



THE UNIVERSITY of EDINBURGH
informatics



| epcc |

EMNLP
2023



AsyLex: A Dataset for Legal Language Processing of Refugee Claims

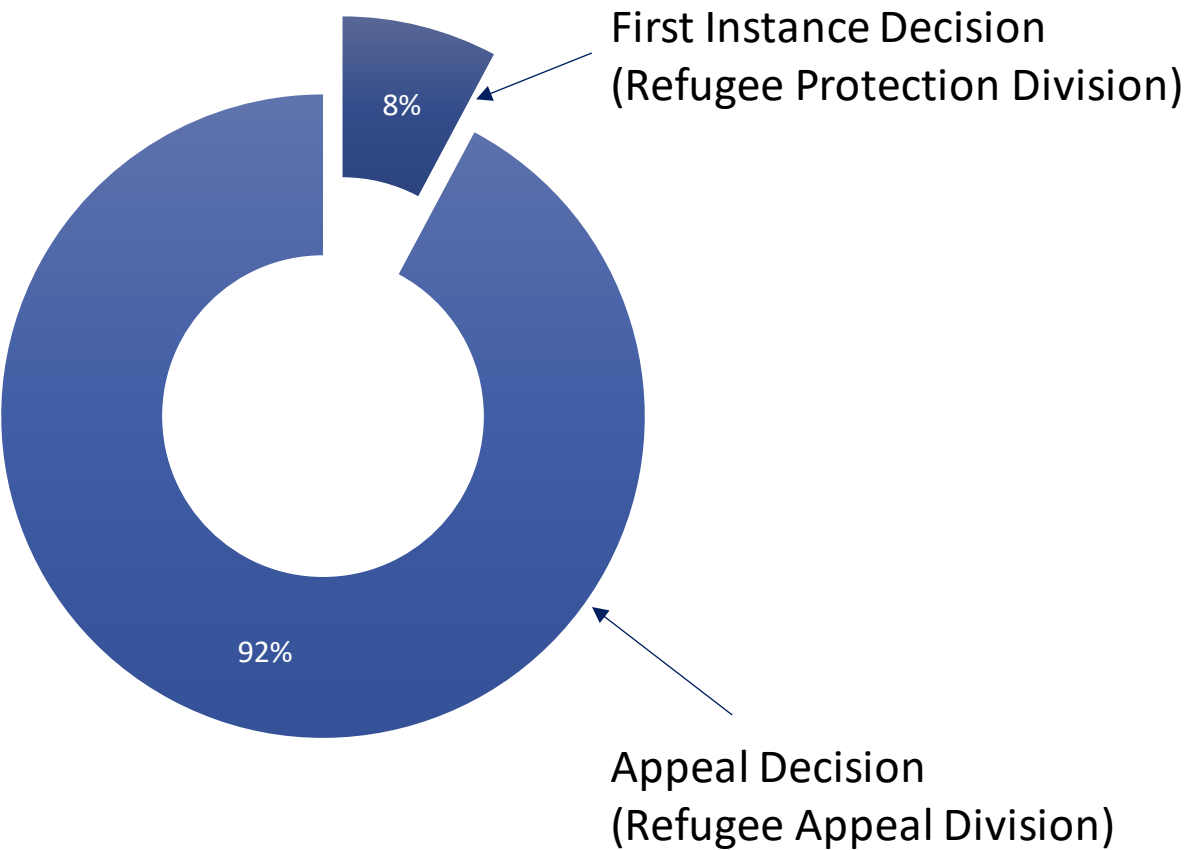
Claire Barale, Mark Klaisoongnoen, Pasquale Minervini, Michael Rovatsos,
Nehal Bhuta

