



COURT DECISIONS AND DATA ANALYSIS

*A Survey Among 22 Member States of the European Union
on
Access to Case Law and Legal Prediction*

MATTIA FALDUTI

i-lex

i-lex. Scienze Giuridiche, Scienze Cognitive e Intelligenza Artificiale
Rivista quadrimestrale on-line: www.i-lex.it
Dicembre 2020
Numero Speciale AI and Justice
Fascicolo 1
ISSN 1825-1927

COURT DECISIONS AND DATA ANALYSIS A SURVEY AMONG 22 MEMBER STATES OF THE EUROPEAN UNION ON ACCESS TO CASE LAW AND LEGAL PREDICTION

MATTIA FALDUTI

ABSTRACT. Ignorance of law excuses no one, according to a quoted adage. Indeed, law is fully available and accessible in Official Gazettes by the general public. On the contrary, Court Decisions (CDs) are not as readily available for several reasons. Access to CDs involves many aspects, such as judicial transparency, data governance, right of the defence and access to public data. The lack of a complete European regulation on the topic implies that every EU Member State provides rules on access to CDs in an autonomous fashion. To draft a European context on the topic, a questionnaire has been sent to each institution involved in access to CDs, such as Ministries of Justice, Judiciary Offices, Judge Associations and Supreme Courts, in order to ask them how access to CDs is managed in the EU. The paper presents the results of the survey.

Keywords: Data access and governance, Court decisions, Predictive justice

1. Introduction

Creation, publication and communication of law require the use of language, mostly written. In the last centuries law and the related application have been reported in documents daily produced both by legislative and jurisdictional powers. The set of these documents is constantly (and inevitably) growing, creating large datasets.

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 different application domains and contexts. This is stimulated, on one side, and made possible on the other side, by the continuous production of data coming from disparate data sources and locations and by the availability of web-based technologies for data storage, integration, analysis and mining, thus enabling behaviour and trend prediction as well as descriptive statistics for facts and events.

The legal domain is one of these different domains and contexts where data science approaches are applied. The legal domain has its knowledge: i.e. the legal knowledge, that is composed by Law, the set of rules and regulations and by case law, the set of Court Decisions (CDs)¹.

Contrary to law, CDs report much more (legal) information. Therefore, CDs are a core component of the legal system since a clear and exhaustive understanding of them influences the future interpretation of law and implies new legal effects. However, quantity, complexity, and articulation of CDs are constantly growing. Considering some real-world data, in Table 1 the details of the incoming cases in the Courts of the five most populated countries in the European Union² are presented. The data are extracted from the Justice Scoreboard 2019, a report realized by the European Commission for the Efficiency of Justice (CEPEJ)³. This report concerns the evaluation of the judicial systems (2016-2018 cycle - the last available)⁴ where the efficiency and quality of justice in the European Union have been analysed.

2. Court Decisions in the European Union

First, all incoming cases inevitably will have to be resolved through some sort of court's decisional act, i.e. a CD. All jurisdictions start from millions of cases in Courts of first instance, followed by a dramatic reduction of cases in Courts of third instance. However, even if the trend follows the traditional jurisdictional direction, which

¹ I agree with the notices from the EU Council, where it is clarified that the term 'court decisions' (CDs) covers all types of intermediary and final judicial decisions, under whatever name, rendered by tribunals or Courts as defined by national law.

² Constitutional Courts, Military Courts and Court of Auditors, and the other European or International Courts have been excluded.

³ European Commission of the Efficiency of Justice - CEPEJ: European judicial systems report on the efficiency and quality of justice. Council of Europe (COE), at <https://www.coe.int/en/web/cepej/country-profiles> (2018)

⁴ Data are available at <https://www.coe.int/en/web/cepej/2016-2018-evaluation-cycle-reply-bey-country>, last visit July 2020.

implies a natural reduction of cases through instances, some clear differences emerged among the countries analysed.

