

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

Elena Montiel-Ponsoda¹, Víctor Rodríguez-Doncel¹, Jorge Gracia^{1,2}

¹ OEG, Universidad Politécnica de Madrid

² Universidad de Zaragoza

Abstract. This position paper describes the vision, objectives and methodology of the LYNX project. The aim of Lynx is to create services to better manage compliance, based on a legal knowledge graph which integrates and links heterogeneous compliance data sources including legislation, case law and standards.

Keywords. Compliance, Legal Knowledge Graph, regtech, regulatory compliance, semantic web

Introduction

The term *compliance* is widely used to refer “to the conformance to a set of laws, regulations, policies, or best practices” [1]. Every company performing almost any activity has some concern with compliance-related problems.

These problems are a bigger concern for small and medium-sized enterprises (SMEs), which cannot afford expensive consultancy services and they are a bigger hurdle for companies trying to sell abroad, as they usually lack the knowledge on the applicable conditions in the target country. According to the European Commission¹, only 7% of European SMEs sell across borders, but those who do, exhibit 7% job growth and 26% innovate in their offering, greater numbers than the 1% and 8% of SMEs that do not go outside their local markets.

The European Commission is aware of these legal and language barriers and is trying to build a single market with less entry barriers. The LYNX project is a European research project funded as an H2020 Innovation Action covering the topic ICT-14: *Big Data PPP: cross-sectorial and cross-lingual data integration and experimentation*. The project is expected to last three years, starting in December 2017. This position paper briefly presents this research project.

¹ http://europa.eu/rapid/press-release_IP-10-895_en.htm

1. The LYNX project

Having identified these compliance problem, and having assumed that technology can help lowering these legal and language barriers, the main objective of Lynx can be stated as “to create an ecosystem of cloud services to better manage compliance, based on a legal knowledge graph which integrates and links heterogeneous compliance data sources including legislation, case law, standards and other aspects”. Other specific objectives follow:

- to provide a platform that helps companies solving questions and cases related to compliance in different sectors and jurisdictions.
- to help European SMEs to reduce costs and effort in organising and monitoring legislation, regulations, and sectorial good practices.
- to provide citizens a better access to legal and regulatory information from multiple jurisdictions and to facilitate their in the legislative processes and in the writing process of standards.
- to draw a legal knowledge graph across different jurisdictions, comprising legislation, case law, doctrine, standards, norms, and other documents.
- to deliver the necessary domain-neutral common services (such as document annotation, interlinking, etc.) as building blocks to orchestrate aggregated business-oriented services.
- to cover some particular business case necessities, validating the project’s developed ecosystem.

The referred *legal knowledge graph* is a collection of structured data and unstructured documents which are densely interlinked. These documents shall be variate in their language, and terminologies and language resources will also be included in order to provide multilingual services. Semantic Web technologies favour the distributed publication of documents and data, and provide the technology to announce and traverse the links. Whereas the relationship between Semantic Web and Law has been a rich one [2], there had been no “Legal Linked Open Data Cloud” (cf. Linguistic Linked Open Data Cloud) nor the term “Legal Linked Data” has gained spread.

The Lynx Project will identify the major legal resources necessary to provide exploitable compliance services, storing their metadata in a uniform manner and running the processing routines to extract, structure, transform, annotate and link the resources whenever needed. The cloud of Lynx services will be built, including middleware microservices as well as end-user services, offered under different access modalities. These services will be configured in three pilots according to the industry needs in three different domains: *labour law*, *data protection* and *oil & gas and energy*. The services will consider compliance in a broad sense, including not only compliance with the law but also compliance with standards and even with good practices and internal policies.

In summary, this project may benefit European business by means of a cloud of services providing mass-customised regulatory information including legislation, regulations, and policies, which will open the path for a new breed of solutions. Faster answer will be facilitated to recurrent questions like: “*What is the equivalent to regulation X in country Y?*” or “*How is the technical question X handled in country Y?*”

How has it been understood by the local courts?'. As a side benefit, citizens will also see legislation published in different manners promoting legal knowledge and development of new applications. Finally, Lynx will also support the creation of a common legal ICT infrastructure that will contribute to unlock the potential of a multilingual, truly single digital market.

Acknowledgements

Lynx has received funding from the Horizon 2020 European Union (EU) Research and Innovation programme under Grant Agreement: 780602. We thank every partner in Lynx for their contribution to the project's inception and the Datos 4.0 project with ref. TIN2016-78011-C4-1-R.

References

- [1] Silveira, P., Rodríguez, C., Casati, C., Daniel, F., D'Andrea, V., Worledge, C. and Taheri, Z. On the design of compliance governance dashboards for effective compliance and audit management. In Service-Oriented Computing. ICSOC 2009, pp.208-217. Springer Berlin Heidelberg, 2010.
- [2] Benjamins, V. R., Casanovas, P., Breuker, J., & Gangemi, A. (2005). Law and the semantic web, an introduction. In Law and the Semantic Web (pp. 1-17). Springer, Berlin, Heidelberg.
- [3] Chiarcos, C., Hellmann, S., & Nordhoff, S. (2011). Towards a Linguistic Linked Open Data cloud: The Open Linguistics Working Group. TAL, 52(3), 245-275.