School of Informatics, EPCC, School of Law
The University of Edinburgh

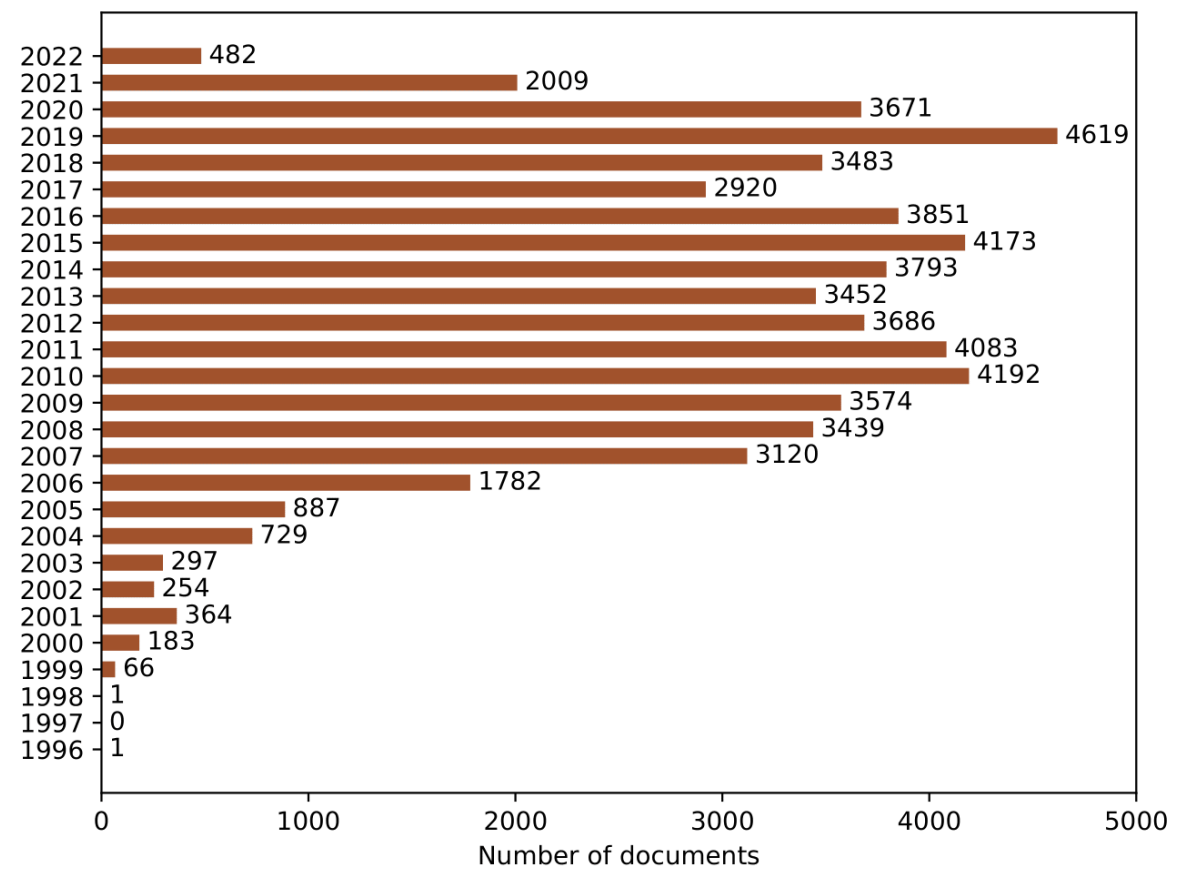
(1) THE FIRST DATASET FOR REFUGEE LAW



59,112 Refugee Status Determination Cases
from the Immigration and Refugee Board of Canada (IRB)



