Foster Transparency of Judicial Decisions and Enhancing the National Implementation of the ECHR (TJENI)

Report on the Workshop

on summarisation of judicial decisions

held on 15-16 May 2023 in Nicosia (Cyprus)

The project "Foster Transparency of Judicial Decisions and Enhancing the National Implementation of the ECHR" (TJENI) is funded by Iceland, Liechtenstein and Norway through the EEA and Norway Grants Fund for Regional Cooperation

Table of Contents

The project "Foster Transparency of Judicial Decisions and Enhancing the National Implementation of the ECHR" (TJENI) aims to support beneficiary countries in the improvement of consistency between national jurisprudence and human rights standards set in the case law of the European Court of Human Rights (ECtHR).

This workshop was planned under the Project in order to bring together IT experts and representatives of the judiciaries of Project's beneficiaries to discuss available IT solutions that can be used by national judiciaries with the aim to improve consistency in national jurisprudence and reinforce ECHR implementation at national level. The workshop was focused on: (1) national initiatives in the area of summarisation and indexing /categorisation of judicial decisions; and (2) testing of existing IT tools that support automated summarisation of judicial decisions, with a special focus on Greek language.

SESSION 1. National initiatives in the area of summarisation and indexing /categorisation of judicial decisions

1. French experience

Representatives of the French Court of Cassation (Auditor, and data scientist) presented the project of the Court of Cassation aimed at detection of jurisprudence divergences. The project started in 2021 as one of the 21 projects selected within the «LabIA» initiative of the interministerial Direction in charge of digitalisation DINUM[1] to support administrations in the deployment of their AI projects benefitting from innovative technologies.

The goal of the French project is identifying judicial decisions with different interpretations of the law (except for the cases when the judges specifically departed from the set jurisprudence which shall be indicated in the reasoning of the judicial decision) through automated comparison of the summaries of the decisions. The Court of Cassation's auditors prepare summaries of the decisions, with several sentences of free text and certain number of keywords. The Court of Cassation uses a set of several thousand keywords (which number is growing) arranged in 12 hierarchical levels. A preliminary step for the identification of divergences is to identify pairs of judgments which deal with a similar matter (i.e. have the same keywords). The tests run under the French project proved that the keywords comparison is quite effective for the set goal.

Published decisions represent only 10% of all decisions adopted. To scale up the summarisation and categorisation of the decisions (which is manual at the moment) an authomatisation is needed to: 1) create a summary of decisions; 2) assign (propose) the keywords on the basis of its summary. Deep machine learning techniques used for translation was tested in order to automate the process. The input for the summary was reduced to a few relevant paragraphs from the judgment, and the summary was further used as an input for the keywords assignment (proposal). As training set used composed ~145.000 of the ~165.000 available summaries, the keywords proposed for ~80.000 published decisions of the Court of Cassation. Then this method was applied to ~800 unpublished decisions with the result confirmed as correct by humans in ~75% of the cases.

Representatives of the French Court of Cassation shared also their experiences with the annotation tool used for the pseudonymisation of judicial decisions based on Name Entity Recognition (NER) approach[2] and the dedicated software "Label" they created for reviewing such annotations and performing corrections[3]. A tool they quickly created[4] to obtain short French-style "titles" for texts in Greek was demonstrated[5].

They explained that the Court of Cassation took in 2019 a policy decision to develop all relevant tools in-house relying only on open source software and libraries, and to make them available online with all the relevant documentation.

[1] See https://www.numerique.gouv.fr/dinum/
[2] See https://github.com/Cour-de-cassation/moteurNER
[3] See https://github.com/Cour-de-cassation/label
[4] See https://huggingface.co/kriton/greek-text-summarization. With a more suitable training set (the one used consisted in newspaper articles) results may improve.
[5] It can be tested at https://huggingface.co/spaces/kriton/greek-title-generator

2. Romanian experience

Representative of Romanian judiciary (judge and Head of the Informatics and Judicial Statistics Unit of the Romanian Superior Council of Magistracy) explained that the publication of Romanian judicial decisions is carried out without AI support. Categorisation is carried out by clerks at the moment of receipt of a case, and the categories may be changed or added during the case processing by courts. Categorisation is used for determining the weight of a case before its allocation, aiming at a balanced workload among judges. It is also used for statistical reports and for searches in the web portal of judicial decisions.

