and aligning with regulations like the EU AI Act. LLMs and foundational models will potentially play a pivotal role in enhancing the functionalities of LEOS. Interestingly, however, the study assesses that most of the prioritised smart functionalities can be implemented without relying on generative AI. This finding thus underscores the importance of exploring mature AI technologies.

One of the main goals of the study was to investigate the integration of smart functionalities within the LEOS system. For this, the study argues that the reliance on proprietary technologies should be reduced, opting instead for an open-source approach. The strategic (re)use of existing open-source toolsets, including those funded by the EC, can expedite developments and ensure seamless integration with legacy systems. Nonetheless, the practicality and cost-effectiveness of integrating such technologies into LEOS depend on strategic choices and the specific approaches that will be selected, e.g., private vs. public cloud, how to go about using and building of datasets, and how to train and reskill staff. In any case, there is a need for flexible customisation and scaling for effective use in large and diverse organisations.

The study argues to commence immediately with the implementation of a select set of functionalities to mitigate potential risks arising from rapidly evolving technologies and changing requirements, This will allow for early experience building and deployment of in-demand features. Such a strategy must establish valuable feedback loops, offer critical insights, and guide further developments.

The AI technologies identified have reached a significant level of maturity and can be readily implemented in LEOS. However, the dynamic nature of AI suggests the necessity to invest significantly in the continuous monitoring of the technological landscape. This fast-paced evolution also highlights the importance of adopting a PoC approach and to only proceed with a standard agile development approach ‘at scale’ upon confirming the viability of the piloted solutions. Moreover, alignment with the recent Interoperable Europe Act and its agenda should be explored, including the exploration of interoperability regulatory sandboxes.

Inevitably, the report touched on the critical role of an open-source strategy, foundational models, and the importance of open standards. An open-source strategy enhances transparency and trust that are crucial in the public sector. However, an open-source approach also brings challenges such as the need for professional support and robust documentation, as well as security and privacy concerns.

In the above context, the study calls for immediate action.

References

Blair-Stanek, A., Holzenberger, N., & Van Durme, 'BLT: Can Large Language Models Handle Basic Legal Text?', preprint, arXiv:2311.09693, 2023. https://arxiv.org/abs/2311.09693

Bommarito, M. J., & Katz, D. M., 'GPT Takes the Bar Exam', preprint, arXiv:2212.14402, 2021. https://arxiv.org/abs/2212.14402

Branting, K., Weiss, B., Brown, B., Pfeifer, C., Chakraborty, A., Ferro, L., Pfaff, M., & Yeh, A., 'Semi-Supervised Methods for Explainable Legal Prediction', in *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*, Association for Computing Machinery, NY, 2019, pp. 22-31. https://doi.org/10.1145/3322640.3326723

Chalkidis, I., Androutsopoulos, I., & Aletras, N., 'Neural Legal Judgment Prediction in English', preprint, arXiv: 1906.02059, 2019. https://arxiv.org/abs/1906.02059

Clavié, B., Gheewala, A., Briton, P., Alphonsus, M., Laabiyad, R., & Piccoli, F., 'LegaLMFiT: Efficient Short Legal Text Classification with LSTM Language Model Pre-Training', preprint, arXiv:2109.00993, 2021. https://arxiv.org/abs/2109.00993

Davies, M., 'The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights', *International Journal of Corpus Linguistics*, Vol. 14, No 2, 2009, pp. 159-190. https://doi.org/10.1075/ijcl.14.2.02dav

Douka, S., Abdine, M., Vazirgiannis, M., Hamdani, R., & Amariles, D., 'JuriBERT: A Masked-Language Model Adaptation for French Legal Text', preprint, arXiv:2110.01485, 2021. https://arxiv.org/abs/2110.01485

Fitsilis, F., Mikros, G., & Leventis, S., 'Overview of Smart Functionalities in Drafting Legislation in LEOS', Joinup, 2024. https://joinup.ec.europa.eu/sites/default/files/document/2024-05/Augmented%20LEOS%20with%20smart%20functionalities%20final%20report%20checked.pdf

Guha et al., 'LegalBench: A Collaboratively Built Benchmark for Measuring Legal Reasoning in Large Language Models', preprint, arXiv:2308.11462, 2023. https://arxiv.org/abs/2308.11462

Hajlaoui, N., Kolovratnik, D., Väyrynen, J., Steinberger, R., & Varga, D., 'DCEP-Digital Corpus of the European Parliament', in N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Ninth International Conference on*