Distribution of cases per year



(2) KEY FEATURES OF ASYLEX



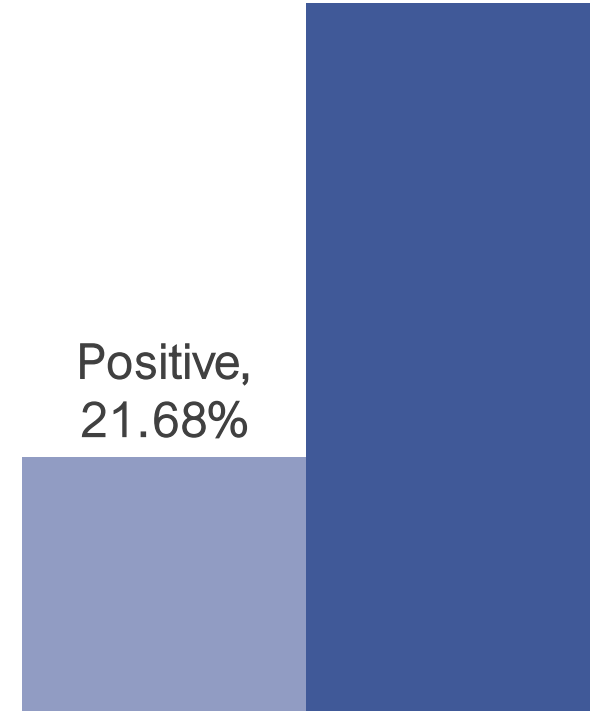
THE UNIVERSITY
of EDINBURGH

	Main text	Case cover	Case outcome
Documents	59,112	45,882	32,627
Sentences	4,946,438	-	53,977
Paragraphs	1,781,240	-	-
Labels	16	8	3
Labeled (human)	16,628	2,487	2,360
Labeled (rule-based)	-	34,001	-
Labeled (inferred)	6,154,226	123,802	31,227

Outcome of the cases

Negative,
60.18%

Positive,
21.68%



20 legally relevant entity types extracted from each case
+ documents labeled with the decision outcome

(2) KEY FEATURES OF ASYLEX



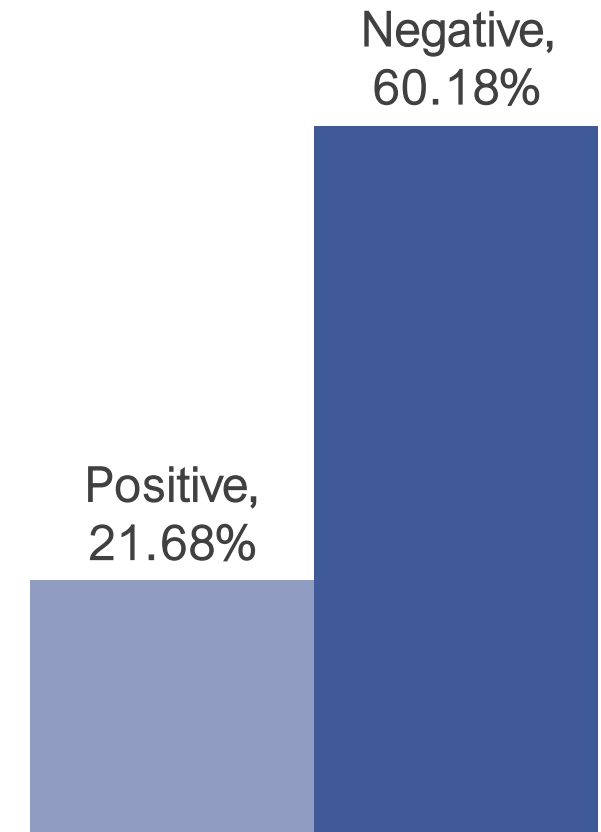
THE UNIVERSITY
of EDINBURGH

	Main text	Case cover	Case outcome
Documents	59,112	45,882	32,627
Sentences	4,946,438	-	53,977
Paragraphs	1,781,240	-	-
Labels	16	8	3
Labeled (human)	16,628	2,487	2,360
Labeled (rule-based)	-	34,001	-
Labeled (inferred)	6,154,226	123,802	31,227

19,115 gold standard annotations

+ 2,360 documents labeled with their case outcome

Outcome of the cases



(2) KEY FEATURES OF ASYLEX



THE UNIVERSITY
of EDINBURGH

	Main text	Case cover	Case outcome
Documents	59,112	45,882	32,627
Sentences	4,946,438	-	53,977
Paragraphs	1,781,240	-	-
Labels	16	8	3
Labeled (human)	16,628	2,487	2,360
Labeled (rule-based)	-	34,001	-
Labeled (inferred)	6,154,226	123,802	31,227