3. Lithuanian experience

Representative of Lithuanian judiciary (judge assistant at the Supreme Court of Lithuania) described the categorisation of cases and judicial decisions at the Court information system, LITEKO. The categories are allocated to the cases by first instance courts from a closed list determined by the Council of Judges. Summaries of the most important cases are prepared manually each month by legal experts in the case law departments of the Supreme Court of Lithuania, Supreme Administrative Court and Court of Appeal of Lithuania.

4. Polish experience

Polish expert in the Modern Technologies Team of the Ministry of Justice of Poland, explained that the publication of selected decisions is a strategic choice of Poland aimed at ensuring open justice and transparency. Approximately 100.000 decisions of the Supreme Court are available online, in addition to ~400.000 of lower courts. Categorisation of cases is performed by clerks for statistical purposes, and judges, who select their decisions for publication, choose the relevant keywords. He stressed the importance of a systematic approach to innovation, learning from use cases, and feedback collection and analysis.

5. Greek experience

Two speakers representing Greece: Presiding Judge at First Instance Administrative Court of Thessaloniki, and Head of the Directorate of Digital Governance of the Ministry of Justice of Greece, explained that judges in Greece have access to digital archives of all judicial decisions. At the same time due to lack of resources for manual anonymisation only selected decisions are being published. AI tools could be used[6] to facilitate anonymisation and association the European Case Law Identifiers (ECLI)'s optional metadata "dcterms: abstract" (summary) and "dcterms: description" (keywords). The latter could be selected from a "Controlled Legal Vocabulary" to be developed for Greek language starting from existing legal ontology standards such as Atoma Nkoso[7] or VocBench[8]. Advanced search tool could allow to judges to have faster access to information on similar cases, which could help to take into account all relevant jurisprudence during the decision-making. A similar tool (or chatbot) could be made available to citizens, self-represented litigants and lawyers.

[6] Law 4961/2022 "On emerging information and communications technologies, strengthening digital governance and other provisions" provides the legal framework for utilising of AI in Greece, waiting for the future "AI Act" of the EU.
[7] See http://www.akomantoso.org/
[8] See https://vocbench.uniroma2.it/

6. Cypriot experience

Legal Officer of the Supreme Court of Cyprus explained that summarisation and categorisation of judicial decisions have so far been carried out only for the decisions of the Supreme Court published in the Cyprus Law Reports. There are expectations that with the support of TJENI project AI based tools could be tested from the point of view of their potential application for automated summarisation.

7. ECtHR experience

Lawyer of the Registry of the European Court for Human Rights (ECtHR) explained how the process of summarisation, which at the ECtHR is carried out manually, given also the relatively limited number of cases to be covered each year (around 5.000) can be faster and more efficient when decisions have the same structure. She described the structure of the ECtHR's judgments used since 2019, including the special box with keywords and short sentences on the main legal issues covered at the top of the text.

SESSION 2. Tests of automated summarisation – analysis and discussion of results

1. Methodology of preparation of the summary with AI tool

CoE expert presented the proposed concept for a methodology for testing summarisation tools for Greek language judicial decisions. He noted that the recent availability of tools based on Large Language Models (LLM, one kind of the so-called "Artificial Intelligence") for effective Natural Language Processing (NLP) constitutes an important enabling factor. Generative pre-trained models, such as the chat-GPT developed by OpenAI, when provided with a task ("prompt") proceed to complete it predicting based on the statistical model encoded in the model with billions of parameters enshrined in the nodes of the neural network, which has been trained on the basis of human-like results. Testing would aim at assessing effectiveness in reproducing human-like summaries of different AI-based tools.

