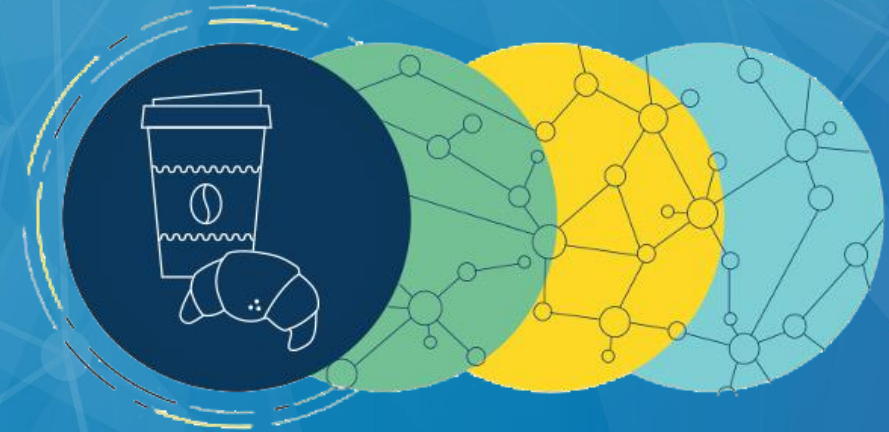


KEY TAKEAWAYS

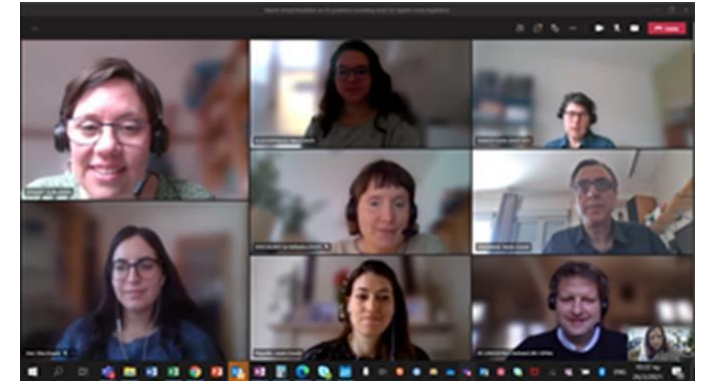


March virtual breakfast on
Better Legislation for Smoother Implementation:

AI-powered screening tools for digital-ready legislation

26 March 2021 from 9:30 to 11:00 CET

- / **01** Introduction
- / **02** AI-powered screening of legislation
- / **03** AI-powered pilots on screening legal texts: SeTA
- / **04** AI-powered pilots on screening legal texts: Manylaws
- / **05** Plenary session
- / **06** Upcoming



The Better Legislation for Smoother Implementation community

A living and continuously growing multi-disciplinary community on better legislation to share good practices and co-create tools with interoperability in mind around

- **Digital-ready policymaking**
- **Legislation and technology**
- **Streamlined regulatory reporting process in the EU**



The Legal Interoperability team



Zsofia Sziranyi



Cecile Guasch



Isa von Kalben





Bertrand DE LONGUEVILLE

SeTA - a Semantic Text Analysis tool

**European Commission
Joint Research Centre (JRC)**

Bertrand is a Team Leader Text Mining Analysis Competence Centre of the Joint Research Centre (JRC). Together with a team of computer linguists, data scientists and IT experts, he aims at delivering to EU institutions a wide range of tailor-made solutions to convert vast amounts of unstructured data into actionable knowledge.

