ISSN 2281-5147

INTRODUCTION TO “LEGALITY ATTENTIVE DATA SCIENTISTS FOR RVERYONE”

by Giovanni Comandè[*]

This special issue of Opinio Juris in Comparatione includes 14 contributions by the Early Stage Researchers (ESRs) of the "Legality Attentive Data Scientists" (LeADS) Project, funded by the European Commission (GA number 956562). These contributions were selected through a rigorous peer-review process conducted by a pool of international and interdisciplinary reviewers.

The review and selection process spanned nearly nine months, during which the researchers were guided in creating contributions that were both scientifically rigorous—based on three years of research within the project—and intentionally accessible to non-experts (at least to those not specialized in the researchers' respective core disciplines). The cultural goal aligns with the objectives of LeADS, which aim, among other goals, to create "a new interdisciplinary professional figure that we call Legality Attentive Data Scientist or LeADS. LeADS will be an expert in data science and law expected to work within and across the two disciplines, a leader in bridging scientific skills with the ethic-legal constraints of their operating environment.”

Credit must also be given to all the ESRs, even those who did not pass the lengthy and rigorous review process, for their intense cultural effort in crafting a language that is scientifically rigorous yet imbued with a popularizing spirit. This approach often required a certain heterogeneity in the notes, or even their removal and simplification.

[*] Full Professor of Comparative Private Law at Sant’Anna School of Advanced Studies and Coordinator of LeADS Project (www.legalityattentivedatascientis.eu). All contributions in this Special Issue were financially supported by the European Union's Horizon 2020 - Innovative Training Networks programme under the grant agreements ID: 956562.

ISSN 2281-5147

Each researcher interpreted the task according to their individual talents and the constraints imposed by their core discipline, resulting in styles that were often very different but all deserving of careful reading and re-reading to appreciate their diverse nuances.

We are confident that all the contributions will inspire the readers. Possibly those articles that explore the sectorial boundaries and overlaps of legislations, may help legislators and policymakers to propose regulatory innovation, support legislative reforms and devise appropriate policies bridging these players with in depth analysis and a less esoteric language. Eventually, theoretical and critical contents may help judges, independent authorities and legal experts to give a turn in their understanding of the emerging digital regulatory framework. Similarly, the technical and empirical results may increase the preparedness of developers, engineers, business owners and governmental organizations to more efficiently design and develop technologies and implement norms in real case scenarios.

Overall, we hope that citizens at large may be directly impacted by the overall results presented in these pages, for example through technological advancements for smoother user-centered privacy-friendly management of personal data or through access to fairer automated decision-making in key sectors such as justice and employment.

This issue is organized in sections around four core arguments.

The first one, devoted to "*Privacy, consumers and competition*", contains 3 rich contributions.

Onntje Hinrichs wrote on "*Why Your Data is not Your Property (and Why You Still End Up Paying With It)?*" exploring three interrelated topics that reveal tensions in the European approach towards the regulation of the data economy: (i) data as property (ii) data and fundamental rights, and (iii) data as currency. Qifan Yang put her twofold skills of statistician and jurist at work to understand the complex relationship between the GDPR and market competition in her article "*Your Data Rights: How does the GDPR Affect the Social Media Market?* "Last but not least Tommaso Crepax, with his very intriguing style drives everyone into the realm of data portability in the contribution entitled "*Bring Your Own Data. The struggle of re-using data in a world of heterogeneous systems*".

The second section, devoted to "*Privacy and security in practice*" is definitively dominated by a team of researchers with heightened technological skills. Cristian Lepore's "*Self-Sovereign Identity: The Revolution of Digital Identity*" drives us through the complex world Self-Sovereign Identity (SSI). Meanwhile, Louis Sahi, through his summary of interviews with experts and background analysis in "*Evaluation and Harmonization of Data Quality Criteria: Insights from Expert Interviews for Legal Application*" escorts the reader in understanding the technical and legal role of data quality criteria and the need for collaborative data processing (CDP) in decentralised environments.

Armend Duzha ("*Extracting Data Value through Data Governance*") and Christos Magkos ("*Personal Health Information Management Systems (PHIMS) for user empowerment: A Comprehensive overview*") continue this section. Mr. Duzha explores a new approach for data governance developed to extract data value respecting the ever-delicate balance between transparency and privacy, relating it to novel technologies such as Artificial Intelligence, Federated Learning and Blockchain, and illustrating how these can be integrated in a data governance program. Mr. Magkos devotes his attention to personal health information management systems (PHIMS) and on how integrating raw data could provide a method for the storage, management, and regulation of personal health data access. The key message is how PHIMS can empower users to take control of their own healthcare.

Privacy risks entailed in the advent of AI is at the core of the last contribution of this Section devoted by Soumia al Mestari to "*What AI is stealing! Data privacy risks in AI*". She discusses that this risk of AI's leaking personal data is not only hypothetical and suggests how to mitigate it.

The third section is devoted to "*Sharing (personal) data*", a title that would have suited a number of the contributions in the previous sections. Yet her ethe focus is more on the sharing *in practice*. Barbara da Rosa's "*Can business-to-government data sharing serve the public good?*" explores a number of regulations enacted by the European Union and their overlaps and analyzes if they indeed assist business-to-government data sharing. Xengie Doan ("*Collective consent, risks and benefits of DNA data sharing*") uses genetic data sharing as a use case to better understand what tools and methods can enhance a user-friendly, transparent, and legal-ethically aware collective consent. Still in the domain of health data is the contribution of Fatma Dogan ("*To Use or Not to Use? Re-using Health Data in AI Development*") focusing on the re-use of health data in the context of AI development, concentrating on regulatory frameworks governing this practice under the European Health Data Space. Her aim is to assess whether health data can be re-used for AI-driven healthcare advancements without undermining individuals' data protection rights.

In the last contribution of the section Maciej Zuziak ("*How to collaboratively use statistical models in a secure way*") empowers the curious reader with a set of links and pointers that would allow them to go deeper into a well of data governance and large AI infrastructure but only after having introduced the reader to the nuanced world of decentralised learning systems and statistical learning explaining the basic technocratic lingo in an engaging way.

The last Section ("*Preparing for AI* ") is opened by Mitisha Gaur's "*Policing the AI Judge: a Balancing Act*" where she analyzes AI backed automated decision-making systems used by public authorities. She advocates for a strict governance framework based on risk management and algorithmic accountability practices focused on safeguarding fundamental rights and upholding the rule of law by adhering to the principles of natural justice.

Robert Poe's challenging "*The perils of Value Alignment*" is a program already by the title. The article vigorously argues that global AI governance risks institutionalizing violations of fundamental rights. It argues that the current ethical foundation of AI governance can lead to conflicts with the rule of law. It calls for a re-evaluation of AI governance strategies, urging a realistic approach that respects citizens, legal precedent, and the nuanced realities of social engineering.

ISSN 2281-5147