Outcome of the cases

Negative,
60.18%

Positive,
21.68%

57,408 rule-based extracted annotations

(2) KEY FEATURES OF ASYLEX

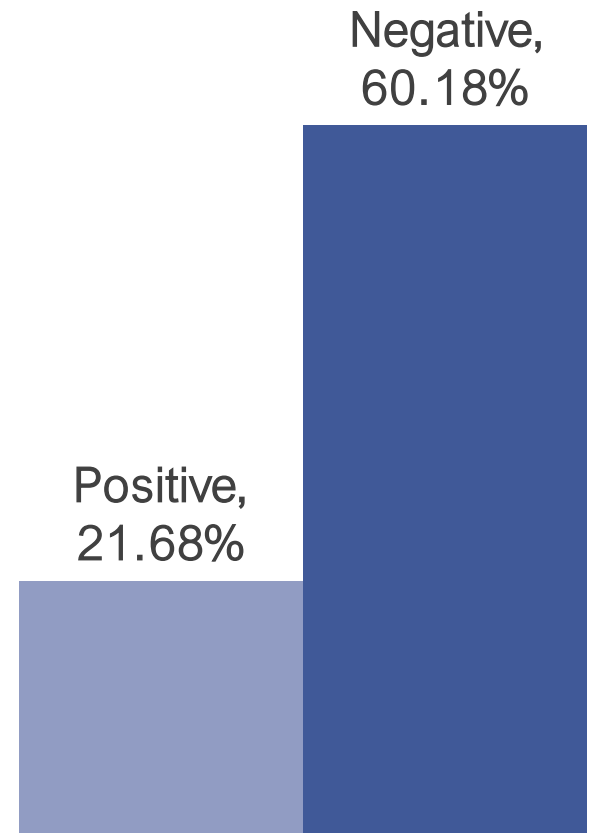


THE UNIVERSITY
of EDINBURGH

	Main text	Case cover	Case outcome
Documents	59,112	45,882	32,627
Sentences	4,946,438	-	53,977
Paragraphs	1,781,240	-	-
Labels	16	8	3
Labeled (human)	16,628	2,487	2,360
Labeled (rule-based)	-	34,001	-
Labeled (inferred)	6,154,226	123,802	31,227


Fine-tuned model for legal entity extraction,
based on LegalBERT
(fine-tuned models available on HuggingFace)

Outcome of the cases



(3) EXTRACTED ENTITIES

Immigration and
Refugee Board of Canada
Refugee Protection Division



Commission de l'immigration et
du statut de réfugié du Canada
Section de la protection des réfugiés

RPD File No. / N° de dossier de la SPR : MB2-06059

Private Proceeding / Huis clos

Reasons and decision – Motifs et décision

Claimant(s)

XXXX XXXX XXXX

Demandeur(e)(s) d'asile

Date(s) of hearing

October 28, 2013

Date(s) de l'audience

Place of hearing

Montréal, Quebec

Lieu de l'audience

Date of decision
and reasons

December 16, 2013

Date de la décision
et des motifs

Panel

Stéphane Morin

Tribunal

Counsel for the claimant(s)

M^{re} Marie-José Blain

Conseil(s) du (de la/des)
demandeur(e)(s) d'asile

Designated representative

N/A

Représentant(e) désigné(e)

Counsel for the Minister

N/A

Conseil du (de la) ministre

2013 CanLII 92087 (CA IRB)

CONCLUSION

DETERMINATION

[23] The appeal **is allowed**. I substitute my own decision that **the Appellant is a Convention**
refugee.