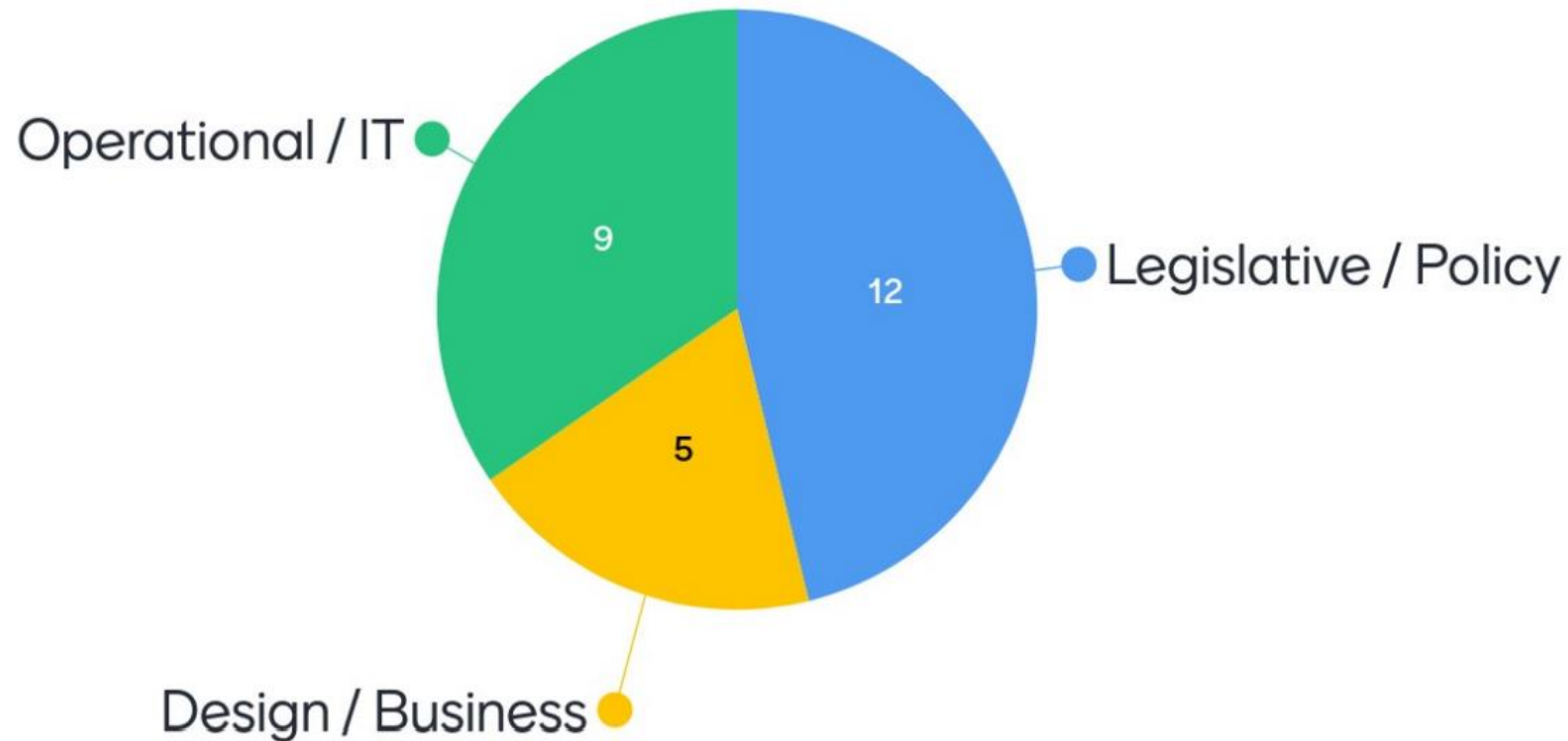


Yannis CHARALABIDIS

ManyLaws - EU-wide enhanced Access to Big Open Legal Data
**Professor of Digital Governance,
University of the Aegean**

Yannis is the founding director of the Digital Governance Research Centre, coordinating policy making, research and application projects for governments and enterprises worldwide.

Your professional background: do you see yourself in any of the below groups?



Provenance of the participants (organisation and country)

France	EC - Belgium	Salzburg State Parliament, Austria
Commission, Germany	Belgium, consultancy	JRC ISPRA
European Commission	EC, Luxembourg	EC - Sweden
Digitisation Agency, Denmark	EC, Belgium	portugal
Norway	Belgium	Belgium
European Commission, Spain	European Commission, Brussels	EC - Italy
Belgium, consultancy	Spain	Commission, Belgium
Luxembourg	European Commission, Germany	Commission, Germany

Participants' expectations:

Keeping up to date	Learn about possible practical application of AI	Learn about use cases and possibilities with opds
Learn more about AI tools	Discover synergies between projects	Food for thought
Learning more about AI and the uses for legal analysis	See practical implementations of AI in legal applications	To understand better the possibilities and limits of text mining.
listen	Have fun and interaction	Sharing!
Catch a glimpse of the future	Input and ideas	Info on your plans
Learn more about AI powered tools	Keep myself updated on new developments	To share best practices
Learn what are our users	EUR-lex search	Eurlex
EuroLex	LexisNexis	



AI-powered screening of legislation

[Link to recording](#)

The WHAT and the WHY

Our subject: AI-powered screening tools for digital-ready legislation

WHAT?

