Summer School LEX2024

Overview of Smart Functionalities in LEOS

Dr. Fotis Fitsilis, Hellenic Parliament

7 September 2024

#LEX2024

European Commission

HELLENIC OCR TEAM

HELLENIC PARLIAMENT

Structure

- Context
- The study & its objectives
- Stakeholders
- Tasks & timeline
- Approach
  - Data collection
  - Questionnaire
  - Interviews
  - Categories & attributes
- Main results

Context
Drafting legislation at the core of any public organisation

Improve legislation

- Quality
- Efficiency
- Effectiveness
- Efficacity

Using

- Readily available IT, including AI
- Digital Transformation

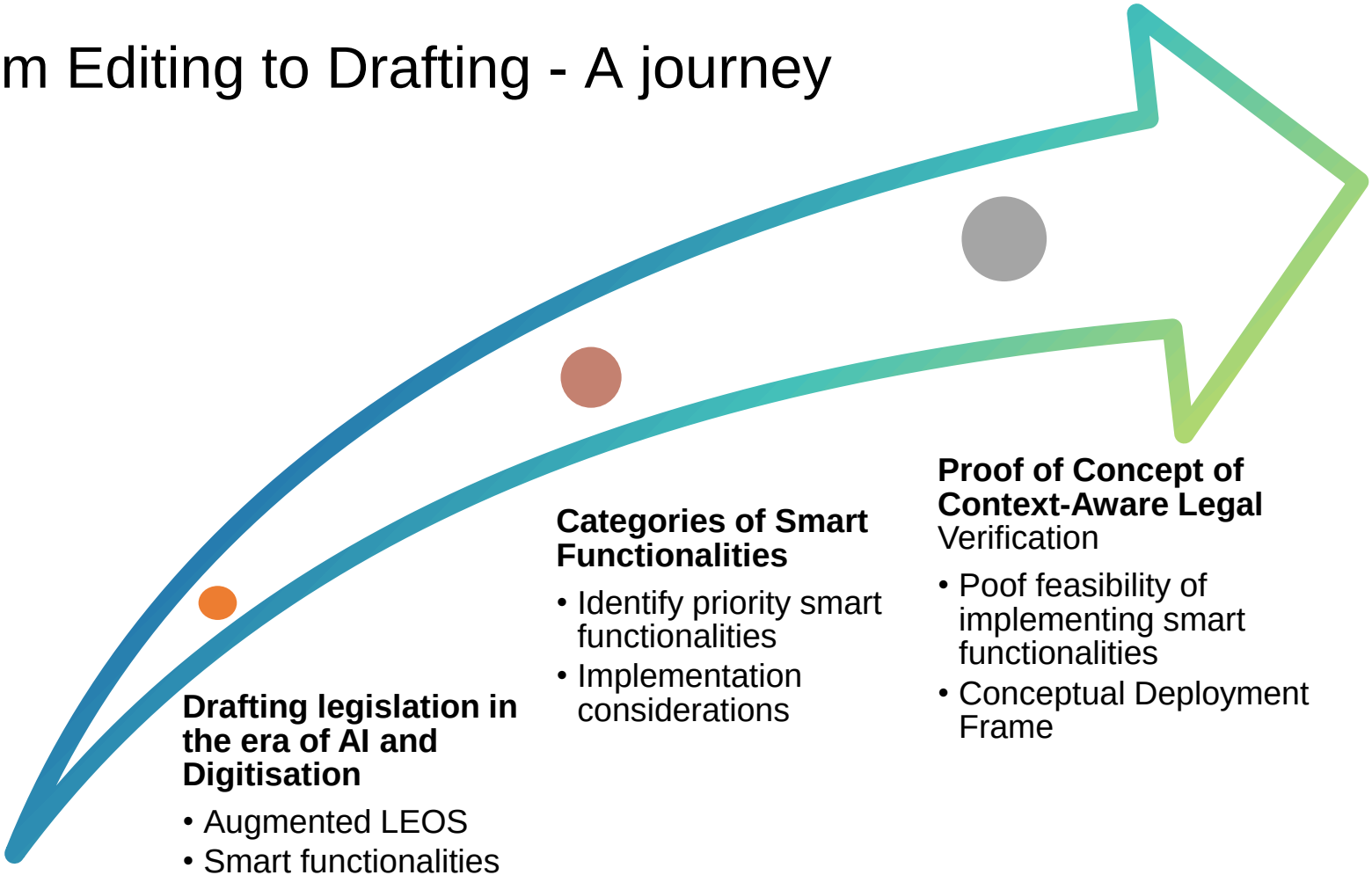
And addressing EU specificity

- Multi-stakeholder
- Subsidiarity
- …

In the near future

…

From Editing to Drafting - A journey



Drafting legislation in the era of AI and Digitisation

- Augmented LEOS
- Smart functionalities

Categories of Smart Functionalities

- Identify priority smart functionalities
- Implementation considerations

Proof of Concept of Context-Aware Legal Verification

- Poof feasibility of implementing smart functionalities
- Conceptual Deployment Frame