CoE expert explained the methodology used to create with GPT-3.5 assistance the summaries used for this test. The text of the judgment was separated into parts and the relevant parts were separately uploaded to the model with relevant summary request (prompts). The answers obtained were used only if considered good enough, and in case they were not sufficiently clear a follow up question was prompted. The resulting text was edited and abridged but without adding any element which did not come out as an answer. The ECtHR decision was not read in advance, but the original summary (produced by human) was used as a model of the result at which the exercises was aimed at.

2. Testing methodology and results

CoE Project officer and legal analyst introduced the methodology to be used. The proposed testing methodology is based on the assessment (scoring) of the quality of the produced summaries and comparison of machine produced summaries with human produced summaries.

The text of two summaries of an ECtHR judgment[9] were distributed along with a scorecard to all participants. One of the two summaries was prepared by human (available on HUDOC), while the other was produced with the assistance of chat-GPT 3.5 (the methodology of its production is described below in section 4 below). The participants were asked to score each of these summaries, separately each of their parts (Topic/Labelling, Facts, Law, Conclusion and Remedies) along two dimensions (correctness and completeness) with a 3-steps scale (0: wrong information/incorrect language; 1: needs improvement; 2: correct information/clear language (understandable).

Following the compilation of the scorecards, the discussion was opened to compare the results. The two English summaries got quite similar scoring, and a slight majority of the participants thought that the summary created with chat-GPT was the one created by a human.

The scoring results are summarised in Figure 1 below (in green are highlighted the best scores) and presented in full in Annex 3.

Figure 1: Scoring of summaries in English and Greek

ENGLISH (N = 11)		Summary A (original)	Summary B (chat-GPT assisted)
Topic / labelling	correct	1,6	1,7
	clear	1,4	2,0
Facts	correct	1,7	1,4
	clear	1,7	1,5
Law	correct	1,7	2,0
	clear	1,3	2,0
Conclusion	correct	1,8	1,8
	clear	1,9	1,6
Remedies	correct	1,5	1,9
	clear	1,6	1,9
Average	overall	1,6	1,8
	correct	1,7	1,8
	clear	1,6	1,8

GREEK (N = 3)		Summary 1 (chat-GPT assisted)	Summary 2 (original)
Topic / labelling	correct	1,3	1,7
	clear	1,7	1,7
Facts	correct	1,3	2,0
	clear	0,7	2,0
Law	correct	1,7	1,7
	clear	1,7	1,7
Conclusion	correct	1,7	1,7
	clear	1,3	2,0
Remedies	correct	2,0	2,0
	clear	1,0	2,0
Average	overall	1,4	1,8
	correct	1,6	1,8
	clear	1,3	1,9

The second test results confirmed that chat-GPT provides much lower performances with the Greek language.[10] The scoring results are summarised in Figure 2 below (in green are highlighted the best scores) and presented in full in Annex 3. The discussion of the results of scoring and their comparison proved to be quite thought-provoking.

In the last session the participants were jointly suggesting how to formulate prompts[11] in order to improve the summarisation of a selected decision of the Cyprus Supreme Court and discussing the results or tested themselves summarisation of pre-loaded decisions both in

[9] Stavropoulos and others v. Greece, Application no. 52484/18.
[10] It has to be noted however that the scorings may be also reflect the fact that in this case participants already knew which was the original summary.
[11] For GPT 3.5 turbo, via https://platform.openai.com/playground?mode=chat.

English and Greek[12]. The dissatisfactory results of such summarisation in Greek language, even for short texts, were again confirmed.

3. Main conclusions

Language models can deal with a limited amount of text in each interaction. Text chunks (parts of words or groups of words) are converted into numbers ("tokens") for further manipulations. The user's prompt and the result produced by a model cannot exceed a certain number of tokens. The maximum token size accepted by the model in most cases is not enough for a summarisation of a whole judicial decision.

Through the mechanism of self-attention, the model has the ability to maintain a context of previous inputs (prompts), and this in its turn affects all further responds (results produced by the model).

