INTRODUCTION

THE USE OF FORMULAS BY TAX COURTS: MYTH OR REALITY?

Piera Santin – Alessia Fidelangeli*

CONTENTS:

1. POLINE: From an idea, a project takes shape

1.1. A brief introduction

This project pursues two interconnected objectives involving the use of Artificial Intelligence (AI) technologies: first, to empirically demonstrate the existence of the specific legal concept of *Judicial Interpretative Formulas* (see par. 2.3.) through their automatic extraction; and second, to organise legal knowledge in a way that enhances its accessibility, particularly for legal professionals with limited IT expertise.

The project was designed to be innovative in both its core

---

* Piera Santin is Research Associate at the Robert Schuman Centre for Advanced Studies of the European University Institute; Alessia Fidelangeli is Research Fellow at the University of Bologna.

domains: legal analysis and computer science. To achieve this, a robust interdisciplinary methodology was developed, which stands as one of the most significant results of our work.

From the outset, we made a deliberate choice to focus on Value Added Tax (VAT), as it represents one of the most harmonised areas of EU law. Over the years, case law from the Court of Justice of the European Union (CJEU) has played a pivotal role in fostering such harmonisation (see Ch. IV). VAT is also the field with the highest number of preliminary rulings, making it a particularly fertile ground for identifying interpretative patterns and verifying their consistency across jurisdictions.

Within the VAT domain, we further narrowed our focus to judgments concerning exemptions for public interest and the definition of taxable amount. This thematic constraint allowed us to manage computational costs effectively, while still ensuring that the methodology is scalable and applicable to broader areas of case law in future developments.

A brief account of the POLINE project's development is essential, as its evolution has directly influenced the project's design. When the proposal was first drafted, Large Language Models (LLMs) had not yet become prominent in public or academic discourse. At that time, the collaboration between legal scholars and data scientists focused on applying other Natural Language Processing (NLP) techniques to identify relevant portions of legal texts, primarily through manual annotation (see Ch. X).

We subsequently adopted both technologies for complementary purposes. LLMs were employed for the initial extraction of JIFs and for linking these with a structured legal ontology (see Ch. IX and XI). Later, NLP was used to explore an alternative method for JIF extraction from new decisions (see Ch. X), this was necessary to implement the automatic detection module of the platform (see Ch. XII).

Our second goal was to build a tool that improves the usability of AI-extracted legal knowledge. Legal research tools often yield unsatisfactory results, largely due to poor classification and retrieval systems. To address this, we created a structured taxonomy as the foundation for organising the automatically extracted information and employed AI techniques to classify each JIF accordingly (see Ch. XI).

1.2. Extracting interpretation: The idea behind JIFs

To clarify the focus of our knowledge extraction process, we must first introduce the concept of JIFs. The starting point for this work was the observation that the Court of Justice of the European Union (CJEU) frequently uses recurring interpretative statements, typically corresponding to a single paragraph, which are reproduced verbatim across multiple decisions. Over time, this practice has generated a corpus of interpretative rules that has significantly influenced national case law.

As discussed in Paragraph 2, although many legal scholars and practitioners intuitively recognise this phenomenon, it remains underexplored in the literature, in part because its existence is difficult to demonstrate through purely qualitative methods. The repetitive and formulaic nature of the phenomenon demands empirical validation.

The POLINE project was initiated to identify the features that characterise these formulas, with the aim of enabling their automatic detection through AI. This involved both theoretical analysis and empirical research, which informed the drafting of detailed annotation guidelines (see par. 2.2.). These guidelines were used for manual tagging of selected judgements (training the NLP models, see Ch. X), and for creating prompts used in LLM-based experiments (see Ch. IX).

Our hypothesis is that the relevance of JIFs extends beyond the EU legal system, impacting national legal systems, especially in harmonised fields of law as VAT. To test this, we included three national jurisdictions in our study (Bulgaria, Italy, and Sweden) following established methods of comparative legal analysis (see Ch. V, VI and VII). Legal-theoretical and empirical research was conducted across the EU and the three national legal systems, allowing us to examine whether similar formulaic interpretative practices could be identified in national case law.

For this comparative work, we adapted the original EU guidelines to fit each national context. This ensured both methodological consistency across datasets and sensitivity to the unique characteristics of each legal tradition.

