iUS PUBLICUM NETWORK REVIEW
www.ius-publicum.com

DIGITALIZATION OF JUSTICE: APPLICATIONS AND OPEN ISSUES

Mattia FALDUTI[1]

INDEX

[1] Ph.D, Postdoctoral Assistant Researcher, Faculty of Computer Science, Free University of Bozen-Bolzano, Italy.

1. INTRODUCING E-JUSTICE, ARTIFICIAL INTELLIGENCE AND LAW

The digitalization of justice and legislation is coming regularly under the spotlight[2] in various field of research, involving complex interdisciplinary issues, related both to computer science and (public) law[3]. Many are the words and the concepts used to express this topic that today is recognized as a proper field of research. A common word to identify the digitalization of justice is *e-justice*[4], intended as a trend for judicial agencies to create their websites with different levels of technological sophistication and functionality, often to open their processes and interact with multiple stakeholders[5].

[2] G. SARTOR, *Intelligenza Artificiale e Diritto. Un'introduzione.*, in *Informatica e Ordinamento Giuridico*, n. 12, Milano, Giuffrè, 1996.

[3] R. CAVALLO PERIN – D. U. GALETTA, *Il diritto dell'amministrazione pubblica digitale*, Torino, Giappichelli, 2020.

[4] https://e-justice.europa.eu/home?action=home&plang=en, accessed in June 2022.

[5] R. SANDOVAL-ALMAZAN, J. RAMON GIL-GARCIA, *Understanding E-Justice and Open Justice Through the Assessment of Judicial Websites: Toward a Conceptual Framework*, in *Social Science Computer Review* 38, n. 3, 334–53.

Differently, *digital justice*[6] is a concept to name the new technological tools introduced to resolve and prevent litigation in cyberspace. In particular, digital justice is intended as the set of technical and legal solutions where those who would never go before a court for assistance can find easy and direct help digitally, via a smartphone or pc. Recently, the words *online courts* and *online judging* have been used to express a new digital trend in the justice system. In particular, firstly, considering that nowadays more people have access to the internet than access to justice, *online courts* are words used to intend the so-called *extended courts*, where access to justice is guaranteed by tools that permit users to understand their rights and duties[7]. Secondly, *online judging* is a concept used to intend all the cases where judges and parties do not gather together arguments in a courtroom, but instead, evidence and arguments are presented to judges through an online platform[8].

From different perspectives, all these topics address the interaction between law and technology, but there is a main topic that is particularly popular at the moment, and this is artificial intelligence (AI)[9]. We are in the 3rd wave of the AI boom[10] and this wave is followed

[6] E. KATSH, O. RABINOVICH-EINY, *Digital Justice: Technology and the Internet of Disputes*, *Digital Justice*, Oxford University Press, 2021.

[7] R. SUSSKIND, *Online Courts and the Future of Justice*, Oxford University Press, 2019.

[8] R. SUSSKIND, *ibidem.*

[9] https://www.liquid-legal-institute.com/workinggroups/legal-text-analytics/

[10] D.H. CHAU, *Data Science Buzzwords*, in *CSE6242: Data & Visual Analytics*, 2019. https://poloclub.github.io/cse6242-2019fall-campus/slides/CSE6242-015-BuzzWords.pdf, accessed in June 2022.

by the ethical topics related to AI, namely, transparency [11], non-discrimination [12], accountability [13] and fairness. These topics are discussed from several perspectives and disciplines, such as sociological causes, legal argumentations, economic models, statistical techniques, computational issues[14], and also justice. For example, access to justice and the future challenges of artificial intelligence and justice[15]. Today, the field of research named *Artificial Intelligence and Law (AI and Law)* appears mature and composed of several sub-fields, but before the access to large legal *corpora* was permitted, many approaches in the AI and Law during the '90 were devoted to realise expert system applications[16] or to elaborate knowledge representation techniques [17], both aiming the making law processable by a

[11] In terms of transparency, the risk is to develop processes and algorithms that are unclear, incomprehensible and unrepeatable. M. TURILLI, L. FLORIDI, *The ethics of information transparency*, in *Ethics and Information Technology* 11, n. 2, 2009, 105–12.

[12] S. BAROCAS, A.D. SELBST, *Big data's disparate impact*, in *California Law Review* 104, n. 3, 2016, 671–732.

[13] R. RODRIGUES, *Legal and Human Rights Issues of AI: Gaps, Challenges and Vulnerabilities*, in *Journal of Responsible Technology*, n. 4, 2016.

[14] A. ROMEI, S. RUGGIERI, *A multidisciplinary survey on discrimination analysis*, in *The Knowledge Engineering Review* 29, n. 5, 2014, 582–638.

[15] Predictive Justice and Artificial Intelligence, European Commission for the Efficiency of Justice (CEPEJ), https://www.coe.int/en/web/cepej/justice-of-the-future-predictive-justice-and-artificial-intelligence. Accessed in June 2022.

[16] G. SARTOR, *ibidem.*

[17] D. SCHMIDT, C. TROJAHN DOS SANTOS, R. VIEIRA, *Analysing top-level and domain ontology alignments from matching systems*, in *Proc. of the 11th int. Workshop on ontology matching co-located with the 15th int. Semantic web conf. (ISWC2016)*, Kobe, Japan, 2016.

machine. These topics are still discussed today, confirming that the field of study, even if mature, is far to be outlived. Indeed, a recent literature review[18] identified two main streams in this field of research, namely i) computational models of legal reasoning[19] and ii) legal text analytics[20].

1.1. Intelligent applications and the legal domain: an overview

In the first set, we can find approaches, devoted to designing computational models of legal reasoning, concerning both i) *modeling statutory reasoning*[21], where legal rules are expressed logically and computers can reason deductively and ii) *modeling case-based legal reasoning*[22], where computational models are designed for interpreting terms and concepts

[18] K. D. ASHLEY, *Artificial Intelligence and Legal Analytics. New Tools for Law Practice in the Digital Age*, Cambridge University Press, 2017.