Languages written in non-Latin scripts, such as Greek, are affected (and have limitations) by English-based tokenizers. Using the standard OpenAI tokenizer (based on UTF-8 encoding) for chat-GPT 3.5 allows conversion into one token of approximately 4 English characters, but only 1.19 Greek characters. This affects the size of the text that could be summarised and increases the costs, as OpenAI charges fees per token. Summarisation of Greek decision thus becomes almost 4 times more expensive than English ones. Using alternative tokenizers that are efficient on the Greek script can improve this.

The use of such tools requires time and effort by beginning users to become accustomed to the system and the process. The composition of the Greek summary, for instance, necessitated an investment of approximately 3 hours and a series of iterative prompts prior to its ultimate manual refinement by the tester.

As noted during the discussion, relying on commercial tools such as the ones produced by OpenAI, let alone the question of the costs, is not only risky from the point of view of business continuity but can also be non-compliant with the GDPR requirements (if, for example, such models are used for the summarisation or anonymisation of decisions containing personal data). Given their lack of transparency, it may not be excluded that they produced results can be manipulated or wilfully biased.

[12] Via https://beta.pickaxeproject.com/ connected to GPT 3.5 API.

Annex 1 – Summaries tested

TEST 1 – Summarisation of ECtHR judgement (English version)

SUMMARY A (original summary)

June 2020

Stavropoulos and Others v. Greece - 52484/18

Judgment 25.6.2020 [Section I]

Article 9

Article 9-1

Manifest religion or belief

Birth certificate revealing parents' choice not to christen their child: violation

Facts – The three applicants are parents and daughter. The daughter's forename was entered into the birth record with the handwritten addition, contained in brackets, of an abbreviation of the word "naming" (ονοματοδοσία). Moreover, the section concerning christening that is included in the birth registration act was left blank. The Supreme Administrative Court rejected the application for the annulment of the birth registration of the daughter as inadmissible.

Law – Article 9: The note "naming" next to the daughter's first name and the fact that the section concerning christening had been left blank implied that the parents had chosen not to have her christened. Such information appearing in a public document issued by the State constituted an interference with the right of all of the applicants not to be obliged to manifest their beliefs. Given the frequent use of the birth certificate, implying one's religious beliefs in it exposed the bearers to the risk of discriminatory situations in their dealings with administrative authorities.

This interference was not prescribed by law. Indeed, the relevant domestic-law provision provided that an individual acquired his or her first name by naming. Therefore, it did not follow, either from that law or from any other piece of domestic legislation brought to the attention of the Court, that registrars needed to write the word "naming" next to the first names of new-born children acquiring their names by the civil act of naming, as opposed to by christening. The interference in issue had resulted from the widespread practice of the registry offices.

Conclusion: violation (unanimously)

Article 41: EUR 10,000 jointly in respect of non-pecuniary damage.

SUMMARY B (AI assisted summary)

June 2020

Stavropoulos and Others v. Greece - 52484/18

Judgment 25.6.2020 [Section I]

Article 9

The use of the term "naming" next to a child's first name on a birth certificate and the omission of information regarding christening was deemed an interference with the right to not manifest one's beliefs.

Facts – The case concerns the registration of the third applicant's forename in the birth register. The applicants argued that the addition of the note "naming" next to their daughter's name constituted a violation of their right to freedom of religion. They claimed that the note revealed their religious beliefs as it suggested that their child had not been christened, which could lead to discrimination. The applicants believed that registry offices were not required by law to specify whether a child was named by a civil or Christian act, and argued that this violated their right to freedom of religion and revealed sensitive personal data, in breach of Article 8 of the Convention. The Supreme Administrative Court rejected the applicants' application for annulment as inadmissible.

Law – The Court decided that there had been an interference with the applicants' rights under Article 9 of the Convention, which protects freedom of thought, conscience, and religion. The interference was caused by the insertion of the word "naming" next to the third applicant's name on her birth registration act, which implied that she was not christened and that her name was given by the civil act of naming. This constituted an interference with the applicants' right not to be obliged to manifest their beliefs. The Court found that the interference was not prescribed by law, as there was no legal requirement for registrars to write the word "naming" next to the first names of newborn children acquiring their names by the civil act of naming, as opposed to by christening.