Keywords search

looking for specific terms;
highlighting excerpts in
legal texts; ...

Artificial intelligence

more than simple text
search; new insights with
big data (text mining
algorithms)

Digital-ready legislation

to be fostered by legal
screening results

WHY?

Legal screening
results

Better legislation

Smoother
implementation

Screening existing legislation

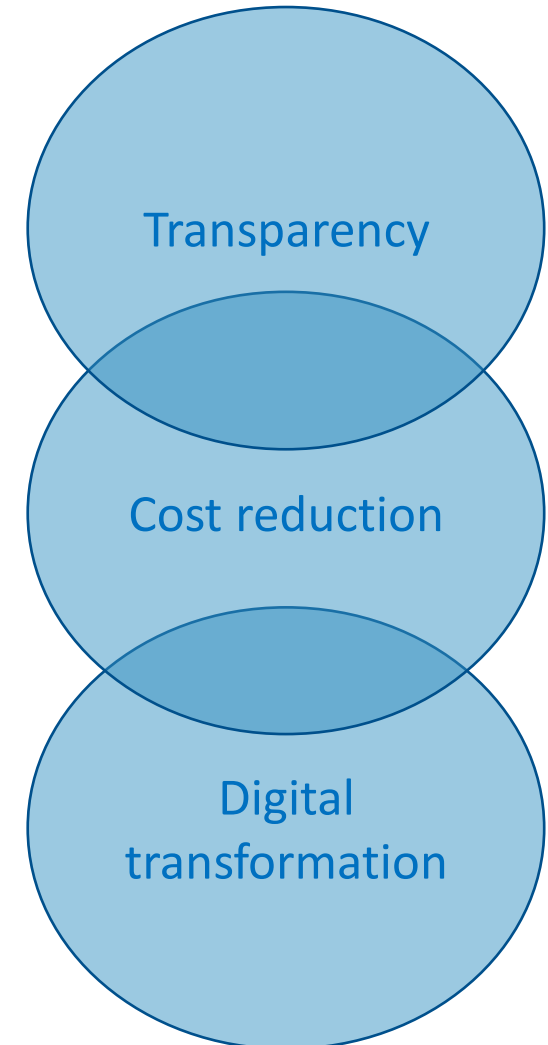
- Identify legal provisions applicable to a situation
- Assess the degree of transposition

Screening for streamlining reporting clauses

- Find already collected data & related reporting terms
- Identify burdensome reporting processes (fax, duplicate, ...)