Table 1: 2016 Incoming Cases

TOTAL				
Country	First Instance	Second Instance	Third Instance	Total
<i>France</i>	3.253.649	329.688	37.667	3.621.004
<i>Germany</i>	5.699.447	205.634	18.635	5.923.716
<i>Italy</i>	5.102.805	259.774	92.177	5.454.756
<i>Spain</i>	2.756.317	244.200	23.987	3.024.504
<i>UK</i>	4.176.393	14.768	60	4.191.221
Total	20.988.611	1.054.064	172.526	22.215.201
CIVIL				
Country	First Instance	Second Instance	Third Instance	Total
<i>France</i>	2.253.976	282.835	30.018	2.566.829
<i>Germany</i>	4.686.504	142.619	15.591	4.844.714
<i>Italy</i>	3.657.690	135.081	39.793	3.832.564
<i>Spain</i>	1.972.326	184.339	19.956	2.176.621
<i>UK</i>	2.540.573	1.012	52	2.541.637
Total	15.111.069	745.886	105.410	15.962.365
CRIME				
Country	First Instance	Second Instance	Third Instance	Total
<i>France</i>	999.673	46.853	7.649	1.054.175
<i>Germany</i>	1.012.943	63.015	3.044	1.079.002
<i>Italy</i>	1.445.115	124.693	52.384	1.622.192
<i>Spain</i>	783.991	59.861	4.031	847.883
<i>UK</i>	1.635.820	13.756	8	1.649.584
Total	5.877.542	308.178	67.116	6.252.836

For instance, the United Kingdom (UK)⁵ reported a number of first instance cases similar to that of other countries, but for the second and third instances, the number of cases is considerably smaller by comparison. More in detail, considering the total amount of cases in the Supreme Court, the UK reported only 60 cases, whereas Italy for instance has 92.177. The causes of this may be twofold. Firstly, the justice system in the UK is structurally different from that of Italy, where instead of a common-law system, based on case law precedents, the Roman tradition has established a civil-law system. Secondly, the UK Supreme Court is the final court of appeal for all civil and criminal cases and hears appeals on arguable points of law of general public importance. On the contrary, a country like Italy has many more cases in the third instance, specially 2,4 times of the cases of France, that reports 37.667 cases. With regards to this, there are two opposing points of view. On the one hand, widespread access to the court of third instance can be seen as a synonymous of guaranteed and complete access to justice. On the other hand, large numbers of incoming cases can be seen as a lack of justice efficiency and certainty. The UK is also the only country that reports such a threshold between civil and criminal cases in Courts of second instance. Indeed, contrary to the trend followed by the other countries, the UK has 13 times more criminal than civil cases at the court of second instance. The percentage of the cases in the three jurisdiction on the total amount of cases has been calculated. First instance cases compose around 90% of the total, and among these, civil cases are the large majority. Indeed, i) commercial, familiar, contractual, and other sub-domains of private law claims and legal actions, and also ii) the administrative cases, are all included in the civil cases set.

The CEPEJ report aforementioned names this set as *other than criminal cases*. Table 2 confirms a common trend among the most

⁵ In UK, the survey was completed during the Brexit procedure.

populated European countries, with the only exception once again being the UK. The percentages confirm the dominance of criminal cases in the Courts of second instance in the UK (93,14%), which are dramatically higher compared to Italy, the second country for number of criminal cases in the same jurisdiction (48%).

Table 2: Percentage of Jurisdiction and Matter

Country	First Instance (%)			Second Instance (%)			Third Instance (%)		
	Impact	Civil	Crime	Impact	Civil	Crime	Impact	Civil	Crime
France	89,86	69,28	30,72	9,1	85,79	14,21	1,04	79,7	20,3
Germany	96,21	82,22	17,78	3,48	69,36	30,64	0,31	83,67	16,33
Italy	93,54	71,68	28,32	4,76	51,99	48,01	1,7	43,18	56,82
Spain	91,13	71,56	28,44	8,07	75,49	24,51	0,8	83,92	16,08
UK	99,64	60,83	39,17	0,35	6,86	93,14	0,001	86,67	13,33
Total	94,48	72	28	4,74	70,77	29,23	0,78	61,1	38,9

Table 3 shows the rounded fraction between the total amount of cases and the population, in order to count the magnitude of justice on society in 2016. Italy counts one case every 11 people, and France one case every 18 people. Germany appears as the first country for civil cases, with one case every 17 people, but the same country counts only one criminal case every 76 people. Contrary to Germany, Italy reports one criminal case every 37 people.

Considering the total amount of cases and the total sum of the population of the countries in question, for 2016 a total of one case every 15 people is calculated.