[19] H. PRAKKEN, G. SARTOR, *The Role of Logic in Computational Models of Legal Argument: A Critical Survey*, in A.C. KAKAS, F. SADRI, (eds) *Computational Logic: Logic Programming and Beyond. Lecture Notes in Computer Science*, 2002, vol 2408. Springer, Berlin, Heidelberg.

[20] J.G., CONRAD, L.K. BRANTING, *Introduction to the special issue on legal text analytics*, in *Artificial Intelligence and Law*, 26**,** 99–102 (2018).

[21] The first approaches recognized the problem of coding the ambiguity of the law. E. A. LAYMAN, C. S. SAXON, *Some problems in designing expert systems to aid legal reasoning*, in *Proceedings of the 1st international conference on Artificial intelligence and law* (*ICAIL1987*), New York, USA, 1987, 94–103.

[22] For example, HYPO (a computer program that models reasoning with cases and hypotheticals in the legal domain, 1987) and CATO (a computer program which employs artificial intelligence techniques to teach first-year law students how to make basic legal arguments with cases, 2000), operating in the domain of US Trade Secrets Law, K. D. ASHLEY, *Reasoning with cases and hypotheticals in HYPO*, *International Journal of Man-machine Studies.*,

through analogical reasoning. Since early 2000, also the speed and ability to query data streams have encouraged the development of knowledge-based approaches. Legal Knowledge-Based System (LKBS) and Judicial Decision-Support Systems (JDSS) devoted to judgment support and justice prediction have been intended as efficient solutions for the jurisdiction.[23]

In the second set, we can find approaches devoted to developing techniques for legal text analytics concerning *extracting information from statutory and regulatory texts*[24], where the aim is to automatically extract information about rules' requirements from legislation texts, electronically stored, and *extracting argument-related information from legal case text*[25], where the aim is to use machine learning (ML), Natural Language Processing (NLP) techniques[26] or manually constructed rules, for extracting information from case law, or other

1991, 753–96. V. ALEVEN, *Using background knowledge in case-based legal reasoning: A computational model and an intelligent learning environmen*t, in *Artificial Intelligence*, 2003, 183–237.

[23] P.L.M. LUCATUORTO, *Teorie e modelli del diritto per il ragionamento giuridico automatico*, LED, Milano, 2009.

[24] J. SAVELKA, M. GRABMAIR, K. D. ASHLEY, *Mining information from statutory texts in multi-jurisdictional settings*, in *Proc. of the 27th int. Conf. on legal knowledge and information systems (JURIX2014)*, Krakow, Poland, 2014, 133–42.

[25] M. GRABMAIR, *Predicting trade secret case outcomes using argument schemes and learned quantitative value effect tradeoffs*, in *Proc. of the 16th int. Conference on Artificial Intelligence and Law (ICAIL2017)*, London, UK, 2017, 89–98.

[26] Machine learning (ML) is a subfield of artificial intelligence, which is broadly defined as the capability of a machine to imitate intelligent human behavior, in https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained, accessed June 2022. Natural Language Processing (NLP) strives to build machines that understand and respond to text or voice data, in https://www.ibm.com/cloud/learn/natural-language-processing, accessed June 2022.

sorts of legal case decisions. Case studies and practical approaches have been tested in and outside Europe[27]. For instance, possible litigation outcomes are assessed by Artificial Intelligence (AI) tools in China[28] and NLP techniques have been tested on the ECHR proceedings to automatically predict (future) judicial decisions[29] or violations of the Convention, with promising results.[30] Such data-driven systems have shorter development timing and permit the ability to discover new information[31] but have highlighted the need for data completeness and transparency of the algorithms[32]. However, even if many approaches appear promising, some issues are still unresolved. The first concerns data completeness, intended as access to case law, and in general, access to any (legal) decisions. In fact, at the

[27] In Europe, for example T. NOVOTNA et al. *Topic Modelling of the Czech Supreme Court Decisions*, in *Proceedings of Automated Semantic Analysis of Information in Legal Text* 2020 and GLASER et al. *Classification of German Court Rulings: Detecting the Area of Law*, in *Proceedings of Automated Semantic Analysis of Information in Legal Text*, 2020. Outside Europe, for example, H. ZHONG, Z. GUO, C. TU, C. XIAO, Z. LIU, AND M. SUN. *Legal Judgment Prediction via Topological Learning*, in *Proc. of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3540–3549, Brussels, Belgium. Association for Computational Linguistics or BRANTING et al., *Predictive Features of Persuasive Legal Texts*, in *Proceedings of Automated Semantic Analysis of Information in Legal Text*, ASAIL, 2020.

[28] R. SUSSKIND, *Online Courts and the Future of Justice*, Oxford University Press, 2019.

[29] M. MEDVEDEVA, M. VOLS, M. WIELING, *Using Machine Learning to Predict Decisions of the European Court of Human Rights*, in *Artificial Intelligence and Law* 28, n. 2, 2019.

[30] N. ALETRAS et al., *Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective*, in *PeerJ Computer Science*, n. 2, 2016, 93.

[31] A. ALZGHOUL et al., *Comparing a Knowledge-Based and a Data-Driven Method in Querying Data Streams for System Fault Detection: A Hydraulic Drive System Application*, in *Computers in Industry* 65, n. 8, 2014, 1126–35.

[32] E. VINCENTI, *Il "problema" del giudice-robot*, in A. CARLEO (ed. by), *Decisione Robotica*, Mulino 2019.

moment, court decisions are only partially accessible, meaning that only few databases permit the download of bulk data, stored in processable forms and enriched with metadata[33]. A similar approach is followed by other decisions written by other institutions, such as independent administrative authorities. Datasets with the decisions produced by these entities are not publicly open nor digitally stored or completely accessible. However, even if desirable, full access to all decisions is not sufficient for legal analytics approaches. Thus, ML and NLP techniques require a sufficient level of quality of data, in order to avoid the first problem of data analysis, the so-called *garbage in, garbage out.* However, permitting full access to high-quality legal data is not yet enough, as demonstrated in real case scenario[34]. The language expressed in legal texts is a particular sub-language, with its unique rules[35]. When accessible, court decisions are usually not annotated or report few useful metadata. Moreover, contrarily to other text analysis domains, such as social media text analysis[36] and

[33] M. FALDUTI, *Court Decisions and Data Analysis: a Survey Among 22 Member States of the European Union on Access to Case Law and Legal Prediction*, in *Journal of Law, Cognitive Science and Artificial Intelligence*, n. 13, 2020.