1.3. The importance of classification: AI for a satisfactory retrieval system

The second major challenge addressed in this project concerns the accessibility of JIFs after their extraction. Beyond simply compiling these interpretative statements and enabling Boolean search functionalities, we placed particular emphasis on their systematic classification according to legal content and domain relevance.

Current legal databases suffer from two persistent limitations: they typically consist of full-text judgments or, less frequently, summaries of entire decisions, and they rarely isolate or highlight key interpretative passages. Moreover, even at the level of full judgments, classification systems tend to be inconsistent, leading to ineffective retrieval mechanisms. Legal professionals often face time-consuming searches that produce limited or imprecise results, especially when seeking specific legal reasoning or formulaic interpretations.

In addressing this gap, we observed that classification is not merely a technical convenience, but a structural necessity in enabling meaningful access to legal knowledge. During our analysis of CJEU case law (see Ch. IV), we noted a recurring practice: judges frequently reference precedents from domains outside the immediate subject matter of the case. This happens to address procedural issues, to invoke general principles (e.g. proportionality), or to bolster reasoning with analogical references. This interpretative practice suggests that access to JIFs, independently from the full text of judgments, could be highly beneficial for practitioners engaged in the application of EU law.

To make such access possible, we developed a legal ontology grounded in the EuroVoc classification system, as adopted by EUR-Lex, and enriched it with domain-specific keywords derived from the case law corpus. However, the massive and consistent classification of extracted JIFs (amounting to several thousand distinct entries) would not have been feasible without the application of LLMs.

Thanks to the inferential and semantic capabilities of LLMs, we were able to assign each JIF a precise label within the ontology, reflecting not only its linguistic content but also its legal substance. This process allowed us to maintain a high level of internal consistency, a key requirement for any retrieval system intended to support structured and comparative legal analysis.

In practical terms, this means that users of the platform can conduct searches not only through Boolean expressions or keywords but also by navigating through an ontology-based classification interface. Each JIF is accessible through thematic categories, and users can move seamlessly from one formula to others within the same classification group. Additionally, the platform allows results to be filtered according to jurisdiction enabling both targeted research and broader comparative exploration in the EU, Bulgarian, Italian, and Swedish legal systems.

Ultimately, the integration of LLMs into the classification process represents an important aspect of the POLINE project. It transforms unstructured legal language into structured legal knowledge, making it possible to retrieve interpretative content with unprecedented precision and coherence. This advancement is crucial for our long-term goal: enhancing legal information systems in a way that supports transparency, consistency, and accessibility across legal systems.

2. The definition of Judicial Interpretative Formulas: Methodological remarks

The methodology underlying this project is complex, as POLINE included different activities: the comparative legal analysis of case law in the field of VAT; the extraction of JIFs; the creation of a multilingual ontology; the association of JIFs with that ontology; the extraction of citations to earlier case law contained in JIFs; network analysis between JIFs; and the examination of similarities between different JIFs. Each of these works followed its own methodology, which will be discussed in detail in the chapters devoted to each topic.

In this introductory chapter, our focus is on several methodological considerations relating to the definition of the concept of JIF (central to the project) rather than on the specific activities undertaken.

Drawing on our prior experience on AI and law, we identified three guiding priorities.

First, because the project concerned civil law jurisdictions, we aimed to develop a novel conceptual framework for identifying relevant interpretative statements. This framework departed from approaches developed in common law systems, particularly those prevalent in the United States (see par. 2.1.), adapting them to the features of civil law legal systems.

Second, before starting the extraction through LLMs and NLP, experts in tax law and legal theory carried out an in-depth theoretical and empirical analysis concerning the features of national case law in the tax domain. This preliminary work required considerable time and amounted to the theoretical foundation for the subsequent stages of the project (see par. 2.2.).

Third, lawyers were involved in the preparation, execution, and validation of every phase of the project, and they revised the underlying assumptions guiding activities where necessary (see par. 2.3.).