Conclusion: There was a violation of Article 9 of the Convention. The decision was unanimous.

Article 41: The Court decided that the respondent State is to pay the applicants jointly EUR 10,000 (ten thousand euros), plus any tax that may be chargeable, in respect of non-pecuniary damage.

TEST 2 – Summarisation of ECtHR judgement (Greek version)

ΠΕΡΙΛΗΨΗ A (AI assisted summary)

Ιούνιος 2020

Σταυρόπουλος και λοιποί κατά Ελλάδας - 52484/18

Απόφαση 25.6.2020 [Τμήμα Ι]

Άρθρο 9

Άρθρο 9-1

Εκδήλωση θρησκείας ή πεποίθησης

Παραβίαση ελευθερίας θρησκευτικής πεποίθησης λόγω σημείωσης στη ληξιαρχική πράξη γέννησης και κενής ενότητας για τη βάπτιση

Πραγματικά περιστατικά - Στην υπόθεση αυτή, οι προσφεύγοντες ζητούσαν την ακύρωση της ληξιαρχικής πράξης γέννησης της κόρης τους, καθώς είχε σημειωθεί η λέξη "ονοματοδοσία" δίπλα στο όνομα της κόρης τους. Οι προσφεύγοντες ισχυρίστηκαν ότι αυτό αποτελούσε αναφορά στο γεγονός ότι το παιδί τους δεν είχε βαπτιστεί, αποκαλύπτοντας έτσι τις θρησκευτικές πεποιθήσεις τους. Επίσης, ισχυρίστηκαν ότι ο νόμος δεν απαιτεί από τα ληξιαρχεία να προσδιορίζουν κατά πόσον το όνομα του παιδιού δόθηκε με πολιτική ή θρησκευτική πράξη, και ότι ο εν λόγω προσδιορισμός αποτελεί παραβίαση του δικαιώματος της ελευθερίας της θρησκείας. Ωστόσο, το Συμβούλιο της Επικρατείας απέρριψε την προσφυγή ως απαράδεκτη, καθώς οι προσφεύγοντες δεν διέθεταν έννομο συμφέρον στην υπόθεση.

Νομοθεσία - Το Δικαστήριο αποφάσισε ότι υπήρξε επέμβαση στα δικαιώματα των προσφευγόντων σύμφωνα με το άρθρο 9 της Ευρωπαϊκής Σύμβασης Δικαιωμάτων του Ανθρώπου, το οποίο προστατεύει την ελευθερία της συνείδησης και της θρησκευτικής πεποίθησης. Το Δικαστήριο χρησιμοποίησε ως κριτήρια την παρουσία της σημείωσης "ονοματοδοσία" δίπλα στο όνομα της τρίτης προσφεύγουσας στη ληξιαρχική πράξη γέννησης, την κενή ενότητα που αφορά τη βάπτιση στην ίδια πράξη γέννησης και τον κίνδυνο διακρίσεων στις σχέσεις του ατόμου με τις δημόσιες αρχές. Το Δικαστήριο επίσης επισήμανε ότι η επέμβαση στα δικαιώματα των προσφευγόντων δεν προβλεπόταν από το νόμο, δεν επιδιώκονταν ένας νόμιμος σκοπός και δεν ήταν απαραίτητη σε μια δημοκρατική κοινωνία.

Συμπέρασμα - Το Δικαστήριο αποφάσισε ότι υπήρξε παραβίαση του άρθρου 9 της Σύμβασης και χορήγησε δίκαιη ικανοποίηση στον παθόντα.

Άρθρο 41: - Το Δικαστήριο αποφάσισε ότι το εναγόμενο Κράτος πρέπει να καταβάλλει από κοινού στους προσφεύγοντες το ποσό των EUR 10.000 (δέκα χιλιάδες ευρώ), συν τους φόρους με τους οποίους μπορεί να επιβαρυνθούν οι προσφεύγοντες, για ηθική βλάβη.