Language Resources and Evaluation (LREC'14). European Language Resources Association (ELRA), 2014, pp. 3164-3171. http://www.lrec-conf.org/proceedings/lrec2014/pdf/943_Paper.pdf

Hosseini, M., Resnik, D. B., & Holmes, K., 'The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts', *Research Ethics*, Vol. *19*, No 4, 2023, pp. 449-465. https://doi.org/10.1177/17470161231180449

Li, H., Ai, Q., Chen, J., Dong, Q., Wu, Y., Liu, Y., Chen, C., & Tian, Q., 'SAILER: Structure-aware Pre-trained Language Model for Legal Case Retrieval', preprint, arXiv:2304.11370, 2023. https://arxiv.org/abs/2304.11370

Mamalis, M., Kalampokis, E., Fitsilis, F., Theodorakopoulos, G., & Tarabanis, K., 'A Large Language Model based legal assistant for governance applications', preprint, OSF, 2024, https://doi.org/10.31219/osf.io/94fmq

Nguyen, H., Fungwacharakorn, W., & Satoh, K., 'Enhancing Logical Reasoning in Large Language Models to Facilitate Legal Applications', preprint, arXiv:2311.13095, 2023. https://arxiv.org/abs/2311.13095

Palmirani, M., Sovrano, F., Liga, D., Sapienza, S., & Vitali, F., 'Hybrid AI Framework for Legal Analysis of the EU Legislation Corrigenda', in E. Schweighofer (Ed.), *Legal Knowledge and Information Systems*, IOS Press, 2021, pp. 68-75. https://doi.org/10.3233/FAIA210319

Palmirani, M., Vitali, F., Van Puymbroeck, W., & Durango, F., 'Legal Drafting in the Era of Artificial Intelligence and Digitisation', Directorate-General for Informatics Solutions for Legislation, Policy & HR, European Commission, Brussels, 2022. https://joinup.ec.europa.eu/sites/default/files/document/2022-06/Drafting%20legislation%20in%20the%20era%20of%20AI%20and%20digitisation%20–%20study.pdf

Sandrelli, A., & Bendazzoli, C., 'Tagging a Corpus of Interpreted Speeches: the European Parliament Interpreting Corpus (EPIC)', in N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk, & D. Tapias (Eds.), *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, European Language Resources Association (ELRA), 2006, pp. 647-652. http://www.lrec-conf.org/proceedings/lrec2006/pdf/174_pdf.pdf

Savelka, J., & Ashley, K. D., 'The unreasonable effectiveness of large language models in zero-shot semantic annotation of legal texts [Original Research]', *Frontiers in Artificial Intelligence*, Vol. 6, 2023. https://doi.org/10.3389/frai.2023.1279794

Steenhuis, Q., Colarusso, D, & Willey, B., 'Weaving Pathways for Justice with GPT: LLM-driven automated drafting of interactive legal applications', preprint, arXiv: 2312.09198, 2023. https://arxiv.org/abs/2312.09198

Sun, Z., 'A Short Survey of Viewing Large Language Models in Legal Aspect', preprint, arXiv:2303.091362023, 2023. https://arxiv.org/abs/2303.09136

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D., 'Chain of thought prompting elicits reasoning in large language models', preprint, arXiv:2201.11903, 2022. https://arxiv.org/abs/2201.11903

Wiratunga, N. et al., 'CBR-RAG: Case-Based Reasoning for Retrieval Augmented Generation in LLMs for Legal Question Answering', preprint' arXiv:2404.04302, 2024. https://arxiv.org/abs/2404.04302

Xu, H., & Ashley, K., 'Question-Answering Approach to Evaluating Legal Summaries', preprint, arXiv:2309.15016, 2023. https://arxiv.org/abs/2309.15016

Yu, F., Quartey, L., & Schilder, F., 'Legal Prompting: Teaching a Language Model to Think Like a Lawyer', preprint, arXiv:2212.01326, 2022. https://arxiv.org/abs/2212.01326

Annex 1: Initial list of smart functionalities

Table 5 – Initial list of smart functionalities

#	Category	Smart functionality
1	Verification – Correct usage of …	Citations
2		Existing references
3		Acronyms, organisations, and other abbreviations
4	Verification – Context aware correct usage of …	Validity and relevance of references
5		Existing legal definitions
6		Specific legal lexicon
7	Granular change tracking – Comparison of documents	Modifications
8		Change Tracking
9	Linguistic support	Use correct linguistic formulations within the structure of the document
10		Correct formulation in accordance with the English Style Guide
11		Detect divergences between different linguistic translations
12		Suggest linguistic formulations in provisions
13	Legal Assistance – within the act	Detect and avoid structures that could create issues in legal interpretation
14		Correlation between recitals and the enacting terms
15		Between previous acts and the new one
16		Incompatibilities in temporal parameters
17		Explicit or implied obligations
18		Detect implicit or incomplete modifications
19		Detect obligations, rights, permissions, penalties
20	Legal Assistance – within the legal corpus	Automatically identify existing legislation relevant for the act under development
21		Identify hidden semantic correlations