Screening draft proposals – Digital-ready policymaking

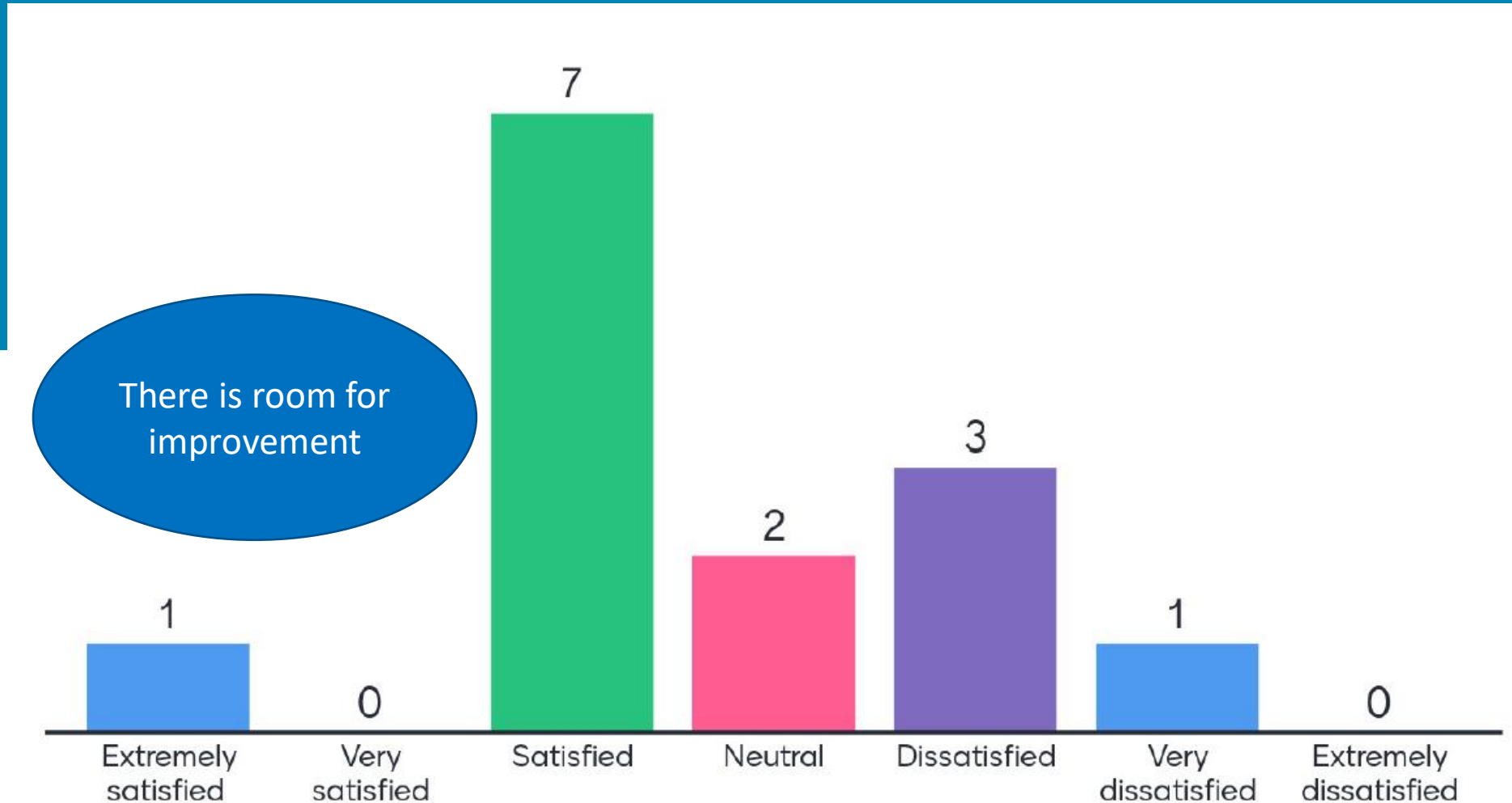
- Spot proposals with digital impacts / support needs
- Spot inconsistencies with digital initiatives



Services participants use searching for legal provisions:

Eur-lex	Eur-lex	EurLEX
Google	EURlex	Google
Eurlex advanced search	eur-lex	EUR Lex as EU system
lexisnexis	Google	Eurlex,
Curia	EUR-Lex	EurLex, RIS (.at), var. .de services
RCD		

Participants' satisfaction with their searching results:



Semantic Text Analysis tool (SeTA)