ΠΕΡΙΛΗΨΗ Β (original summary)

Ιούνιος 2020

Σταυρόπουλος και λοιποί κατά Ελλάδας - 52484/18

Απόφαση 25.6.2020 [Τμήμα Ι]

Άρθρο 9

Άρθρο 9-1

Εκδήλωση θρησκείας ή πεποίθησης

Πιστοποιητικό γέννησης αποκαλύπτει την επιλογή των γονέων να μην βαπτίσουν το παιδί τους: παραβίαση

Πραγματικά περιστατικά - Οι τρεις προσφεύγοντες είναι γονείς και κόρη. Το όνομα της κόρης καταχωρήθηκε στο πιστοποιητικό γέννησης με τη χειρόγραφη προσθήκη, σε παρένθεση, μιας συντομογραφίας της λέξης "ονοματοδοσία". Επιπλέον, η ενότητα που αφορά τη βάπτιση και περιλαμβάνεται στην ληξιαρχική πράξη γέννησης έμεινε κενή. Το Συμβούλιο της Επικρατείας απέρριψε ως απαράδεκτη την προσφυγή για την ακύρωση της ληξιαρχικής πράξης γέννησης της κόρης.

Νομοθεσία - Άρθρο 9: Η σημείωση "ονοματοδοσία" δίπλα στο μικρό όνομα της κόρης και το γεγονός ότι η ενότητα που αφορά τη βάπτιση είχε αφεθεί κενή υπονοούσαν ότι οι γονείς είχαν επιλέξει να μην τη βαπτίσουν. Τέτοιου είδους πληροφορίες που εμφανίζονταν σε δημόσιο έγγραφο που είχε εκδοθεί από το κράτος συνιστούσαν επέμβαση στο δικαίωμα όλων των προσφευγόντων να μην υποχρεώνονται να εκδηλώνουν τις πεποιθήσεις τους. Δεδομένης της συχνής χρήσης του πιστοποιητικού γέννησης, η αναφορά των θρησκευτικών πεποιθήσεων σε αυτό εξέθετε τους κατόχους του στον κίνδυνο να υποστούν διακρίσεις στις συναλλαγές τους με τις διοικητικές αρχές.

Η επέμβαση αυτή δεν προβλεπόταν από το νόμο. Πράγματι, η σχετική διάταξη του εσωτερικού δικαίου προέβλεπε ότι ένα άτομο αποκτούσε το μικρό του όνομα με την ονοματοδοσία. Ως εκ τούτου, δεν προέκυπτε, ούτε από τον εν λόγω νόμο ούτε από άλλο εσωτερικό νόμο που τέθηκε υπ' όψιν του Δικαστηρίου, ότι οι ληξίαρχοι έπρεπε να γράφουν τη λέξη "ονοματοδοσία" δίπλα στα μικρά ονόματα των των νεογνών που αποκτούν τα ονόματά τους με τη ληξιαρχική πράξη της ονοματοδοσίας, και όχι με βάπτιση. Η επίμαχη επέμβαση είχε προκύψει από την ευρέως διαδεδομένη πρακτική των ληξιαρχείων.

Συμπέρασμα: παραβίαση (ομόφωνα)

Άρθρο 41: 10.000 ευρώ από κοινού για ηθική βλάβη.

Annex 2 – Summaries' scorecard

2 : correct information / clear language (understandable)
1 : need improvement (please add a brief comment on what improvement is needed)
0 : wrong information / unclear language

	Summary A	Summary B
Topic/ labelling	0 1 2 - correct 0 1 2 - clear	0 1 2 - correct 0 1 2 - clear
Facts	0 1 2 - correct 0 1 2 - clear	0 1 2 - correct 0 1 2 - clear
Law (legal analysis)	0 1 2 - correct 0 1 2 - clear	0 1 2 - correct 0 1 2 - clear
Conclusion	0 1 2 - correct 0 1 2 - clear	0 1 2 - correct 0 1 2 - clear
Remedies	0 1 2 - correct 0 1 2 - clear	0 1 2 - correct 0 1 2 - clear