2.1. The need to depart from common law concepts and approaches

A central challenge of the project concerned the definition of the object of our analysis, i.e. what, precisely, we sought to extract. Our experience in analysing both national and European case law made us certain that both jurisprudences make use of interpretative statements concerning the interpretation and application of VAT legislation. Extracting and classifying these statements could be very useful to interpreters for many purposes: granting legal certainty; investigating the coherence between national and European case law; examining the correctness of citations of previous case law; etc. The existence of these statements is mentioned here and there in literature (from the perspective of the doctrine of precedent; to criticise the creative role of case law; to show that the excessive workload of supreme courts). However, these statements are not analysed in themselves and, especially, are not given a name. This difficulty is reminiscent of a long-standing problem in American jurisprudence, famously captured in the observation that “dictum is one of the commonest yet least discussed of legal concepts. Every lawyer thinks he knows what it means, yet few lawyers think much more about it”.

A bibliographic analysis made it clear that there were well-developed conceptual tools within the common law tradition for identifying and categorising judicial interpretative statements, but none of these frameworks could be transposed without distortion into the civil law context in which our work was situated.

In the United States, the debate has long revolved around the distinction between *holding* and *dictum*[2]. A *holding* is generally defined as that part of a judicial opinion which is "necessary" to the outcome of the case, whereas *dictum* is any statement in a judicial opinion that does not form part of the holding[3]. Yet, as the literature repeatedly acknowledges, this distinction is far more elusive in practice than such concise definitions suggest. For nearly a century, American lawyers, judges, and scholars have debated where the boundary lies, and what practical consequences should follow from it[4]. The intensity of the *holding/dicta* debate in common law jurisdictions stems from the structural role of *stare decisis*: in a system where the holdings of higher courts bind future decisions, accurately identifying the "holding" is essential, as it determines the scope of a decision's precedential effect. Indeed, there is an existing literature focusing on the possibility of their automatic extraction through LLM[5].

By contrast, the interpretative statements we initially sought to identify are not embedded within a doctrine of binding precedent. The CJEU does not operate under a strict *stare decisis* regime, although it accords substantial weight to its earlier decisions, frequently citing them verbatim through what might be termed a "copy-paste" technique, and never expressly admitting to overruling itself[6]. National courts, too,

---

[2] A very famous work is *Dictum revisited*, in *Stanford Law Review*, 1952, p. 509 ff. See also M. Abramowicz, M. Stearns, *Defining Dicta*, in *Stanford Law Review*, 2005, p. 953 ff.; L. Alexander, *Constrained by Precedent*, in *Southern California Law Review*, 1989, p. 1 ff.; H.J. Friendly, *In Praise of Erie - and of the New Federal Common Law*, in *New York University Law Review*, 1964, p. 383 ff.; A. Goodhart, *Determining the Ratio Decidendi of a Case*, in *Yale Law Journal*, 1930, p. 161 ff.

[3] K. Greenawalt, *Reflections on Holding and Dictum*, in *Journal of Legal Education*, 1989, p. 431 ff.

[4] For a reconstruction of the debate, see J.M. Stinson, *Why Dicta Becomes Holding and Why it Matters*, in *Brooklyn Law Review*, 2010, p. 219 ff.

[5] C. Arvin, *Identifying Legal Holdings with LLMs: A Systematic Study of Performance, Scale, and Memorization*, arXiv preprint – arXiv:2505.02172, 2025.

[6] F.X. Millet, *In the name of analogy: Judicial copy-pasting and*

increasingly cite prior decisions, including those of the CJEU, especially in harmonised areas of law (see Ch. IV).

Nonetheless, the absence of a binding precedent rule in these jurisdictions means that transplanting [7] the common law *holding/dicta* dichotomy into our context would have been misguided. For this reason, we developed a new conceptual category, initially termed *Judicial Principles of Law* (JPOLs), inspired by the Italian doctrinal notion of *principio di diritto* as applied to the Italian Supreme Court (see Ch. V). As will be explained in the following section, we later replaced this terminology with *Judicial Interpretative Formulas* (JIFs), a conventional concept with a more explicitly European origin and a closer fit with the aims of our project.

Yet one feature is common to both the *holding/dicta* debate and the JPOL/JIF methodology and has been highlighted in American scholarship [8]. In earlier decades, legal research typically began with index searches, followed by reading entire cases to determine whether they addressed the relevant legal issue. Contemporary legal research, however, is mediated by electronic databases, which deliver search results by pinpointing keywords or phrases within opinions. This allows researchers to find an isolated sentence that appears to support their position without necessarily reading the remainder of the judgment to understand the significance of the quoted statement in relation to the decision as a whole.