Bertrand de Longueville

European Commission, the Join Research Centre

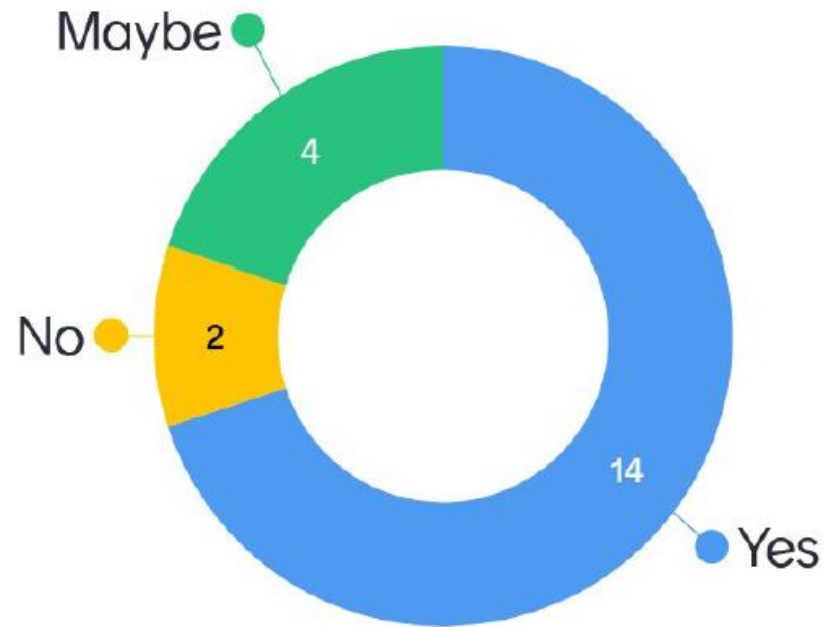
[Link to recording](#)

- **Semantic analysis** is key to deal with **massive amounts of information** – it allows converting raw information into knowledge
- The tool applies **semantic text analysis principles** in legislation (policy in the broader domain)
- Injected 500k EU legal documents and publications from the EU in SeTA
- The tool can understand **contextual aspects** and classify the information. It also helps to understand biases and extract key knowledge.
- There is a **massive processing chain** that extract terms and measures the distance between each other. This creates **concepts maps**.
- Right now, SeTA handles only English.

- Two key use cases were demonstrated during the session: **Search** that provides to the user means to identify excerpts of documents containing searched concepts or a similar document combining linked open data search with text mining algorithms and **concept browsing** providing a visual representation of concept similarity
- The tool is **European Commission internal** for now. And is looking for its way forward.
- There is a **report** available that explains in detail the SeTA tool [here](#).

See the specific presentation from SeTA available [here](#) for more details.

Potential of SeTA in its current version to help participants in their work:



Participants' ideas for improvement of the existing features or for additional functionalities of SeTA:

Yes	Retrieve case-law	Multi-lingualism
depends on how it works.	Compare to concepts	Would be useful if we could access it outside of the EC.
Links to most relevant source texts	Searches in different parts of documents individually	needs to be tested in reality
Extension to different language versions?	It would be very useful if we could go directly to the part of the document or legal text that is related to the key word.	Share GitHub link with public
I would like to spend more time on it.	Why not making this available as OSS?	uhmm

- **How does SETA make difference between broader or narrower concepts?**

There is no ontology in SeTA. But there is high complementarity with the ontology approach. JRC would like to make it possible for the user to establish his own ontology. The graph model has its value because it is flat and based on co-occurrence. We would like to use EU ontologies to extract knowledge from SeTA or to offer linked concepts to build ontologies. SeTA proposes rich sets of keywords and it can show that two words are linked but the meaning of the link is to be provided by the user.

- **Is there a commercial plan to make the tool publicly available?**

The mandate of JRC is to support the EC with the tool, but of course, it is paid by taxpayers so the EC would be eager to share it publicly. However, the “how” still needs to be decided, as many aspects need to be considered.

- **Any plan to publish the tool as Open Source Software (OSS)?**

This is one of the possible SeTA futures. However, this would require preparatory work to make the software "OSS-ready". The JRC would be happy to do this, if it would find the right context for it.

- **Does the algorithm measure quantitative qualities like occurrence or proximity and does it learn from search results?**

The algorithm measures the context similarity. It does not yet learn from the search results. Feedback loops and ML are not yet implemented, and information extraction is made from real-life examples of the end users.