Table 3: Population Index

Country	Population ⁶	Cases	Index	Civil Cases	Index	Crime Cases	Index
<i>France</i>	66.809.816	3.621.004	1/18	2.566.829	1/26	1.054.175	1/63
<i>Germany</i>	82.521.653	5.923.716	1/14	4.844.714	1/17	1.079.002	1/76
<i>Italy</i>	60.483.973	5.454.756	1/11	3.832.564	1/16	1.622.192	1/37
<i>Spain</i>	46.658.447	3.024.504	1/15	2.176.621	1/21	847.883	1/55
<i>UK</i>	66.273.576	4.191.221	1/16	2.541.637	1/26	1.649.584	1/40
Total	322.747.465	22.215.201	1/15	15.962.365	1/20	6.252.836	1/51

3. Motivation

CDs provide a lot of information about the application of the law and legal interpreters, such as judges, public prosecutors and lawyers, are daily involved in analysis and evaluation of them. Moreover, besides the typical legal professionals, many other subjects are interested in extract/derive possible suggestions for incoming case applications by relying on the experience of past applications of law. For instance, policymakers and public administrators, enterprises, labor unions, trade unions, police forces, self-employed and individuals are constantly involved in the legal knowledge research from CDs related to their area of interest. Furthermore, academic institutes and research centers are involved in CDs monitor and analysis. For example, the *Oxford Reports on International Law in Domestic Courts* (ILDC)⁷ selects relevant CDs which are relevant to the identification and interpretation of rules of international law, as applied in the domestic Courts of around 70 jurisdictions. The *Human Rights Centre of Ghent University*⁸ developed a project where the main objective is to study the *European Court of Human Rights* case law with the aim of proposing innovative solutions to strengthen the consistency and persuasiveness of the Court's legal reasoning to improve its accountability and transparency.

⁶ <https://ec.europa.eu/eurostat/web/population-demography-migration-projections/data/database>; last visit July 2020.

⁷ <https://opil.ouplaw.com/page/212>; last visit July 2020.

⁸ <https://hrc.ugent.be>; last visit July 2020.

Other institutional bodies, both national and international, are involved in analysis and evaluation of CDs. Examples are many. First, the *French National Observatory of Crime and Criminal Justice* (ONDRP)⁹ created in 2003 aims to produce, collect and disseminate data on crime, criminal justice, and safety issues. Secondly, the *European Intellectual Property Office Observatory* (EUIPO)¹⁰ performs CDs monitoring on the copyright. Moreover, the *United Nations Office on Drugs and Crime* (UNODC) with the *SHERLOC* project¹¹ faces the issues of analyse case law, creating a comprehensive CDs database. But, due to the large number of published CDs, it is physically impossible to read, analyse, systematize and extract information from all the international and national CDs. The need for automation clearly emerges both from real-world applications and from the literature comparison. In this age of legal big data, more and more researchers start to notice that combining traditional doctrinal legal methods and empirical quantitative methods is a promising approach, as reported in a recent work applied on CDs of the ECHR¹². These research approaches need to face access to CDs at first. Access to CDs has been discussed among European institutions. For instance, the representatives of the Governments of the member states meeting within the EU Council for discussing the best practices regarding the Online Publication of Court Decisions (2018/C 362/02). This work¹³ has emphasized that in modern democracies, the Rule of Law requires a transparent judiciary, where citizens have adequate access to the sources of law. The publication of CDs provides insight in how the law is applied by the judge and to this end, continues the Council, modern technologies have revolutionized the way information can be disseminated to the public and retrieved by citizens. However, the online publication of CDs needs to balance a variety of interests. The reported best

⁹ <https://inhesj.fr/ondrp/the-ondrp>; last visit July 2020.

¹⁰ <https://euipo.europa.eu/ohimportal/it/home>; last visit July 2020.

¹¹ <https://sherloc.unodc.org/cld/v3/sherloc/index.html?lng=en>, last visit July 2020.

¹² M. Medvedeva, M. Vols, M. Wieling, *Using Machine Learning to Predict Decision of the European Court of Human Rights*, in *Artificial Intelligence and Law*, 2, 2019, 1-30.