The study

- Towards an ‘augmented LEOS’
- Duration: June ‘23 - January ‘24
  8 months
- Break down: 4 tasks/objectives

Objectives

- Further group & specify categories of smart functionalities
- Analyse & document their added/business value
- Identify, detail & explore the techno-business feasibility of implementation/deployment
- Provide a high-level roadmap

Stakeholders

Implementation

- Fotis Fitsilis, project manager
- Sotiris Leventis, ICT expert
- George Mikros, AI expert
- Hellenic OCR Team, support

Contracting authority - DG DIGIT

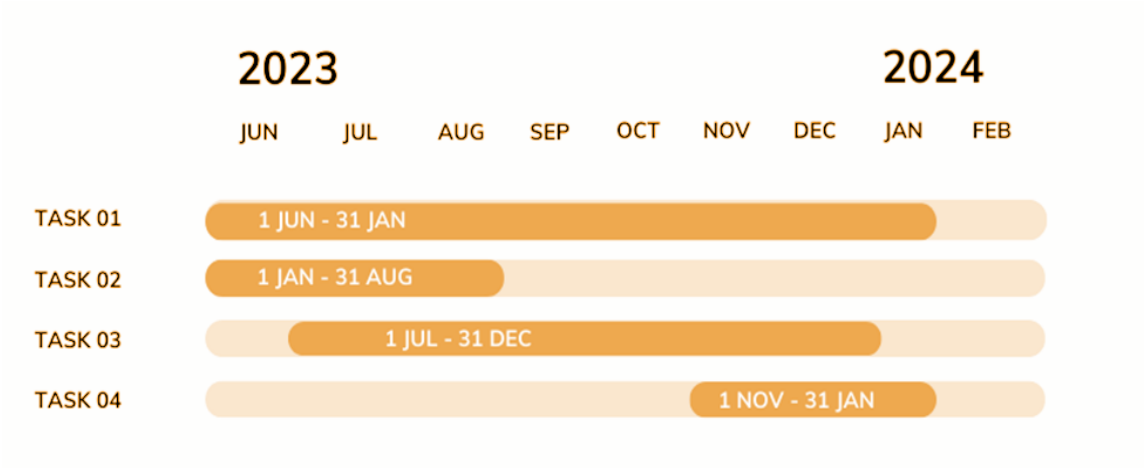
- Fernando Nubla Durango
- Willy van Puymbroeck

Monitoring - SG

- Cristina Stanciulescu

Tasks & timeline

Task-01: Project Management
Task-02: Elaborate on categories of Smart Functionalities
Task-03: Assess the business value & techno-business feasibility of the implementation & deployment of categories of smart functionalities
Task-04: Provide a high-level roadmap



Data collection

Desk review

- Reference study: Legal Drafting in the Era of AI and Digitisation
- Literature study and assessment

Field data

- Set of 11 interviews (July to September 2023)
- Technical workshop on LEOS (27 July 2023)
- Questionnaire on original smart functionalities (response rate: 73%)