ManyLaws

EU-wide Enhanced Access to Big Open Legal Data

Yannis Charalabidis

University of the Aegean, Greece

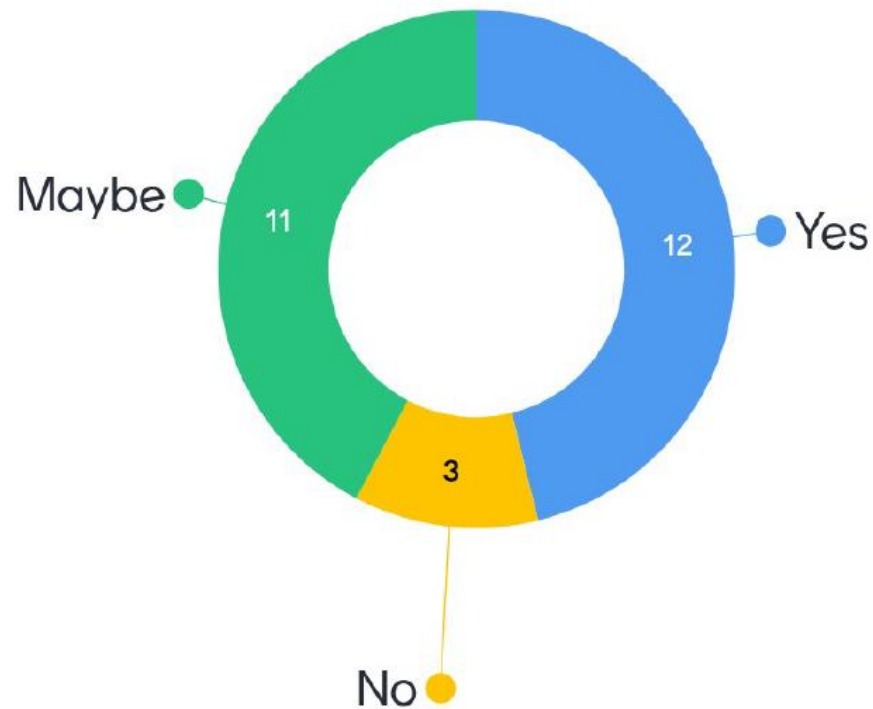
[Link to recording](#)

- Part of a call from CEF Telecom in order to **facilitate the access and use of legal information**.
- It is based on **textual** and not numerical data (the differentiation of the project)
- It covers **three languages** and **three sources**.
- All the machine processable documents are found via **EUR-LEX and 2 national databases** (GR, AU) and then, the processing phase starts.
- The code is deployed at a **Greek HPC supercomputing infrastructure** (ARIS) - 40k cores (1million times faster than a computer)
- The machine does certain processes **one by one** (multiple runs of this action), then the law is downloaded and becomes available in an **XML format**.
- It decomposes the law, process and **translates** the law **automatically** using an EU infrastructure (CEF DSI – eTranslation), which is not yet available to the public.
- Processed law will be published in the **EU Data portal as Open Data** (feed the open data portal). There are intermediate results that can be reused to make your own tool.

- The idea is to **make law accessible** to everyone. The first step is to retrieve current legal information, then identify the correlations and use automated translations and it also does an analysis of the citizens' comments on the specific legislation.
- The tool allows to follow the **evolution of legislations** over time.
- There is **no human intervention** in the process. If the algorithm does not work, it will be fixed.
- The tool is not only open, but also extendable.
- Here are some possibilities for the future of the tool:
 - Include case law
 - Include more countries
 - Extend functions, e.g. get closer to automated justice
 - Interconnecting with other tools, e.g. [Leos](#)
 - Further use the power of hyper computing

See the specific presentation from ManyLaws available [here](#) for more details.

Potential of ManyLaws in its current version to help participants in their work:



Participants' ideas for improvement of the existing features or for additional functionalities of SeTA:

Not clear what technologies are used only XML was mentioned

mobile functionality

integration in eurlex, leos...?

More languages

comparability of code and court decisions.

Ingest EU 27 legislation!

LEOS - Manylaws intergration

Extent to other parliamentary functions

Extend to more MS and languages

Go to the text level - finding keywords

More filters for results

Advance basic functionality

The more legislation it gets the better. For drafters it is very useful to look at other national law

- **What could be the relation of ManyLaws with EUR-Lex?**

Ideally, ManyLaws starts where legal processing and legal publishing finishes. EUR-Lex is supported by many actors at the same time. ManyLaws is much smaller than EUR-Lex and does not have the same capacity and resources.

The tools codify a predetermined sets of words that exist in our languages and we search our documents. ManyLaws would like to move towards a use case where the user would speak about the law and what is he looking for and ManyLaws would execute the request in the user's desired language.