¹³ EU Parliament and Council of European Union: Official Journal of the European Union C 362 (2018).

practices suggest activating the publication of CDs on the internet, specifying that, in the case where only a subset of CDs is published on the internet, publishing the selection criteria is required for transparency reasons. Furthermore, the Council indicates that regarding the protection of personal data in CDs published on the internet, each Member State has to take into account the implications of data protection, considering, for instance, obscuring data while preserving readability and comprehensibility. Moreover, in the CEPEJ report aforementioned, a comparison focused on litigious civil, commercial and administrative cases is presented, in order to assist the Member States in their efforts to generate more investment and businesses as well as creating a more citizen-friendly environment. In this report, access to CDs is listed as one of the four categories selected for measuring the quality of justice. As considered by the Commission, online access to CDs increases the transparency of the justice system and helps citizens and businesses understand their rights. The EU Commission confirmed that the publication of CDs requires balancing a variety of interests, but at the same time, the Commission supports open data initiatives from the public sector, including the judicial system. Access to CDs is directly and indirectly influenced by several factors, such as type of jurisdiction, technological resources and legislative environment. Online accessibility to published CDs and the related arrangements for publication of CDs are reported in the CEPEJ report as well. However, details and differences regarding publication criteria have yet to be addressed. This also goes for issues related to access to CDs with respect to the General Data Protection Regulation (GDPR) and justice outcomes prediction. For these reasons, it emerges the need to understand how access to CDs is managed among all the European Member States.

4. Methodology

I have asked about CDs access to every administration involved, such as Ministry of Justice, Supreme Court, Judiciary Office, and Judge Association. To complete this analysis, a detailed questionnaire to the institutional offices involved has been sent, in order to receive an official answer regarding four main research topics: i) Data

Completeness, ii) Data Access, iii) Data Protection, iv) Data Prediction. The obtained data have been first transferred to an answer sheet, and then, aggregated by topic. The results are discussed below with the help of pie charts and maps.

4.1. Survey Research Topics

Usually, access to CDs is evaluated only considering if the general public has a free and direct access to CDs. But, from an analytical point of view, this information may not be sufficient. For this reason, I have considered the details regarding completeness of the data, actual access modalities and the data protection strategies provided by the. Finally, I also wanted to understand if EU Member States have introduced some regulations concerning prediction of judicial outcomes based on past CDs analysis.

Data Completeness (DC)

Data completeness is rated by quality level and low-quality of data is a serious problem for any mining approaches¹⁴. Incomplete data are a subclass of low-quality data and reduces the usability as well as compromising the analysis. For this reason, data completeness is the first aspect to consider for evaluating CDs datasets, given how access to incomplete data may lead to incomplete analysis and partial results. With respect to access to CDs, I have first considered the level of Data Completeness of the published CDs in Europe, and then researched what the publication criteria are in the countries where only a subset of CDs is published. By Data Completeness I obviously mean the grade of completeness of the available CDs provided by data producers (e.g., Courts) for data consumers (e.g. professionals, institutions, or individuals).

Data Access (DA)

Data access can be considered as the actual possibility of consulting or acquiring detailed information, electronically and for free, without barriers and without being drowned into the maelstrom

¹⁴ Liu, Yong-Nan and Li, Jian-Zhong and Zou, Zhao-Nian, *Determining the Real Data Completeness of a Relational Dataset*, in *Journal of Computer Science and Technology*, 31, 2016, 720-740.

of locating data in a myriad of complex systems¹⁵. Accessing a large data collection of CDs implies two main aspects: *where* the data producer provides access to CDs and *who* can access them. For this reason, it has been emerged the need to investigate if the judiciary allows access CDs on dedicated public databases and if the general public, or conversely only some categories of people, have access to CDs. In other words, by Data Access I mean the access modalities designed and implemented by data producers in order to allow access CDs by data consumers.

Data Protection (DPRO)

The European General Data Protection Regulation (GDPR) touches upon issues related to CDs publication. Data needs to be protected when they are processed, because according to GDPR, almost anything done with data counts as processing, including, for instance, collecting, recording, storing, using, analysing, combining, disclosing or deleting it¹⁶. CDs are data that are constantly processed and provide a great deal of information. CDs may report personal data, such as: identification data of the involved people, criminal records, ethnic background, language and nationality, race and ethnic origin, religious or philosophical beliefs, political opinions, trade union memberships, biometric data used to identify individuals, genetic data, health data, and data related to sexual preferences, as well as sexual habits, and/or sexual orientation. GPDR impacts on court and on justice administration differently. CDs are data that need to be handled according to the GPDR indications, even if each Member State has the power to develop particular strategies that respond to different needs. For this reason, in this survey the issues related to CDs anonymization and to the role of the Data Protection Officers (DPO) inside court administration in the EU countries has been faced. In other words, by Data Protection I mean the strategies

¹⁵ S. Kierkegaard, *Open Access to Public Documents- More Secrecy, Less Transparency!* in *Computer Law & Security Review*, 25, 2009, 3-27.