Evaluation / questionnaire

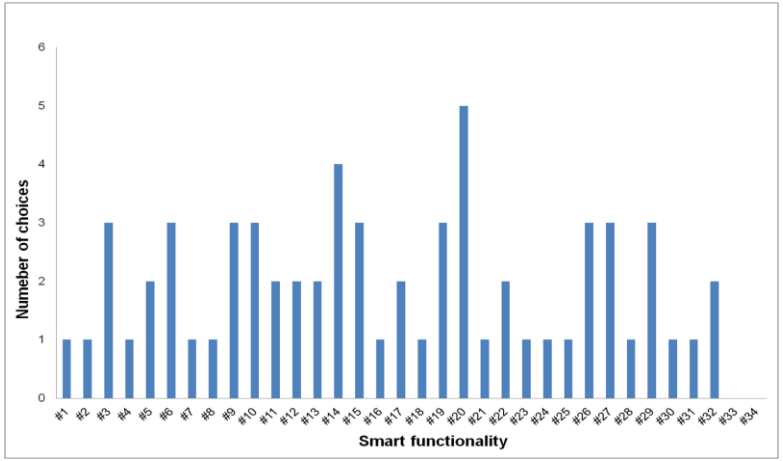
Smart functionalities

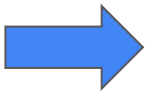
Interviewees	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	#23	#24	#25	#26	#27	#28	#29	#30	#31	#32	#33	#34
#1																	o		o	o						o		o						
#2														o	o					o		o				o	o		o			o		
#3									o				o													o	o		o					
#4	o	o	o	o	o	o			o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o										
#5																			o	o										o		o		
#6			o						o	o	o	o		o	o										o		o		o		o			
#7					o	o								o						o														
#8			o			o	o	o		o																								
#9							o	o	o	o	o	o	o	o											o	o								o
#10	o	o	o				o				o	o	o	o	o	o			o	o			o	o								o	o	o
#11																																		