- **When would an API be available?**

A REST API will be available in April.

- **Is there a parliamentary function?**

There is, with Greece and Austria, and it covers what happens before the law is adopted, e.g. the drafting, the consultation, etc.).



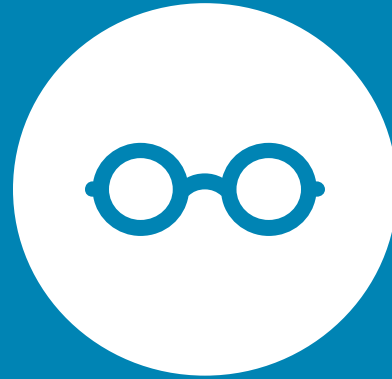
Plenary

- **Multilingualism** is a very important for a legal screening tool.
- It is interesting to explore legal concepts in **different countries** and **political contexts** (for instance gender equality).
- **Different interpretations** across all legislation and concepts.
- The assessment of **sub-document** level (like articles) of ManyLaws is very powerful.
- **Contextual analysis** is also very important and tackles the issue of understanding (e.g. when extracting the data and the subjects of the legislation).
- **Keyword lists** used in the SeTA are a wholly connected thing (graph creation).
- This is the kind of **lexicon** we need and a way to navigate across the not only thousands, but **billions of ideas and notions**.
- **Collaboration** between SeTA and ManyLaws – and possible other solutions – is key.

In brief, the session gave **food for thought** for both pilots and for participants!

Participants' ideas for usage of the presented tools:





Upcoming

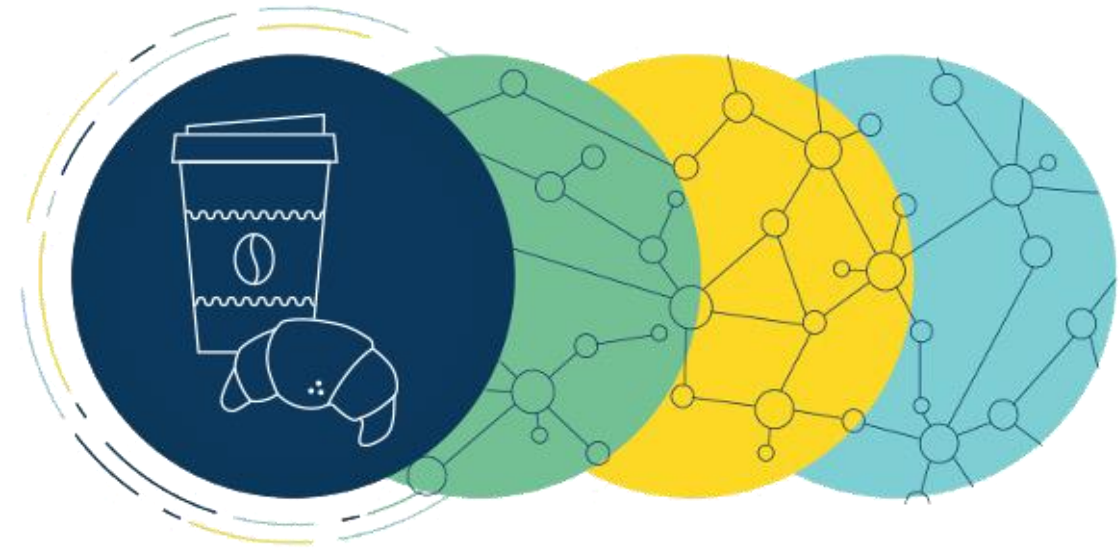
April virtual breakfast: Multidisciplinary policymaking

*Why **digital-ready** needs multidisciplinary teams?*

*What are **the multidisciplinary teams** and what is **their added-value in policymaking**?*

*What are **the challenges** to multidisciplinary policymaking and **how** to address them?*

*What **tools** can facilitate the work of multidisciplinary teams?*



The event will be held on **21 April from 9 am** as part of the [DigitAll conference](#).