¹⁶ G. Wilms, *Guide on Good Data Protection Practice in Research*, in <https://www.eui.eu/Documents/ServicesAdmin/DeanOfStudies/ResearchEthics/Guide-Data-Protection-Research.pdf>, European University Institute (EUI), 2019.

designed and implemented by data producers to protect the personal data contained in CDs.

Data Prediction (DPRE)

Data can be analysed both with descriptive and predictive purposes. Description provides information about something that presently exists, existed in the past or may exist in the future. Prediction is the task of estimating the value of a target attribute for a given instance based on the values of other attributes for that instance¹⁷. Thus, predictive modeling and machine learning based on big datasets are currently being applied to almost any domain of research and justice predictive approaches have also been tested. With respect to CDs, it is now possible to predict Courts' outcomes, based on the previous CDs, as proposed in a recent application applied on US Supreme Court's decisions¹⁸. But approaches of predictive justice rise two fundamental issues: i) legal authorization and ii) structural feasibility. For this reason, first it has been investigated whether countries rule these issues with a dedicated regulation. Secondly, w.r.t all the topics presented above (data completeness, access and protection) it has been analysed if approaches of predictive justice can be tested with the current state of affairs. By Data Prediction I mean the permission and the feasibility given to data consumers to test predictive models of Courts' outcomes, based on collections of CDs.

4.2. Survey Questionnaire

Data Completeness (DC): with respect to this research topic, I have investigated i) if Courts publish all their CDs or only a subset of them and ii) if only a subset of CDs is published, I have inquired which are the publication criteria.

Data Access (DA): with respect to this research topic, I have investigated the facilities and the authorization for accessing CDs,

¹⁷ J.D. Kelleher, B. Tierney, *Data Science*. MIT Press, 2018.

¹⁸ D.M. Katz, M.J. Bommarito, J. Blackman, *A General Approach for Predicting the Behavior of the Supreme Court of the United States*, in *PloS One*, 12, 1-18, 2017.

considering which user category has access and where CDs are stored and able to be accessed.

Data Protection (DPRO): with respect to this research topic, I have asked to describe the strategies adopted for protecting the personal data reported in CDs. First, I have asked if the published CDs are completely anonymized. Secondly, considering all the rights and the administrative issues involved in CDs publication and data protection, I have asked whether or not Courts have nominated a Data Protection Officer (DPO).

Data Prediction (DPRE): with respect to this research topic, I have investigated the state of the art for regulations concerning approaches of predictive justice.

Table 4: The List of the Submitted Questions

Research Topics	Questions
DC 1	Do Courts publish all their decisions?
DC 2	What are the publication criteria?
DA 1	Do courts publish CDs on dedicated and public databases?
DA 2	Which user category can access the published CDs?
DPRO 1	Are the published CDs completely anonymized?
DPRO 2	Have Courts nominated a Data Protection Officer (DPO)?
DPRE 1	Do any regulations or restrictions exist regarding the automatic prediction of case outcomes?

These questions have been included in a questionnaire which has been designed as closed-ended questions, with only one exception (DC2). More in detail, **DC1** reports the following list of Courts, *Court of first instance*, *Court of second instance* and *Court of third instance* with three possible answers choices for each court: i) *yes*, ii) *no*, iii) *only a subset*. **DC2** is an open-ended question. **DA1** is a *yes* or *no* question and **DA2** presents a closed set of several options, such as

everybody / only lawyers / only judges or prosecutors / researches / other (please specify). Questions about Data Protection (DPRO) and about Data Prediction (DPRE) are closed-ended questions: **DPRO1**, **DPRO2** and **DPRE1** are all *yes* or *no* questions. To understand how access to CDs is provided in all the 28 European countries I have decided to contact the involved administrations directly. I have identified the followings administrative bodies as those that could answer to the research topics:

Ministry of Justice, because it represents the governments department for the administration of the entire justice system.

Council of the Judiciary, because is the body that allocates jurisdiction and guarantees the autonomy and independence of magistrates, it intervenes in aspects linked to the organisation and good functioning of justice-related services.

Supreme Court, because is the highest court within the hierarchy of Courts in many jurisdictions and it has an overview on the court of lower instances, and may thus have central administrative and statistical offices.

Judge Associations, such as the *Deutscher Richterbund* in Germany or the *Union Syndicale des Magistrats* in France, because as representative bodies of the magistrates, they are able to provide the point of view of the category with regard to access to CDs.