[34] M.T. SAGRI, T. AGNOLONI, L. BACCI, *Legal Keyword Extraction and Decision Categorization: a Case Study on Italian Civil Case Law*, in *Proc. of the 5th Workshop on Semantic Processing of Legal Texts (SPLeT2014)*, Reykjavik, Iceland, 2014, 1–7.

[35] M. CECI, A. GANGEMI, *An OWL Ontology Library Representing Judicial Interpretations*, in *Semantic Web Journal* 7, n. 3, 2016, 229–53.

[36] I. RIZWANA et al. *A survey on text mining in social networks*, in *The Knowledge Engineering Review*, 2015, p. 157-170.

biomedical text mining[37], researchers in the legal domain can rely only on a few international resources.

1.2. Intelligent applications and the legal domain in practice

From a European perspective, many projects addressed the justice system and the legal matters following an ontological approach[38]. A statute-specific legal ontology of the Polish Commercial Companies Code (PCCC)[39] has been developed for defining concepts, properties and axioms of the commercial law domain. The authors of this work point out that a problem with all legal knowledge is the fuzziness of reality compared with the restricted language of the law.[40] A second example, again for the application of civil law, is an ontological model aimed at formalising the Croatian Family Legislation with the design of a

[37] A. M. COHEN, W. R. HERSH, *A survey of current work in biomedical text mining*, in *Briefings in bioinformatics*, 2005, p. 57-71.

[38] Studies about legal ontologies are discussed in C. GRIFFO et al., *A Systematic Mapping of the Literature on Legal Core Ontologies*, in *Ontobras*. 2015 and in M. FALDUTI, *Law and Data Science: Knowledge Modeling and Extraction from Court Decision*s, 2021, https://air.unimi.it/handle/2434/799875.

[39] P. STOLARSKI, T. TOMASZEWSKI, *Modeling and Using Polish Legal Knowledge - Commercial Companies Code Ontology*, in W. ABRAMOWICZ, AND D. FENSEL, (eds) *Business Information Systems*, 2008, vol 7. Springer, Berlin, Heidelberg.

[40] P. STOLARSKI, T. TOMASZEWSKI, *Modeling and Using Polish Legal Knowledge - Commercial Companies Code Ontology*, in *Proc. of the 11th int. Conf. on Business Information Systems (BIS2008)*, Innsbruck, Austria, 2008, 83–94.

legal expert and a family law judge[41]. Furthermore, the Consumer Protection law has been also the object of various approaches. An ontology developed within the DALOS (Drafting Legislation with Ontology-based Support) [42] project has been realised for supporting legislative drafting by providing legal drafters and decision-makers control over the legal language at the European level[43]. Furthermore, the DALOS KOS has been presented as a middle-out legal ontology and another type of lexical relationship, the so–called *fuzzynym*, has been developed in *two layers*, namely, the Ontological Layer, containing the conceptual modelling at a language independent level, and the Lexical Layer, containing multi-lingual terminology conveying the concepts represented at the Ontological layer. Classes and properties have been implemented on the basis of the terminological knowledge extracted from the chosen European Directives on consumer protection law[44]. Moreover, in civil law, the Dutch tort law has been formalised within an ontology capable to capture the knowledge

[41] S. LOVRENCIC, I. J. TOMAC, *Managing Understatements in Legislation Acts When Developing Legal Ontologies*, in *Proc. of the 10th int. Conf. on Intelligent Engineering Systems (INES2006)*, London, UK, 2006, 69–73.

[42] The Italian Research Council clarifies that, in a multilingual environment, and in particular in EU regulations, only the awareness of the subtleties of legal lexicon, in the different languages, can enable drafters to maintain coherence among the different linguistic version of the same text. To this end, DALOS – Drafting Legislation with Ontology-based Support Project - intends to provide law-makers and European citizens with linguistic and knowledge management tools to be used respectively in the phase of legislative drafting and in the retrieval procedures. http://www.ittig.cnr.it/progetti/dalos/

[43] T. AGNOLONI et al., *Building an Ontological Support for Multilingual Legislative Drafting*, in *Proc. of the 20th Annual Conference on Legal Knowledge and Information Systems (JURIX 2007)*, Leiden, The Netherlands, 2007, 9–18.

[44] E. FRANCESCONI et al., *Integrating a Bottom–Up and Top–Down Methodology for Building Semantic Resources for the Multilingual Legal Domain*, in *Semantic Processing of Legal Texts: Where the language of law meets the law of language*, Springer, 2010, 95–121.

of entities subject to the law (a legal person, a natural person) and objects in tort law (motor vehicles, animals, product).[45]

On the topic of data protection, many works have been proposed. For instance, the Spanish data protection law has been formalised with an ontology called LegLOPD (Legal Ontology Domain)[46], composed of five top concepts directly extracted from a model named LRI-Core ontology[47], a core ontology that covers the main concepts that are common to all legal domains[48]. This project has been followed by another one, again on Spanish law, where the legal knowledge has been formalised as a modular ontology based on the knowledge acquired and organised by legal experts, with the aim of modeling data protection concepts for a reasoning system[49]. More recently, the General Data Protection Regulation (GDPR) constituents and relationships among them have been described in a bottom-up ontology[50].

[45] R. LAARSCHOT et al., *The Legal Concepts and The Layman's Terms Bridging the Gap Through Ontology-Based Reasoning about Liability*, in *Proc. of the 18th int. Conf. on Legal Knowledge and Information Systems (JURIX2005)*, Brussels, Belgium, 2005, 115–25.

[46] H. A. MITRE et al., *A Legal Ontology to Support Privacy Preservation in Location-Based Services*, in *Proc. of int. Workshop on Web Semantics (SWWS2006)*, Montpellier, France, 2006, 1755–64.

[47] http://www.leibnizcenter.org/previous-projects/lricore

[48] J. BREUKER et al., *Law and the Semantic Web: Legal Ontologies, Methodologies, Legal Information Retrieval, and Applications*, in *Law and the Semantic Web*, Springer, 2005, 36–64.