The same issue arises in the context of JIFs: an isolated judicial statement may seem authoritative when read in isolation, but its actual relevance can only be determined in light of the decision. This is why, in the POLINE platform, for each JIF we always add the link to the whole decision.

2.2. From theory to practice: The need for a robust comparative framework before extraction

Too often in legal analytics projects, the practical phase,

---

competence creep in the connection data case law, in *Common Market Law Review*, 2024, p. 1289 ff.

[7] On legal transplants see A. WATSON, *Legal Transplants: An Approach to Comparative Law*, University of Georgia Press, 1974.

[8] T. FOWLER, *Holding, Dictum... Whatever*, in *North Carolina Central Law Journal*, 2003, p. 139 ff.

particularly the use of LLMs for data extraction, commences without a sufficiently precise definition of the object of study. In our case, a wide was essential to determine, from the very beginning, the precise target of extraction.

The theoretical analysis consisted in a detailed examination of the national legal systems of the participating partners. Each partner was tasked with preparing a study of their domestic tax law system, including a concise description of which courts have jurisdiction over the merits of tax cases, the identity of the court of last instance, whether that court is administrative or judicial in nature, and whether it may decide the merits of a case or is confined to reviewing questions of law. They were also required to describe the structure, content, and decision-making style of the tax judgments of their supreme courts [9].

Given the absence of a uniform concept of relevant interpretative statements in case law across jurisdictions, the core of this analysis was a report on the existence (or absence) of the concept of a "judicial principle of law" in each national system. This inquiry was inspired by the Italian concept of *principio di diritto*, a doctrinally notion linked to the nomophylactic function of the Italian Supreme Court (see Ch. V). Partners were asked to determine whether analogous concepts existed in their jurisdictions, whether public databases containing such principles (or the *rationes decidendi* of supreme courts) were available, and how such resources were used by tax judges and practitioners.

This work also addressed the ongoing debate over whether the CJEU functions more like a common law or a civil law court, particularly in terms of the weight accorded to previous decisions. Partners were therefore asked to describe the role of precedent in their national tax case law. They were also asked to report on any national debate concerning the creative role of judges in tax law, especially where judge-made law might raise constitutional concerns in light of the separation of powers or the principle, present in some systems, that tax legislation must have a statutory basis.

In addition, partners were asked to provide information on

---

[9] Taking inspiration, as to the strcture, from the famous work performed by N. MACCORMICK, N.S. SUMMERS, *Interpreting precedents. A comparative study*, Routledge, 1997.

the role of case law in the correct interpretation of national tax law and the role of CJEU and national courts in ensuring uniform application of EU law in the VAT field. Particular attention was paid to whether national supreme courts cite CJEU judgments directly and to the frequency and manner of such citations.

The outcomes of this research amount to the basis of the analysis presented in Ch.s IV, V, and VI.

The second stage was an empirical analysis in which partners examined judgments in a selected subdomain, identifying examples of JIFs and analysing structural features of decisions, such as whether they were divided into paragraphs and whether interpretative statements appeared within single or multiple paragraphs. The empirical analysis also assessed the frequency and uniformity of citations to precedent, both national and EU, and determined whether the identified JIFs concerned VAT specifically or addressed procedural or other issues. Further inquiries examined the level of abstraction of JIFs, the presence of factual references, and the frequency with which courts overruled or distinguished prior JIFs. Finally, partners reported on whether such statements were explicitly labelled as judicial principles within their domestic systems, albeit under different terminology.

Only after these theoretical and comparative analysis we defined the concept of what we called JIF. We defined a JIF as a portion of text, extracted from the argumentative part of a judgement which contains the:

- interpretation of a rule, of the portion of a rule, or of a general principle; or
- consequences stemming from the interpretation/application of a rule or a principle in a legal system; or
- subsumption of a fact within a rule; or
- qualification of a factual hypothesis as a concept contained within a rule.