All the offices have been contacted through the official website of the institutions. Some of them have only one general contact or a contact form to fill, whereas others present a list of internal sub-offices mail addresses. In these cases, I have searched for the dedicated internal sub-offices such as,

i) *general secretary* because, being a general office, it could forwards the questions to the appropriate office, otherwise not reported or not accessible from the official website;

ii) *international offices* because the questionnaire has been part of a non-national request of information;

iii) *statistical offices* because often such offices, reporting statistics, have an overview on collection and accessibility of CDs;

iv) *public-relations offices* because as a general office they can forward the questions to the appropriate office, not reported or not accessible from the official website;

v) *data protection officer (DPO) offices* for the institutions listed above, because knowing how data are processed and accessible in the justice system, I have assumed that they would be facing the majority of the research topics.

To this end, I have sent a few questions to the official contact details reported on the institutional webpage of the 28 European Countries.

5. Analysis of the Survey Results

I have received informative answers from twenty-two countries¹⁹, four countries have responded by refusing the questionnaire²⁰ and two countries did not respond at all²¹. For the only open-ended question DC2 I have received various answers, which in turn have been aggregated in three categories:

- *Public interest*, for all the cases where respondents have reported that the Courts select decisions that they account to be of interest for the public or that are important for the interpretation of a law;

¹⁹ Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech, Denmark, Estonia, Finland, Germany, Ireland, Latvia, Lithuania, Luxembourg, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Sweden, United Kingdom.

²⁰ France, Hungary, Italy, Malta.

²¹ Spain, Greece.

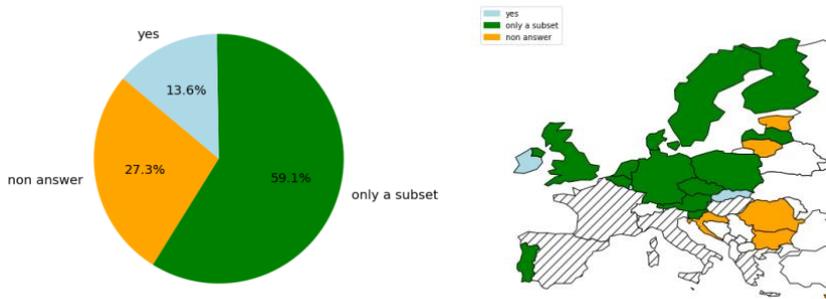
- Jurisdiction policies, for all the cases where respondents have reported that i) some CDs are not published due statutory restrictions, ii) only a subset of CDs is published because each court has its internal policies; and
- Leave to appeal, that occurred only once, for the case where only a subset of decisions is published because parties have their leave to appeal and the trial is not concluded yet.

Moreover, even if DPRO1, DPRO2 have been composed as closed-ended (yes or no) questions, I have received varied answers. In particular, for **DPRO1** some countries have reported: *only a subset*, where only a subset of CDs has been anonymized, and *pseudonymized* where CDs have been pseudonymized instead of anonymized. For **DPRO2**, some countries indicated the office where the DPO was nominated, such as: all federal Courts in Germany and only the Supreme Court in Estonia.

5.1.Data Completeness (DC)

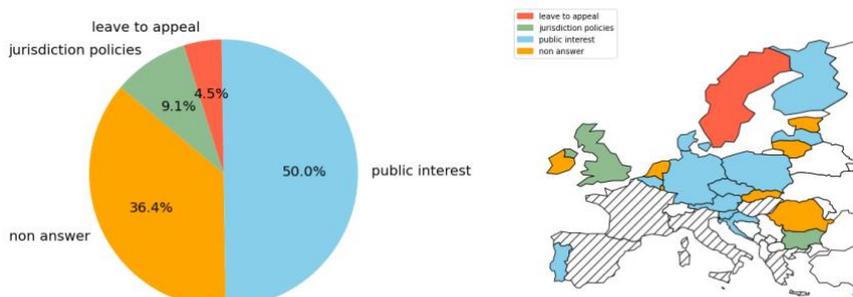
I have decided to group the answers received for DC 1 in one pie chart and one map, in order to summarise the results. The aggregation method was as follows. Based to the possible answers (yes, no, only a subset), I have indicated: *yes* to the countries where all CDs are published in every instance, *only a subset* to the countries that have responded that only a subset of CDs is published, and to last, *non-answer* to the countries that have not answered the question, even for one court.

DC 1 - Do Courts publish all their decisions?