[49] N. CASELLAS et al., *Ontological Semantics for Data Privacy Compliance: The NEURONA Project*, in AAAI Press Technical Reports Series, 2010.

[50] C. BARTOLINI et al., *Using Ontologies to Model Data Protection Requirements in Workflows*, in *New frontiers in Artificial Intelligence*, ed. by M. OTAKE et al., Springer 2017, 233–48.

Then, using Simple Knowledge Organisation System SKOS[51], concepts and obligations of GDPR have been formalised with the set of attributes and terms provided by European Legislation Identifier (ELI) metadata[52]. Finally, the PrOnto ontology presented the GDPR main concepts (data types, documents, processing purposes, legal bases, processing operations) with the aim of supporting legal reasoning and compliance checking[53].

1.3. Risks and side effects

The described projects are mostly knowledge-based. Usually, the advantages of such approaches are the following: i) full control over the system, based on the ontology, usually built (and updated) with the help of a legal domain expert, ii) explainability of the system and, iii) portability, because such system do not require necessarily big datasets. This is particularly relevant to the legal domain, where the available datasets are language-based, difficult to be shared, and are mostly concerning law instead of case law. In fact, in the era of big data, research activity on data science focuses on the collection, processing, and interpretation of large datasets to produce knowledge for decision-making processes in

[51] https://www.w3.org/TR/skos-primer/ Accessed June 2022.

[52] H. J. PANDIT et al., *GDPR as a Linked Data Resource*, in *The European Semantic Web Conference*, Springer 2018, 481–95.

[53] M. PALMIRANI et al., *Pronto: Privacy Ontology for Legal Reasoning*, in A. KŐ - E. FRANCESCONI (ed. by), *Electronic Government and the Information Systems Perspective*, Springer 2018, 139–52.

different application domains and contexts, such as finance[54], healthcare[55] and social science[56].

The legal domain is one of these different domains and contexts where data science approaches can be applied.[57]

However, the development of data-driven approaches may imply side effects and risks, in some cases. For instance, there are already evidences of human rights violations in EU and in the US, with regard to discrimination within the context of AI, such as banks using postcodes to predict problems for repaying the loan;[58] or premium car insurances asking for men to pay more than women for the same type of insurance as statistics showed women are

[54] C. LONGBING et al., *Data Science and AI in FinTech: an overview*, in *International Journal of Data Science and Analytics*, 2021, Springer, p. 81-99.

[55] J. ARCHENAA et al., *A Survey of Big Data Analytics in Healthcare and Government*, in *Procedia Computer Science*, 2015, p. 408-413.

[56] N. CARLO LAURO et al., *Data Science and Social Research, Epistemology, Methods, Technology and Applications*, Springer Cham, 2017.

[57] M. FALDUTI, *Law and Data Science: Knowledge Modeling and Extraction from Court Decisions*, Ph.D. thesis, Università degli studi di Milano, Dipartimento di Informatica Giovanni Degli Antoni, 2021.

[58] F. ZUIDERVEEN BORGESIUS, *Discrimination, Artificial Intelligence, and Algorithmic Decision-Making*, Technical Report, Council of Europe, 2018. https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73, accessed in June 2022.

more careful drivers,[59] or risk assessments predicting recidivism in US courts,[60] or an algorithm used by the Dutch government to predict who is likely to wrongly child benefits,[61] or an AI-system used for recruitment showing bias against women,[62] or a racial bias in the facial recognition process.[63]

2. DATA-DRIVEN APPROACHES FOR THE JUSTICE SYSTEM

The development of algorithms able to judge a case would implies choosing, through the available legal interpretation theories, the one (and the best) to use for the case. Such a decision sounds more political than technical[64]. A way to overcome this obstacle is to

[59] Case C-236/09, *Association des Consommateurs Test-Achats ABSL, Yann van Vugt, Charles Basselier v. Conseil des ministres*, 2011 E.CR. I-00773: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62009CA0236&qid=1612293472333#ntr1-C_2011130EN.01000401-E0001, accessed on 2 February 2021.

[60] J. ANGWIN et al., *Machine Bias. There's Software Used Across The Country to Predict Future Criminals. And It's Biased Against Blacks*, ProPublica, 23 May 2016: Machine Bias — ProPublica, accessed in May 2021.

[61] For more details: https://www.politico.eu/article/europe-artificial-intelligence-blindspot-race-algorithmic-harm/amp/, accessed on May 2021.

[62] J. DASTIN, *Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women*, Reuters, October 2018.

[63] C. GARVIE, J. FRANKLE, *Facial-Recognition Software Might Have a Racial Bias Problem*, The Atlantic, the 7th April 2016.

[64] M. LUCIANI, *La decisione giudiziaria robotica*, in A. CARLEO (ed. by), *Decisione Robotica*, Mulino 2019.

include, inside the justice prediction system, the knowledge inferred by caselaw datasets.[65] However, once again the limit of such inclusion appears more political than technical. For instance, if the caselaw datasets, used to predict new outcomes, should be populated with past decisions, the issue here is how old these decisions should be and who has the power to determine the timeframe. Furthermore, the risk of deciding a case and basing the argumentation on statistical analysis might standardise the judgments, compromising the peculiarities of both the cases and the involved individuals.[66] Additionally, intelligent systems based on standardised judgments appear not capable to deal with open texture concepts and the vagueness of the law.[67]

Differently, with respect to criminal law, scholars analysed the legal implications of softwares for law enforcement authorities[68] aim at predictive policing with data analysis and profiling techniques.[69] Examples of such systems are *Key-crime* and *XLAW* tested with the

[65] C. CASTELLI, D. PIANA, *Giustizia Predittiva, La qualità della giustizia in due tempi*, in Questione Giustizia, 2018.

[66] L. DE RENZI, *Primi passi nel mondo della giustizia "high tech": la decisione in un corpo a copro virtuale tra tecnologia e umanità*, in A. CARLEO (ed. by), *Decisione Robotica*, Mulino, 2019.

[67] F. PATRONI GRIFFI, *La decisione robotica e il giudice amministrativo*, in A. CARLEO (ed. by), *Decisione Robotica*, Mulino, 2019.