This definition emerged from the collective observations of the partners and was formulated jointly by experts in tax law and legal theory. The definition was validated by all partners and subsequently presented to the supreme courts of the participating states during testing events. On the basis of these findings, guidelines for the manual annotation of JPOLs were drafted. Only then did we proceed to the technical phases of the project: the extraction of JIFs, the construction of the ontology, the network analysis, and the related computational activities.

2.3. From JPOL to JIF: Towards a concept aligned with European and national frameworks

In the initial stages of the project, we employed the term *Judicial Principle of Law* to designate the category of interpretative statements we sought to identify. However, it soon became apparent that this terminology risked confusion. In particular, it could be mistaken for the broader legal notion of a "principle of law" in the sense of legal principles such as legal certainty, subsidiarity, or proportionality. This would have been conceptually problematic, as our focus was on interpretative statements emerging from judicial reasoning.

As the project progressed, the analysis of both national and European case law revealed an important characteristic of VAT-related case law: the recurrence of identical or near-identical textual passages across different decisions. This phenomenon is striking in EU case law, but it is also increasingly present in national decisions, particularly recent ones. National courts, especially in tax matters, frequently reproduce interpretative statements verbatim over time, and in many cases, they quote European case law directly and literally (see Ch. VIII).

The relevance of this observation became clearer considering scholarly debates on the so-called "competence creep" of the European Union [10]. This is the gradual and sometimes opaque expansion of EU competences which can be discerned in the case law of the CJEU. The VAT field provides a clear example: both European and national legal literature acknowledge the role of CJEU case law in driving harmonisation, and virtually all scholarly works on VAT devote attention to judicial decisions (see Ch. III). The huge volume of VAT-related case law has even prompted an

---

[10] S. GARBEN, *Competence Creep Revisited*, in *Journal of Common Market Studies*, 2017, p. 1 ff.; M.A. POLLACK, *Creeping Competence: The Expanding Agenda of the European Community*, in *Journal of Public Policy*, 1994, p. 95 ff.; S. WEATHERILL, *Competence Creep and Competence Control*, in *Yearbook of European Law*, 2004, p. 1 ff.

institutional shift, with jurisdiction over certain VAT matters being transferred from the CJEU to the General Court [11].

Another well-documented feature of CJEU jurisprudence is its distinctive drafting style, frequently described as "copy-pasting" (see par. 2.2.). This practice, which emerged in the late 1970s as the length of the Court's judgments increased and a mass of prior case law became available for citation, involves reproducing substantial passages from earlier judgments, verbatim or with slight modifications, into later ones, irrespective of their precise legal status (e.g. *ratio decidendi* or *obiter dictum*) [12]. In some judgments, block quotations from earlier cases dominate the reasoning.

Legal scholarship has examined a specific aspect of this practice of reproducing literally earlier binding interpretative statements in subsequent case law using the concept of "formulas" [13]. This stylistic approach, now pervasive in CJEU decisions [14], is beginning to appear in national case law as well, particularly in courts facing heavy caseloads, such as the Italian Court of Cassation (see Ch. V).

In light of these considerations, we concluded that the concept of "formula" more accurately captured the phenomenon we intended to analyse. We therefore adopted the term *Judicial Interpretative Formula* in place of *Judicial Principle of Law*. In a harmonised field such as VAT, it is perhaps inevitable that concepts originating within the European legal system are better suited to the analytical task than those grounded exclusively in domestic traditions. Consistent with the existing literature, we chose to extract all

---

[11] Regulation (EU, Euratom) 2024/2019 of the European Parliament and of the Council of 11 April 2024 amending Protocol No 3 on the Statute of the Court of Justice of the European Union.

[12] F.X. MILLET, *In the name of analogy*, cit., p. 1291.

[13] L. AZOULAI, *La fabrication de la jurisprudence communautaire*, in P. MBONGO, A. VAUCHEZ (eds.), *Dans la fabrique du droit européen. Scènes, acteurs et publics de la Cour de justice des Communautés européennes*, Bruylant, 2009, p. 153 ff.

[14] On the linguistic–semiotic perspective regarding the drafting style of the Court of Justice and its impact on the development of EU case law, see K. MCAULIFFE, *Hybrid Texts and Uniform Law? The Multilingual Case Law of the Court of Justice of the European Union*, in *International Journal for the Semiotics of Law*, 2011, available at SSRN: https://ssrn.com/abstract=2208686 (last access: September 1st, 2025).