More in detail, CDs from Courts of third instance are mostly published (72,7% of countries). Differently, the percentage drops to 63,9% for CDs of Courts of second instance, but for CDs of first instance the percentage falls dramatically to 27,3%. From this overview, it has been emerged that only the 13,6% of the countries publish all their CDs with respect to every instance, whereas 59,1% publish only a subset of CDs. For this reason, I have decided to investigate what are the publication criteria, i.e. how Courts choose the CDs to publish. I have here reported directly the received answers for DC2. In the majority of the cases, the involved administrations have responded that the main criterium for choosing which CDs to publish is the public interest, intended as the importance of CDs with respect to two main aspects: i) the interpretations of law and ii) the legal issues involved. Only Sweden has reported briefly leave to appeal. I understand this as the fact that in this case the trail is not yet concluded and the CDs cannot be published, until the judgement has become final. Lastly, I have reported jurisdiction policies for all the countries that based the decision of which CDs to publish on internal jurisdiction policies. These could be either legislative, such as regulations on the topic, or administrative, such as Courts discretion. For example, in Bulgaria, the CDs of the Court of Cassation in criminal matters are all published except those containing classified information, and in Germany, each Federal State decides on its own for CDs of lower instance tribunals, but public interest is the main selection criterium. The UK has confirmed that each Court/Tribunal has its own policies on CDs publication.

DC 2 – What are the publication criteria?



5.2.Data Access (DA)

From the obtained answers about **DA 1** it has been emerged that all Courts have a public dedicated database where CDs are published and stored. Denmark has confirmed that the Danish Court Administration was developing at the time a united database for CDs. Concerning user access to Courts public CDs databases, the answers about DA 2 have confirmed that the general public has access to the published CDs, even if many private actors, such as legal publishers, provide access to CDs through pricing plans. Many countries have confirmed the fact that private publishers have access to court decisions too and are populated as same as public databases.

5.3. Data Protection (DPRO)

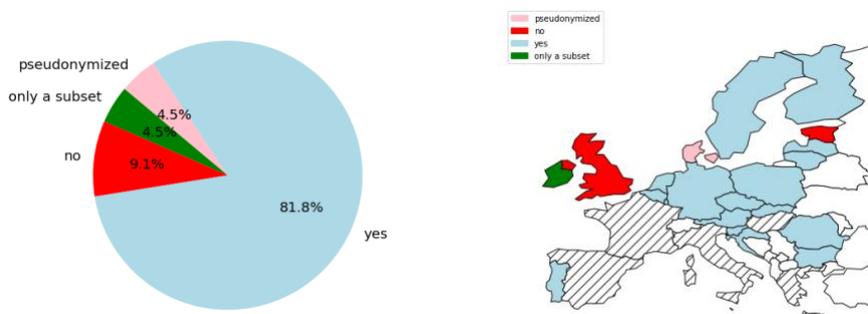
CDs publication guarantees transparency of the judicial action, but at the same time, demands considerations about data governance and privacy in the public administration. The regulation defines that data processing, such as CDs processing, *shall be carried out only under the control of official authority or when the processing is authorised by Union or Member State law providing for appropriate safeguards for the rights and freedoms of data subjects. Any comprehensive register of criminal convictions shall be kept only*

under the control of the official authority. From the answers received about **D PRO 1** it has been emerged that most of the countries anonymise their published CDs (81,8%) confirming the relevance of the involved issues and the practical solution adopted. Only Denmark has reported that personal data in CDs are protected through pseudonymization strategies. The UK and Estonia have adopted a full transparency policy, whereas Ireland has decided to anonymise only a subset of CDs. The answers received present a fragmented approach to data protection.

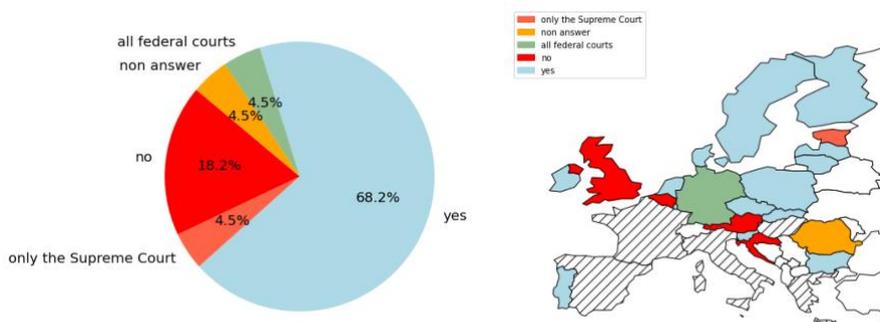
Furthermore, GDPR imposes for every public administration the adoption of a Data Protection Officer (DPO).