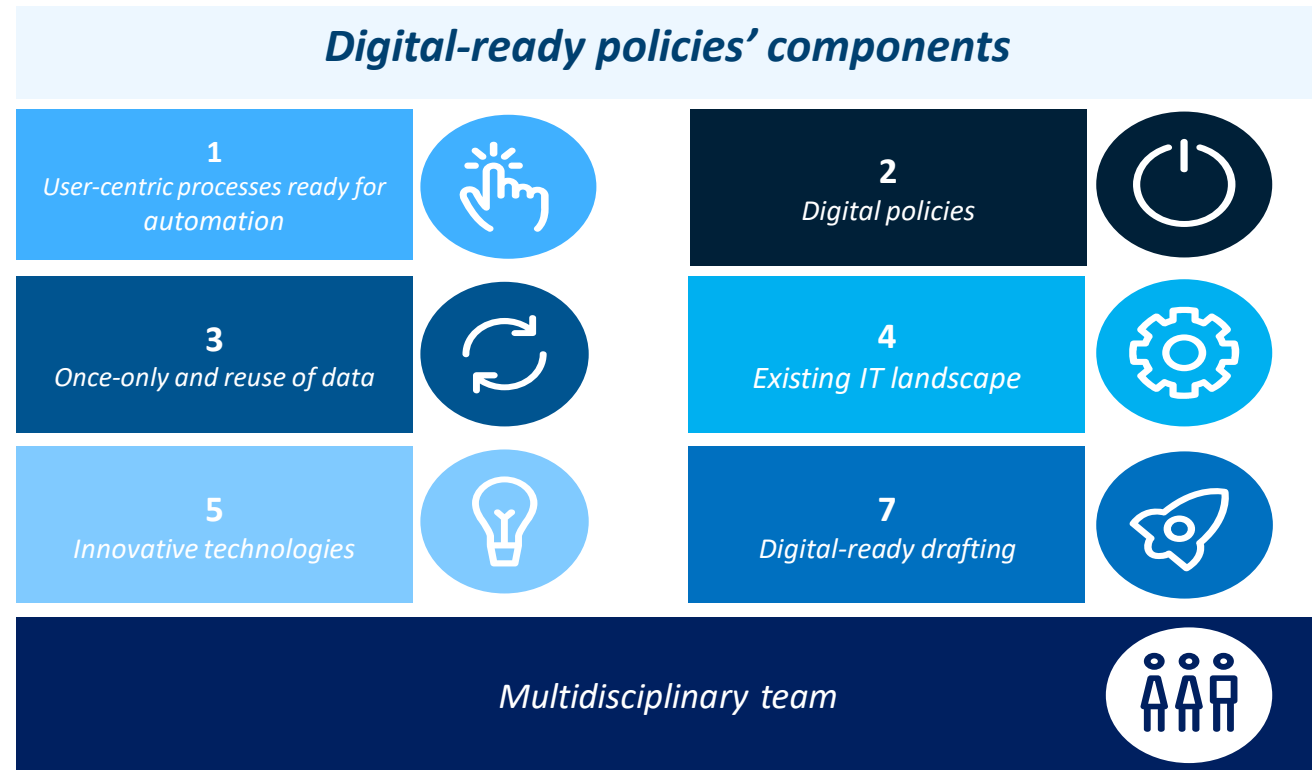
[Subscribe to our newsletter](#) to stay tuned on what is coming and [become a member](#) of our community!

Digital-ready policymaking on our BSLI community

We are currently developing our approach to **digital-ready policymaking**, which will feed into the ongoing revision of the European Commission's [Better Regulation Agenda](#).

Feel free to contribute:

- [Contact us](#) to share any thoughts and feedback
- [Become a member](#) of our BLSI community to join our [discussion directly on JoinUp!](#)



Shaping together

EIF



the future policy on

public sector

interoperability

TAKE PART
in the **PUBLIC**
CONSULTATION
until 26.04.21





ISA² programme

You click, we link!

Stay in touch
ec.europa.eu/isa2

 @ EU_isa2

 ISA² Programme



EU-legal-interopability@ec.europa.eu

Run by the Interoperability Unit at DIGIT (European Commission) with 131€M budget, the [ISA² programme](https://ec.europa.eu/isa2) provides public administrations, businesses and citizens with specifications and standards, software and services to reduce administrative burdens.