JIFs, including those not constituting the *ratio decidendi* of the case.

3. Conclusion and overview of the book's structure

The project shows how combining comparative legal studies and NLP techniques allows identification and extraction of JIFs. Even though often unnoticed by doctrine, they play a structural function in the transmission of legal meaning across courts and jurisdictions. Their recognition and classification, therefore, are not merely technical achievements but epistemological steps toward making the architecture of judicial interpretation empirically observable. The following chapters develop this premise, each addressing a specific dimension of the relationship between interpretation, technology, and the law.

Ch. I traces the historical trajectory of artificial intelligence in judicial activities through the lens of legal informatics. It reconstructs how successive computational approaches have interacted with legal reasoning. The chapter shows that far from being an external instrument, AI has evolved within the conceptual and procedural structures of the law itself.

Ch. II explores the modification of legal interpretation in the age of AI. It examines how interpretative reasoning continues to support legal practice as machine-assisted systems increasingly support decision-making.

Ch. III introduces the legal and institutional context of VAT law in the European Union, explaining how the CJEU has progressively shaped the system through its interpretative activity. The chapter provides the doctrinal foundation for understanding why VAT constitutes an ideal field for the study of JIFs.

Ch. IV offers a diachronic analysis of JIFs within VAT case law. It examines how the drafting style of the CJEU and the practice of citation and repetition have contributed to the emergence of interpretative formulas as building blocks of the *acquis communautaire*. The chapter also discusses the institutional implications of the 2024 reform transferring VAT preliminary rulings to the General Court, which may further consolidate the circulation of such formulas.

Ch. V focuses on the Italian legal system, analysing the interaction between the newly defined JIFs and the traditional notion of *principio di diritto* elaborated by the Court of Cassation. It explores whether the extraction and classification of interpretative formulas may support greater coherence in national case law and contribute to the broader debate on the creative function of the judiciary.

Ch. VI examines the Bulgarian context, where the Supreme Administrative Court plays a crucial role in ensuring uniform interpretation of tax law. The chapter analyses interpretative judgments in VAT matters as expressions of judicial reasoning and legal certainty, while also addressing the national discussion on the role of judges in shaping administrative and fiscal law.

Ch. VII investigates the Swedish Supreme Administrative Court's approach to interpretative stability and doctrinal change in VAT jurisprudence. Drawing on a comprehensive corpus of judgments, it analyses the techniques of overruling and distinguishing and shows how interpretative authority is maintained in a system without formal *stare decisis*.

Ch. VIII explores the concept of judicial interaction and its analytical connection with JIFs. It demonstrates how the mapping of JIFs across jurisdictions show the patterns of dialogue and mutual influence among European courts. The chapter also identifies possible applications of JIF-based analysis for comparative law, judicial cooperation, and the study of transnational reasoning.

Ch. IX describes the creation of the multilingual dataset underpinning the entire project. Curated by experts in tax law, it compiles VAT-related decisions from the CJEU and national courts in Italy, Sweden, and Bulgaria, thus forming the empirical basis for the extraction and classification experiments.

Ch. X presents the methodological and technical foundations of the machine-learning framework developed for the automatic extraction of JIFs. It critically reviews earlier approaches to identifying legal principles and explains how new models based on large language technologies can balance automation with transparency and expert validation.

Ch. XI relies on that framework to construct a multilingual ontology of VAT law and to perform the classification and network analysis of extracted JIFs. By combining LLM-based inference with ontology engineering, the chapter illustrates how interpretative formulas can be connected, compared, and studied as interlinked nodes in a semantic and citation network.

Ch. XII concludes the volume by presenting the POLINE Pilot Tool, the practical implementation of the project's research. The tool integrates NLP, ML, and ontology-based reasoning to provide a transparent, multilingual, and user-oriented platform for exploring interpretative reasoning in tax law. The chapter discusses its potential to improve consistency in judicial decision-making, foster transparency in the use of AI within courts, and promote more equitable access to legal information across jurisdictions.

This book brings together theoretical, comparative, and computational perspectives to examine how interpretation can be improved by tools developed using AI. It may contribute to a new branch of empirical legal research and to a deeper understanding of how legal meaning circulates within the European judicial space.