[68] G. CONTISSA et al., *Quando a decidere in materia penale sono (anche) algoritmi e AI: alla ricerca di un rimedio effettivo*, in Riv. Trim. diritto di internet, 2019, 4. 610.

[69] F. BASILE, *Intelligence artificiale e diritto penale*: *quattro possibili percorsi di indagine*, in Diritto Penale e Uomo, 2019.

police department of Milan and Naples, respectively.[70] On the other hand, software for judgment prediction or support have been addressed only from a theoretical point of view. Someone pointed out that the error rate would decrease[71], whereas others suggest that intelligent systems would resolve the sentencing disparity[72] or maybe reach the perfect justice.[73] A very famous example of an intelligent system tested in the criminal court is COMPAS (Correction Offender Management Profiling for Alternative Sanctions) used for calculating the defendant's recidivism risk index.[74] This system is usually recalled in the literature as an example of a system with a discrimination and racial/gender bias[75].

[70] C. PARODI, V. SELLARODI, *Sistema penale e intelligenza artificiale*, in Diritto Penale Contemporaneo - Rivista Trimestrale, 2019.

[71] C. BONA, *Sentenze imperfette, gli errori cognitivi nei giudizi civili*, Mulino, 2010.

[72] V. MANES, *L'oracolo algoritmico e la giustiza penale: al bivio tra tecnologia e tecnocrazia*, in U. RUFFOLO, *Intelligenza artificiale, il diritto, i diritti, l'etica.* Giuffrè 2020.

[73] M. LUCIANI, *La decisione giudiziaria robotica*, in Rivista AIC, 2018, 872.

[74] T. BRENNAN, et al., *Correctional offender management profiles for alternative sanctions (COMPAS)*, in *Handbook of recidivism risk/need assessment tools*, 2018, 49-76.

[75] A. M. PIERSON. *Validation of the Correctional Offender Management and Profiling Alternative Sanctions (COMPAS)*. Fordham University, 2018.

The depicted scenario has not discouraged the development of justice predictive systems. On the contrary, the discussion around the issues of discrimination[76], transparency[77], privacy[78] and explainability[79] has positively impacted the whole AI and Law community, both in the US[80] as well as in Europe. For instance, in Italy, today the discussion reached a sufficient level of maturity, as demonstrated by the dense literature[81] on the topic of AI and Law and all the projects that are under development in academia, in the judiciary and in the government. This multilateral approach confirms the hybrid nature of the justice system, partly administrative and partly judiciary. In this sense, the role of the public administration is a key factor in pursuing and achieving EU goals.[82]

[76] S. WACHTER et al., *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*, in *Computer Law & Security Review*, n. 41, 2021.

[77] P. KSIĘŻAK, S. WOJTCZAK. *Causation in Civil Law and the Problems of Transparency in AI*, in *European Review of Private Law* n. 29, 2021.

[78] M. FALDUTI, *Court Decisions and Data Analysis: a Survey Among 22 Member States of the European Union on Access to Case Law and Legal Prediction*, in *Journal of Law, Cognitive Science and Artificial Intelligence*, n. 13, 2020.

[79] A. BIBAL et al. *Legal Requirements on Explainability in Machine Learning*, in *Artificial Intelligence and Law* n. 29, 2021.

[80] https://ainowinstitute.org, Accessed June 2022.

[81] A. CARLEO, *Decisione robotica*, il Mulino, 2019; U. RUFFOLO, *Intelligenza artificiale, il diritto, i diritti, l'etica. Tech e-Law*, Giuffrè, 2020.

[82] R. CAVALLO PERIN, G. M. RACCA, *The Plurality and Diversity of Integration Models: the Italian Unification of 1865 and the European Union Ongoing Integration Process*, in *The Changing Administrative Law Of An EU Member State*, Springer, Torino, Giappichelli, 2021.

3. DIGITALIZATION OF JUSTICE: NEXT STEPS

From a European perspective, the first tangible result of the political commitment to making domestic and European e-justice more accessible was the adoption of the first multi-annual e-Justice Action Plan 2009-2013.[83] A few years passed after this Action Plan, and today the topics concerning digital transformation, deployment of innovative technologies and big data have been crucial on the EU's agenda. The e-justice strategy and the action plan for the 2019-2023[84] period are active. In this context, the European Commission published the final report of the study on the use of innovative technologies in the justice field identifying the priority areas of the use of AI in the justice field.[85]

The aims of this work are several. Firstly, it presents the relevant existing EU legal and policy framework. Secondly, it summarizes all the aspects that need to be taken into account in terms of innovative technologies in the justice field, presented in a coherent and narrative way. The focus is on the business problems tackled during the implementation of the projects carried out by public authorities and the judiciary in the Member States, and by legal professional organisations. From this analysis, the following eight categories of problems have been identified.

[83] Multi-Annual European E-Justice Action Plan 2009-2013, https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:075:0001:0012:en:PDF. Accessed in June 2022.

[84] The European e-Justice Strategy https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019XG0313(02)&rid=6. Accessed in June 2022.

[85] European Commission, Study on the use of innovative technologies in the justice field - Final Report, April 2020, ISBN 978-92-76-21347-5, https://op.europa.eu/en/publication-detail/-/publication/4fb8e194-f634-11ea-991b-01aa75ed71a1/language-en. Accessed in June 2022.

(1) Processing high volume of data, such as high volumes of structured and unstructured data and documents manually or with simple digital tools, in order to find relevant information for the case, deduct patterns, search for specific words or cases.

(2) Processing high volume of video, audio and images, such as high volume of video files, audio files and/or images in order to make an analysis of the content, for tasks such as identification of persons/victims, or monitoring of behaviour, detecting illegal activities or transcription to text.

(3) Linking information across different sources, in other terms extracting and analysing information from multiple sources usually because they are not centralised or connected.

(4) Access to justice/public services, intended as making judicial information or public services available to the citizens/the general public in a user-friendly and easily accessible way. Access to case law, case information, legislation, treatment of citizens' questions are included topics.

(5) Data protection compliance, it means making documents (usually court judgments and decisions) compliant with the personal data protection legislation with the aim of making those documents publicly available.