#	Category	Smart functionality
22		Detect suspended, repealed, derogated, delegation of power
23		Passive and active references
24		Life cycle of an article
25	Support ‘automatic’ legal drafting	Drafting transitional measures
26		Large Language Model (LLM) based legal text generation
27		Construct the consolidation text applying amendments
28	Policy dimension	Measure impact of a legislative act
29		Consistency in definitions
30		Repository of legal knowledge
31		Cluster legislative documents
32	Discovery of Practices/Enabling	Automatically extract metadata
33		Classification of corrigenda
34		Discover concrete practices of different styles of drafting

Annex 2: Additional list of smart functionalities following expert consultation

Table 6 – Additional list of smart functionalities

#	New smart functionality	Category
1	Automatic recognition whether a given text belongs to a recital or into the explanatory memorandum	Legal assistance
2	Smart templates - Detection/suggestion of the proper legal template for any legal text	Legal assistance
3	Detection/evaluation of the legal basis (or bases) in view of the content of the document	Legal assistance
4	Coherency check if an act transposed correctly or in line with union, international obligations etc.	Policy dimension
5	Create a smart search facility	Legal assistance
6	Functions allowing to visualise information out of a basic act	Policy dimension
7	Automatically draft legal text using imported text from identified data sources	Automated drafting
8	Create and update a database of legal bases	Verification
9	Avoid common errors based on the predictability of drafting customs	Automated drafting
10	Create a table of content of any act	Legal practices
11	Summarisation of large legal texts	Policy dimension
12	Create and update terminology databases	Verification
13	Automatic e-briefing (and other types of secondary text generation: reporting; fact sheets; Q&As; etc.)	Policy dimension
14	Terminology extraction tool	Legal practices
15	Keeping track of the origin of data	Change tracking
16	Style and quality feature validation based on predefined rules and conventions (joint handbook, inter-institutional style guide, etc.)	Verification
17	Conduct interoperability assessments for digital ready legislation	Legal assistance
18	Legal processes visualisation & gamification	Policy dimension

#	New smart functionality	Category
19	Auto drafting from hints in track-changes and notes from the collaborators	Automated drafting
20	Detect liabilities	Legal assistance
21	Maintain style formatting in LEOS/EdiT when importing text (including comments) from external sources	Verification
22	Automatic switch from American to British English	Linguistic support
23	Presentation of examples, e.g., alternative dispute settlement procedures	Legal assistance
24	Watermark or entirely block printouts to prevent leaking	Policy dimension
25	Filter out/cluster style guide changes and accept them in one batch	Verification
26	Detect deviations in legal jargon and replace this in-house jargon	Linguistic support
27	Use of inclusive language in terms of gender, religion, disabilities, race etc.	Policy dimension

Annex 3: Attributes of categories

Table 7– Attributies by categories

	UX	Business value	Technology stack	Datasets	Performance
Legal verification	Non-intrusive Proactive & Reactive	Quality; Validated outcome	Local & external repos of legal data	Local & externsl repos of legal data	Disruption of drafting during verification
Change tracking	Side by side comparisons or inline visualisation	Unnoticed accidental changes	Existing current and past versions of drafted legal data	Existing current and past versions of drafted legal data	Support for large documents
Linguistic support	Non-intrusive Proactice & Reactive Suggestions in modal form/side panels	Quality; Validated outcome	Linguistic ref data	Linguistic ref data	Disruption of drafting during data set retrievals
Legal assistance			Legal ref data Cross referencing	Legal ref data Cross referencing	
Automated drafting		Consistency; Out of date data	Repositories of prebuilt templates & amendments	Repositories of peebuilt templates & amendments	
Policy dimension		Different policy perspectives	Policies & guidelines repos	Policies & guidelines repos	Infrastructure wokload
Legal practices		Quality; Algo reviewing; Review efforts	Predefined suggestions & patterns repos	Predefined suggestions & patterns repos	Disruption of drafting during verification

Getting in touch with the EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696,
- via the following form: european-union.europa.eu/contact-eu/write-us_en.

Finding information about the EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

EU open data

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

Publications Office
of the European Union