Considering the European inclusive perspective of the definition of public administration, Courts can be considered not only as the expression of the judicial power, but also as a public administration, in particular as justice administrations, considering all the public services and administrative powers that Courts offices have, and that are not strictly jurisdictional activity. For example, Courts sometimes express also typical administrative powers, especially when they i) provide access to criminal records or to other documentation, ii) are proposing public competitions, iii) provide public recognition of qualifications, or finally iv) handle the issues related to political elections. Taking into account the large definition of 'public administration, drafted by the European Courts, for identifying an institution as a public administration, it is necessary to consider the real power expressed by this institution. It is clear that Courts, for their typical judicial action, are not included in the recipients of the General Data Protection Regulation (GDPR), but for the additional administrative powers they have, Courts could be included with respect to the DPO nomination. From the answers received, the large majority of the Courts have nominated a DPO, with some specification. In Germany, all Federal Courts have nominated a DPO whereas in Estonia only the Supreme Court.

D PRO 1 – Are the published CDs completely anonymized?



D PRO 2 – Have Courts nominated a Data Protection Officer (DPO)?



5.4. Data Prediction

The answers received to **D PRE 1** show how the large majority of the countries has no regulation on the topic of automatic prediction of judicial cases. For the moment, I can exclude only France from the set, considering the last regulation that has been introduced in the country regarding prediction systems based on the analysis of the

judges' behaviours²². From the answers received only Romania has not responded to the question. The reported lack of regulation in this field would potentially allow the implementation of automated models capable of predicting judges' behaviours and judicial outcomes. Such approaches, as discussed above, have been already tested. However, the development of justice prediction approaches requires two practical conditions: i) full access to all CDs and ii) a compatible regulatory environment. If both these conditions are realized, it will be possible to move forward on that field of research.

6. Conclusion

As a conclusion, I can affirm that access to CDs across the European countries is not completely guaranteed. Only some countries provide full access to all CDs. Considering the results for Data Completeness and Data Access, I can state that every Court (of each country that has responded) has a public database that is accessible by the general public, but full access to CDs is guaranteed only where Courts publish all their decisions (13,6%). Generally, citizens, as well as lawyers, prosecutors, judges and administrators, have access to public databases provided by the administrations but do not have access to all CDs at every Court level. Considering the impact that cases in Courts of first and second instances have on the total amount of cases, access to Supreme Court decisions permits access to less than 2% of all CDs, and even access to the most relevant CDs of second instance Courts is not enough to fill the gap. From this survey emerged the lack of a common European administrative strategy on CDs open data. Additionally, simply publishing CDs does not imply the availability of a ready-to-use dataset which can be used for data mining approaches. Usually, public facilities provided by the administrations, do not allow the direct download of bulk data, metadata or, more in general, a dedicated Application Programming Interface (API). Indeed, other platforms do not allow automatic and massive downloads, literally, as a term of use. Moreover, from the answers received I have noticed

²² LOI n 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice (1) - Article 33.

some conflicts on the same questions, where different institutions of the same country reported contradictory responses. For example, on Data Protection, three countries reported conflicting responses on the question regarding the DPO nomination, where one administration confirmed the nomination and the other denied it. In these cases, I have considered the nomination as confirmed, assuming the reported contradiction as a result of miss-understanding or ineffective communications between administrations. It appears likely that one administration ignores the nomination done by another one. Regarding access to CDs, it is useful to mention that the Estonian Official Gazette is directly connected to the Court's information system²³, allowing direct and free access to both law and case law.

To sum up, the main aim of this research has been to draft the European state of the art concerning access to CDs, believing that the access to CDs is crucial, not only for the development of quantitative legal research, or for legal data analytics for the justice system, but mostly for the dissemination of the legal knowledge throughout the society and the consequential development of the democracy.

²³ More in detail, the Estonian Circuit Court pointed out that *data is available and systematized analysis of CDs is provided in order to make it possible to determine judicial practices, which in the past could be objectified via hypotheses. According to their report, the mentioned analysis gives litigants greater predictability as regard CDs. The same analysis gives judges, either individually or as part of the judiciary, viewed from a more collective perspective, an awareness of failings or assumptions that they have not identified in their practice, thereby making progress possible.*