(6) Preparing high volume of data, treating data manually, or with simple digital tools, in order to obtain a final output. This involves tasks such as translation of documents, typing of protocols in court hearings or interviews, preparation of contracts, judicial decisions.

(7) Administrative/facilities management. The issue of managing the court administration processes performed by the judicial personnel (clerks, judges, lawyers, etc.), with tasks such as planning the agendas, court hearings, booking and allocation of courtrooms and infrastructure, organising interviews and doing the facility management.

(8) Lack of authenticity and traceability. An insufficient level of traceability imposes actions to ensure a sufficient level of authenticity, trust and integrity of data and documents during their process flows.

From this study emerged that in the EU, Italy reports the highest number of projects (35) followed by Sweden (13). However, the majority of the other European countries, reported only three or less projects. Moreover, the final report confirms that only 46% of the EU member states defined a strategy/policy addressing the use of AI in the justice field. In Italy, institutions are conducting few projects devoted to justice prediction, where both Faculties of Law and Courts are involved.

4. DIGITALIZATION OF JUSTICE IN ITALY

As pointed out, in the last three years, in Italy there are several projects under development, running not only in the judiciary, but also in conjunction with universities or with the government institutions, such as the bar associations[86].

4.1. Projects in the Judiciary

[86] https://www.altalex.com/documents/news/2021/11/04/giustizia-predittiva-progetti-italiani-sentenza-cds. Accessed in June 2022.

The Court of Florence is developing a project called "*the city of simple justice*" aiming to simplify and reduce the administrative burdens in the context of the resolution of civil disputes[87]. One of the key objectives of the project is the creation of models or algorithms capable to incorporate the preventive assessments of mediators, as well as the ability to assess disputes in order to anticipate the probability of successful mediation for the benefit of the parties and/or the judge. From a technical point of view, this project will combine not only expert systems and rule-based systems (manually defined rules in a knowledge-based) but also machine learning, NLP and speech recognition techniques.

A project devoted to criminal justice is conducted by the prosecutor's office at the Court of Cosenza, where since December 2019 a project focused on conceptual modelling of justice data is performed. The aim is to develop a taxonomy of the previous procedures and to design an IT system to support these procedures based on raw data. Moreover, a definition of similarity metrics among the same procedures, together with the design of a dashboard to monitor the interpretative behaviour of real-time changes are planned. The Court of Cosenza aims to automate the legal workflow using data mining and ML techniques to identify similarities among the recalled procedures. The case study will be the emergency of gender violence[88].

Court of Appeal of Reggio Calabria together with the Mediterranean University of Reggio Calabria and other institutions presented *JustitIA*[89] a joint research project with a

[87] https://www.cittametropolitana.fi.it/wp-content/uploads/report.pdf. Accessed in June 2022.

[88] European Commission, Study on the use of innovative technologies in the justice field - Final Report, April 2020, ISBN 978-92-76-21347-5, https://op.europa.eu/en/publication-detail/-/publication/4fb8e194-f634-11ea-991b-01aa75ed71a1/language-en. Accessed in June 2022, p. 57.

[89] http://www.iustit-ia.it. Accessed in June 2022.

twofold aim. Firstly, using NLP techniques they aim at reducing the duration and the costs of the trials. Secondly, using data science and text mining approaches will be applied to open and big data, presumably large corpora of courts decisions. Unfortunately, further technical details are not presented yet.

Furthermore, the Court of Appeal of Milan is conducting a project called "*Milan Antitrust Justice*" on competition law. The project is focused not only on collecting case law in the competition law field and automating case law reviews in the field of competition law, but also on digitalising civil and criminal proceedings as well as administrative requests to fund justice expenses. The final goal of this project is to reduce the length of court proceedings and also to ensure that a larger number of cases can be handled, increasing both efficiency and productivity. A positive effect is the opportunity to acquire insights from the processed data and monitor the results. Again, it is indicated that to achieve these results, both expert and rule-based systems, combined with NLP techniques, will be applied to the data[90].

The Court of Appeal in Bologna is conducting a project devoted to tort law and to family law. The aim is to identify the criteria for quantifying personal injury in tort cases, as well as quantifying maintenance allowances in divorce cases. Moreover, the project aims at automating and facilitating the processes related to the quantification of harm and damages. The final goal is to reduce the length of court proceedings and to ensure that a larger number

[90] European Commission, *Study on the use of innovative technologies in the justice field - Final Report*, April 2020, ISBN 978-92-76-21347-5, https://op.europa.eu/en/publication-detail/-/publication/4fb8e194-f634-11ea-991b-01aa75ed71a1/language-en. Accessed in June 2022, p. 123.

of cases could be handled, increasing consistency, in terms of repeatability and reproducibility of court decisions[91].

Moreover, the Court of Appeal of Salerno is developing a management system of courtrooms using AI approaches, such as expert systems and rule-based systems with the aim to improve both the efficiency of the management of courtrooms and the organisation of court hearings[92].

Since (year) the Court of Appeal in Brescia, the faculty of law and the faculty of statistics of the University of Brescia have been conducting a project with the aim of sharing a public database with court argumentations, case study and all the details that can be of interest not only for legal practitioners about labour law[93]. The final goal is to ensure consistency (predictability) in the decisions taken. Moreover, the data collected would encourage the sharing of the court decisions between the courts of first and second instance.

The University of Pisa and the Court of Pisa, together with the Court of Genova, are developing practices for anonymizing the decisions, with additional techniques and solutions across disciplines. The aim is threefold. Firstly, the system aims to create innovative tools

[91] European Commission, *Study on the use of innovative technologies in the justice field - Final Report*, April 2020, ISBN 978-92-76-21347-5, https://op.europa.eu/en/publication-detail/-/publication/4fb8e194-f634-11ea-991b-01aa75ed71a1/language-en. Accessed in June 2022, p. 206.

[92] European Commission, *Study on the use of innovative technologies in the justice field - Final Report*, April 2020, ISBN 978-92-76-21347-5, https://op.europa.eu/en/publication-detail/-/publication/4fb8e194-f634-11ea-991b-01aa75ed71a1/language-en. Accessed in June 2022, p. 207.