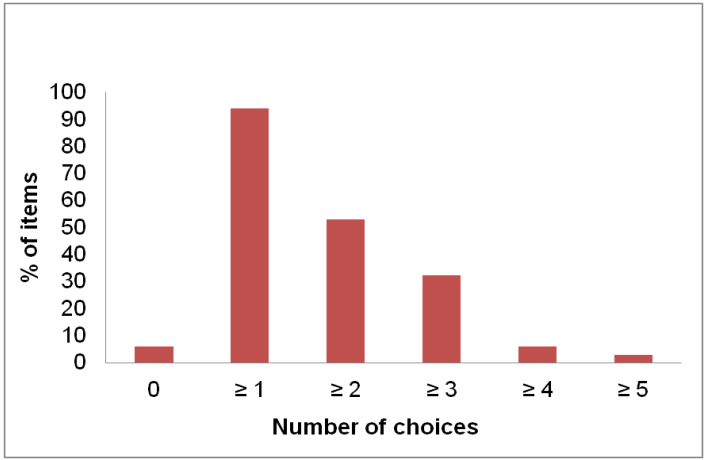
Note: #11 - no response

Prioritisation of smart functionalities

Evaluation / questionnaire







Evaluation / top and no picks

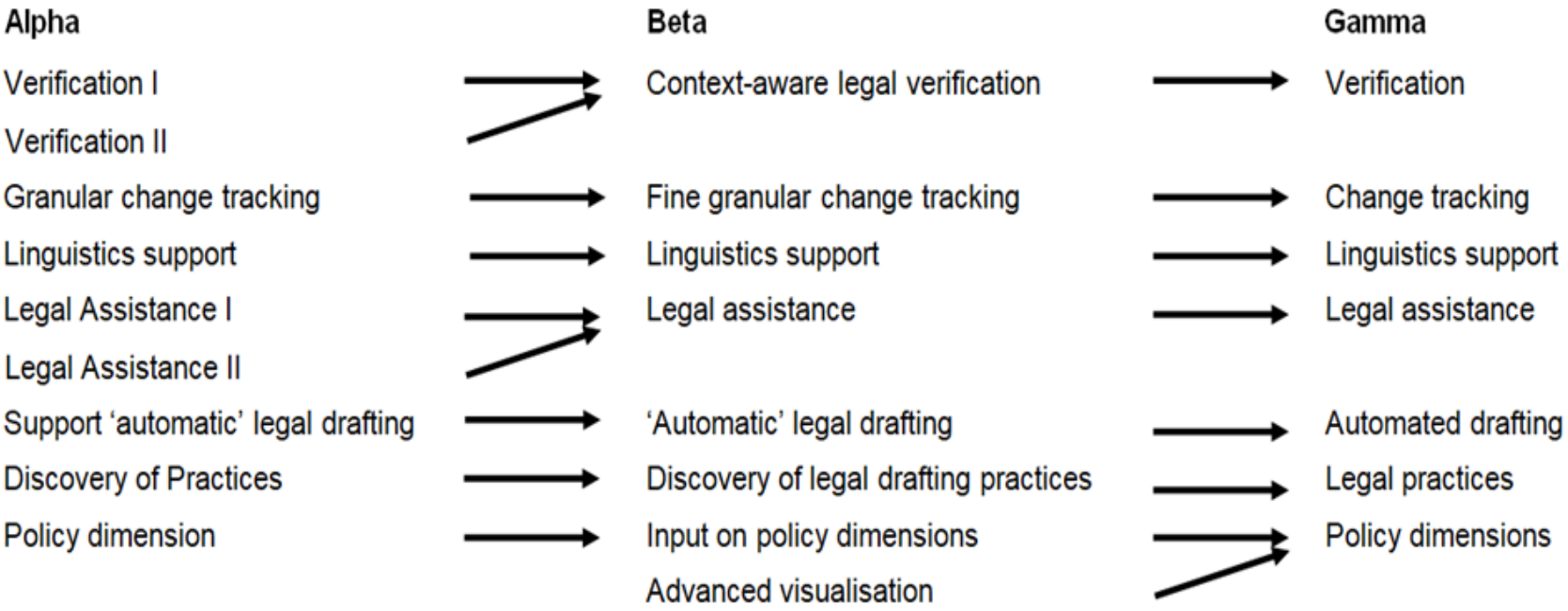
Top picks

- #14: Correlation between recitals and the enacting terms
- #20: Automatically identify existing legislation relevant for the act under development

No picks

- #33: Classification of corrigenda
- #34: Discover concrete practices of different styles of drafting

Categories / alpha, beta, and gamma



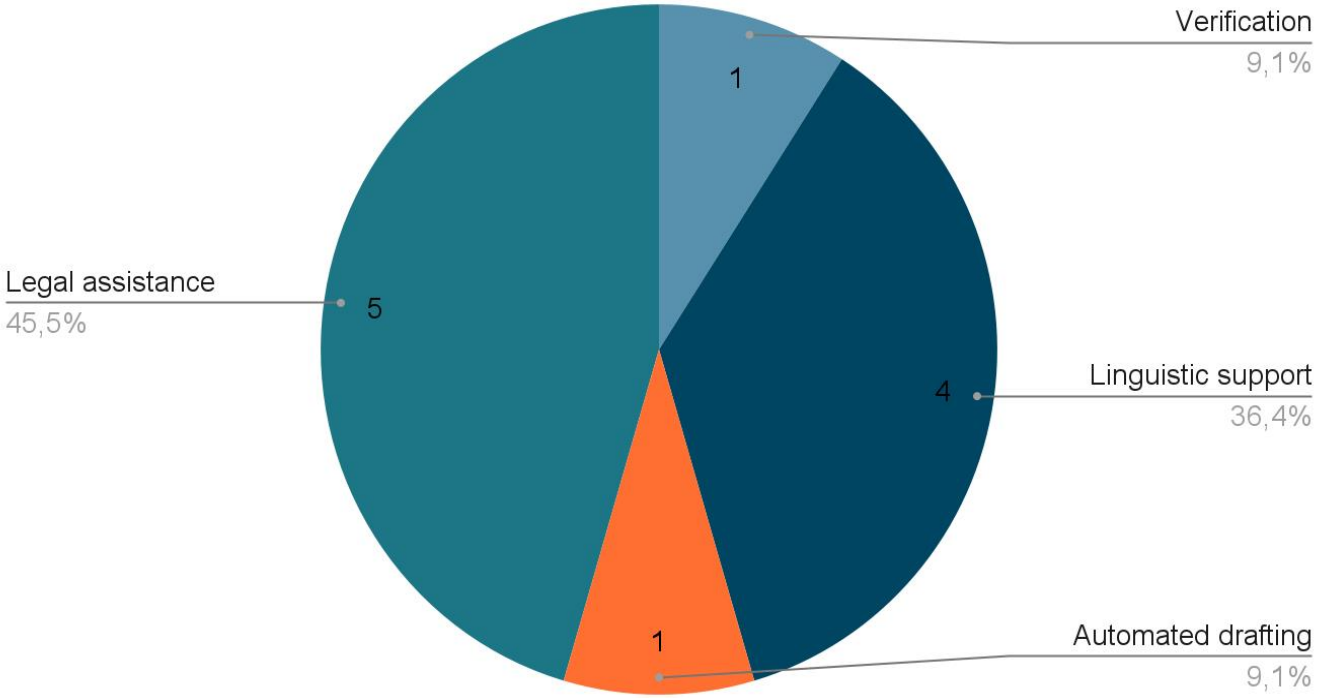
Interviews

- Audiovisual files digitally processed to extract accurate transcripts
- Technologies:
  - *Whisper v2 (Open AI)*
  - *Wav2vec2: ML technique for training speech recognition models*
  - *NeMo (NVidia) for distinguishing the speakers*
- Output anonymised and main points extracted
- Qualitative rather than quantitative analysis