[93] https://giustiziapredittiva.unibs.it. Accessed in June 2022.

for querying legal materials through their automatic annotation. Secondly, the construction of predictive tools based on data science and artificial intelligence[94]. The third goal is to ensure the necessary knowledge of the algorithm. In other words, the organisation intends to develop not only a few tools for legal analytics, but also to explain how these tools work. For instance, many of the data science tools that can be used to extract knowledge from data produce results whose logic is difficult for humans to understand given the number of variables used. The project intends to construct analytical algorithms for suitable tools, capable to explain their operating logic.

The University of Bologna is conducting the LAILA Project, namely, Legal Analytics for Italian Law.[95] Their aims are several. Firstly, they aim to apply analytics technologies—including supervised, semi-supervised, and unsupervised learning—for building an ontology, classifying legal documents, analysing both legislation and case law, extracting "*massime*" (*rationes*) and principles, question-answering, and predicting trends in court decisions. Moreover, they aim also at providing methodological analyses and guidelines for the efficient and ethical deployment of LA technologies and at expanding the understanding of the structure, logic, and dynamic of Italian law in its connection with EU law, using LA tools.[96]Furthermore, the University of Bologna and the AI4Justice laboratory[97]

[94] https://www.predictivejurisprudence.eu. Accessed in June 2022.

[95] https://site.unibo.it/laila/en/people. Accessed in June 2022.

[96] https://site.unibo.it/laila/en/project. Accessed in June 2022.

[97] https://centri.unibo.it/alma-ai/en/news/ai4justice-a-new-lab-for-applied-research-on-ai-and-the-judiciary-system. Accessed in June 2022.

started to create a *corpus* of decisions of the Court of Audit, administrative acts and decisions of the Constitutional Court to discover valuable legal information. However, the aim of this project is to ensure the explainability[98] of the AI systems and the so-called "right of auditability" intended as the right to access code.

4.2. Projects with other Public Entities

The Italian Bar Council is completing a project named *Avvocatura dello Stato 2020*[99] focused on the recognition and classification of documents. The aim is to increase the productivity and the efficiency of the bar council tasks. The approach is devoted to, firstly, rebuilding the organisation processes and human resources (both employees and lawyers) and secondly, simplifying the communication between administrations and stakeholders.

Since September 2017, the Ministry of Justice has been conducting a project titled *Aut Dedere Aut Judicare*[100] from September 2017 concerning criminal law enforcement. The main aim of this project is to detect certain data in different documents (such as arrest warrants, transfers, and extraditions) by applying data analysis and statistics to the field of international judicial cooperation for criminal matters. The Italian Ministry of Justice

[98] I. SHEIKH RABIUL et al. *Explainable Artificial Intelligence Approaches: A Survey*. In *ArXiv Preprint*, n. 2101.09429, 2021.

[99] https://performance.gov.it/performance/piani-performance/documento/1247. Accessed in June 2022.

[100] https://www.camera.it/leg18/410?idSeduta=0278&tipo=stenografico. Accessed in June 2022.

addressed recently an important issue for the AI and Law community, i.e. the semi-automated anonymisation of particular data, named entities in text documents in both civil and criminal legal documents, as well as criminal proceedings. This project focuses on automatically identifying named entities (both physical persons and legal entities) and related information, candidates to be anonymised, by utilising innovative NLP and AI techniques. The ambitious final goal is to solve the common and traditional problem of manual identification and deletion of personal data through legal workflow automation.[101]

5. THE OPEN ISSUES IN THE DIGITALIZATION PROCESS

As described above, today the process of the digitalization of the justice system includes technical and administrative solutions, both crucial for moving from paper-based document flow to a machine-readable document flow, where administrations and private entities can communicate information and share documents in a more efficient way. The achievement of this stage would imply several positive side effects. For instance, it would permit the general public to access a public dataset of legal documents, acts, claims and judgments, in the best scenario, also annotated with keywords and other useful metadata. Thus, these annotated datasets would permit the application of AI and ML techniques, as shown by the literature, where several techniques have been applied for predicting court

[101] The Italian Ministry of Justice participated at the seminar titled: *Finnish Project on the Anonymization of Court Judgments with Language Technology and Machine Learning Apps*, https://www.coe.int/en/web/freedom-expression/finnish-project-on-the-anonymization-of-court-judgments-with-language-technology-and-machine-learning-apps as reported by Prof. Giorgis, https://www.camera.it/leg18/410?idSeduta=0278&tipo=stenografico. Accessed in June 2022.

outcomes and legal argumentation. Besides these innovative approaches, e-justice traditional tasks are only partially completed.

5.1. Data: Access to Caselaw

Indeed, the importance of these tasks in the justice systems of the EU Member States is confirmed by the Justice Scoreboard 2020 of the European Commission.[102] This report is an annual comparative information tool for improving the effectiveness of national justice systems by providing objective, reliable and comparable data on a number of indicators relevant for the assessment of the efficiency, quality and independence of justice systems in all Member States.

One of these indicators is the *online access to published judgments by the general public* is intended as the availability, for each court instance, of all judgments for civil/commercial and administrative and criminal cases online. In particular, accessibility is required throughout the whole justice chain to enable people to obtain relevant information so that the judgment can be swiftly accessed online.

Moreover, the same document intends the *arrangements for producing machine-readable judicial decisions* as to the permission to download judgments and their associated metadata free of charge in the form of a database or by other automated means. Furthermore, in this document are mentioned, not only the approaches for anonymization and

[102] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0306&from=EN. Accessed June 2022.

pseudonymization assisted by algorithms, but also the urgency of putting in place rules to determine whether or not personal data are revealed in online published judgments.

Finally, the study considers if metadata, such as citations, references to national or EU law, keywords, date of the decision are associated with judgments, or if the European Case Law Identifier (ECLI) is assigned. The accessibility to all judgment and the creation of machine-readable judicial decisions (completed with the recalled meta-data) can be seen as two parameters of the efficiency of the justice systems.

5.2. Tech: Communication and Document Management

The urgency for improving traditional tasks in the justice field is confirmed by the Council of the European Union. For instance, the Council adopts new rules to modernize judicial cooperation in taking evidence and service of documents.[103] The position of the Council confirms that it is necessary to further improve and expedite the transmission and service of judicial and extrajudicial documents between the Member States while ensuring a high level of security and protection in the transmission of such documents.[104] Thus, the council points out that efficiency and speed in judicial proceedings in civil matters require that judicial and extrajudicial documents be transmitted directly and by rapid means, and in

[103] https://www.consilium.europa.eu/en/press/press-releases/2020/11/04/digital-europe-council-adopts-new-rules-to-modernise-judicial-cooperation-in-taking-of-evidence-and-service-of-documents/. Accessed in June 2022.

[104] https://data.consilium.europa.eu/doc/document/ST-9890-2020-INIT/en/pdf. Accessed June 2022.

order to enhance electronic cross-border transmission of documents through the decentralized IT-system, such documents should not be denied legal effect and should not be considered inadmissible as evidence in the proceedings solely on the grounds that they are in electronic form.

The issues concerning digitalization of justice are under the spotlight also outside the EU borders. A team within the General Services Administration (GSA), carried out an 11-week path analysis on the federal judiciary's Case Management and Electronic Case Files (CM/ECF) system.[105] The research focuses on user needs, business agility, organization and processes, and the Administrative Office of the U.S. Courts' (AO) culture and legal mandates.

From a user experience, the team observed several instances of application crashes and sluggish response times and gathered several participants' reports about similar issues. Incidents diminish user experience and create distrust and dissatisfaction with the system. Moreover, the same team observed several places throughout the CM/ECF interface that are difficult to use, and some tasks that require repetitive clicking or that may be inaccessible to users with disabilities. More in particular, from this analysis, it emerged that issues, difficulties and questions are similar to every (digital) justice system indeed, judges want to be able to see where they need to be next, and understand the context of a case quickly, all while potentially managing multiple cases, hearings, and appointments each day. Literally,

[105] https://aboutblaw.com/XFW. Accessed June 2022.

this report confirms that "*a white whale for the US judiciary over the last several years has been CM/ECF calendaring*"[106].

Finally, the Law Society presented recently a project called the Future Worlds 2050[107], where also the scenarios of the future of the legal sector are addressed. In this work, it is confirmed again that technology is currently expected to deliver the greatest changes in the legal sector. The involved experts declare that impact of technology on the legal profession particularly the rise of automated self-service legal tools and the implications around global data usage and ownership will tremendously impact the legal sector in the near future.

6. CONCLUSION

To conclude, as described, the digitalization of justice is (and most probably will remain) an enormous ongoing project. Considering the addressed problematics emerged by the mapping of all the projects, it is possible to isolate three main streams, one theoretical and two more technical. Firstly, many difficulties and legal questions about the future of AI and Law are still under the spotlight. Administration of justice needs innovation as same as the judiciary, and the discussion on the topic is still open[108]. The discussion on privacy,

[106] A. BIELEN et al, *Administrative Office of the U.S. Courts CM/ECF, Path Analysis*, Report, 2021, https://www.uscourts.gov/sites/default/files/18f_path_analysis_on_us_courts_cmecf_march_2021_opa_0.pdf

[107] https://www.lawsociety.org.uk/topics/research/future-worlds-2050-images-of-the-future-worlds-facing-the-legal-profession-2020-2030. Accessed June 2022.

[108] R. CAVALLO PERIN, *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale,* in *Quaderni del Dipartimento di Giurisprudenza dell'Università di Torino,* 2021, R. CAVALLO PERIN, *Ragionando come se la digitalizzazione fosse data,* in *Dir. Amm.*, 2/2020, 305-328.

explainability, control, access and human-based judicial decision is mature but yet animated. The two technical streams concern more the administrative daily need of the justice system. The first one is the transition from a paper-based process to a (full) machine-readable process. The main aim for the next future may be the disappearance of paper forms and documents and the completion of public platforms for legal practitioners to deposit (digital) acts, instances and requests. This goal would enhance the second stream, which is the collection, storage and publication of data. Court decision details, such as norms, outcomes and parties are shared in dedicated databases, where not only legal practitioners can access this information. Even if many tasks and procedures will be fully digitalized, a full digital justice system has not been realized yet and it may be a long-term process yet. The authors of *Giustizia 2030* precise that a digitalized justice system is a system that can improve legal analytics and justice prediction, in terms of court outcomes. Such a system would provide judges and prosecutors with informative tools capable to assist legal professionals in their decisional process by providing updated trends in case law, based on the (justice) big data analytics[109]. The depicted state-of-the-art confirms that a complete and generalized digital transition in the justice system is complex and needs a common strategy of each involved stakeholder and the strong impulse of the policymakers.

Abstract. *The digitalization of justice and legislation is coming regularly under the spotlight. Many are the words and the concepts used to express this topic, but today, during the third Artificial Intelligence boom, the field of research named Artificial Intelligence and Law*

[109] During the sharp shock of the first Covid-19 wave in 2020, professionals, members of the judiciary, lawyers and scholars shared the need for a strategy for the entire Italian justice system. Their aim was to develop a general perspective for real solutions capable to transform justice and support the public recovery after the emergency. To this end, they elaborated a white paper, presenting for pillars for the future of justice, that should be, i) (inter) connected, ii) technologically built-in, iii) organized and innovative, and iv) accessible, simple, sustainable. VV.AA. GIUSTIZIA2030, *Un libro bianco per la giustizia e il suo futuro*, 2020. https://irp.cdn-website.com/458fa343/files/uploaded/Giustizia-2030.pdf. Accessed June 2022.

gained attention and relevance. This field of study appears mature and composed of several sub-fields, from legal ontologies for modeling the legal knowledge and expert systems for legal reasoning and arguments to legal analytics and data science approaches. In particular, intelligent tools for assisting the judiciary are currently based on text analysis techniques, which are constantly under development. To test these techniques on real cases, there are promising projects on going, where academia and courts work together for improving the innovation of the entire justice system. In this work, I present the state-of-the-art of the AI and Law approaches, considering lessons learned, weak points and future lines of research, by presenting an overview of the Italian projects and the goals drafted for the justice system for the next decade.