Prioritization of Smart Functionalities

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
#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	Large Language Model (LLM) based legal text generation	

Categories of top picks



Business value and techno-business feasibility

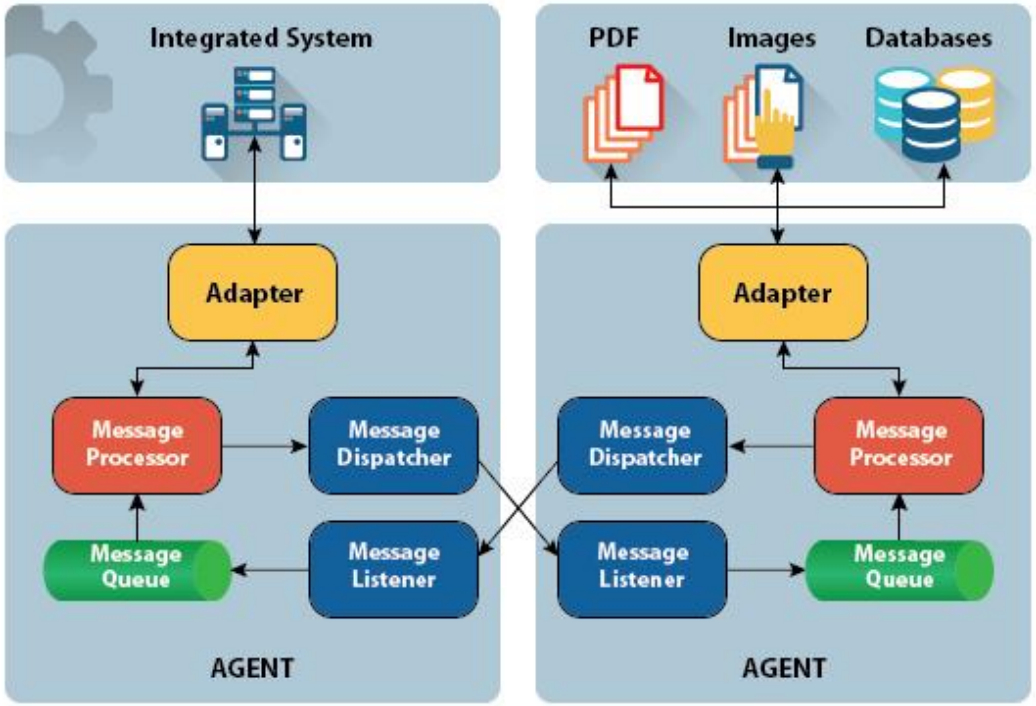
	UX	Business value	Technology stack	Data sets	Performance
Legal verification	Non-intrusive; Proactive & reactive	Quality; Validated outcome	UX event handlers; Asynchronous API calling & throttling	Local & external repos of legal data	Disruption of drafting during verification
Change tracking	Side by side comparisons or inline visualisation	Unnoticed accidental changes	Aligned front-end technologies; 3rd party visualisations	Existing current and past versions of drafted legal data	Support for large documents
Linguistic support	Non-intrusive; Proactive & reactive; Suggestions in modal form / side panels	Quality; Validated outcome	UX event handlers; Asynchronous API calling & throttling	Linguistic ref data	Disruption of drafting during data set retrievals
Legal assistance				Legal ref data Cross referencing	
Automated drafting		Consistency; Out of date data	On-premise LLM; Trusted vendor via secure API	Repositories of prebuilt templates & amendments	
Policy dimension		Different policy perspectives	Heavy processing in the business layer; API endpoint acting as aggregator to multiple repositories	Policies & guidelines repos	Infrastructure workload
Legal practices		Quality; Algo reviewing; Review efforts		Predefined suggestions & patterns repos	Disruption of drafting during verification

Technology analysis and clustering

- Five main technologies or the implementation of the prioritised smart functionalities
  - Advanced Language Editing and Correction (ALEC)
  - Named Entity Recognition (NER)
  - Semantic Similarity
  - Natural Language Generation (NLG)
  - Information Extraction

- Two additional technologies for the accommodation of the rest
  - Legal Ontology and Terminology Management (LOTM)
  - Text Classification

Integration technology



Basic integration architecture

Main results

- Develop a better understanding of the task
- Map priorities
- Match smart functionalities with technologies
- Highlight implementation issues
- Discuss considerations
- Develop a roadmap

Thank you!

fitsilisf@parliament.gr
http://fitsilis.gr