DETERMINATION

(signed)

JUDGE

DATE

December 17, 2020

Date

DATE OF THE HEARING	CLAIMANT EVENT	CLAIMANT INFO
DATES	LOCATIONS	ADJECTIVE LOCATIONS
DATE DECISION	PROCEDURE	CREDIBILITY
PUBLIC PRIVATE HARING	DOC EVIDENCE	EXPLANATION
IN CHAMBER VIRTUAL	LEGAL CITATIONS	CASE PRECEDENT
JUDGE NAME	COUNTRY REPORTS	LEGAL GROUND
TRIBUNAL		

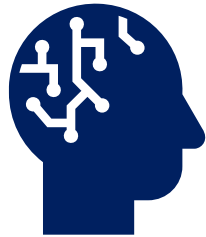
(4) WHAT CAN YOU DO WITH ASYLEX?



THE UNIVERSITY
of EDINBURGH

2 tasks presented in our paper:

1. Legal Entity Extraction
2. Judgement Outcome Classification



For the NLP Community

- Legal Entity Extraction (NER models)
- Unstructured → Structured Dataset
- Text Classification
- Case and Bias Analysis

Access to **the raw text** as well as the **structured data**



For legal practitioners

- Case Review and enhanced legal search
- Unstructured → Structured Dataset
- Past Cases retrieval
- Insights for new applications

Access to the raw text as well as the structured data



THE UNIVERSITY
of EDINBURGH



Data available on HuggingFace

THANK YOU!

AsyLex: A Dataset for Refugee Claims

<p>IAD File No. / N° de dossier de la SAI : MB0-03574 Client ID No. / N° ID client : 5276-3336</p> <p>3</p> <p>Background</p> <p>[5] The appellant was born in Colombia and is 26 years old. He became a permanent resident on August 6, 2004, at the same time as his mother, as a dependent included in his mother's immigration application as a member of the entrepreneur class. The appellant's father was also included in the application for permanent residence filed by the appellant's mother, and he became a permanent resident in October 2004.</p> <p>[6] On September 20, 2007, an immigration officer assessed the file of the appellant and his parents in order to decide whether they had met the conditions of the entrepreneur class regarding their permanent residence.</p> <p>[7] After an analysis, he recommended that an admissibility hearing be held before the ID because he was of the opinion that the conditions set out in section 98 of the IRPR had not been met. The appellant, as well as his parents, were notified to appear at a hearing before the ID.</p> <p>[8] At the hearing on May 11, 2010, only the appellant appeared; his parents had returned to Colombia. He acknowledged that he had not met the conditions, and a removal order was issued against him that same day.</p> <p>[9] On June 10, 2010, the panel received the notice of appeal of the removal order.</p> <p>Analysis</p> <p>[10] The panel is guided in its decision by the factors set out and established by the case law,³ given the adaptations required in this case. The panel is of the opinion that these factors, although not exhaustive, are useful in deciding whether to exercise its discretion to grant special relief. These factors are the following:</p> <p><small>³ <i>Ribic, Marilda v. M.E.I.</i> (IAB 84-9623), D. Davey, Benedetti, Petryshyn, August 20, 1985 <i>Chien, Hsue v. Canada</i> (M.C.I.), 2002 SCC 3 <i>Al Saghan, Ahmad Abdulaziz v. Canada</i> (M.C.I.), 2002 SCC 4.</small></p>	<p>IAD File No. / N° de dossier de la SAI : MB0-03574 Client ID No. / N° ID client : 5276-3336</p> <p>4</p> <ul style="list-style-type: none">• the seriousness of the offence leading to the removal order;• the possibility of rehabilitation;• the length of time the appellant has spent in Canada and the degree to which he or she is established here;• the likelihood that would be caused to the appellant's family members in this country if the appellant was removed;• the family and community support available to the appellant; and• the degree of hardship that could be caused to the appellant by his return to his country of nationality. <p>[11] The weight to be given to each of these factors varies from one appeal to the next, depending of the circumstances of the case.</p> <p>[12] The appellant testified at the hearing. His brother, a Canadian citizen, was also present and testified. His mother testified by telephone. Having heard all of the testimony, the panel has few doubts as to credibility of the appellant and his brother. Their testimony was frank, direct and spontaneous, aside from the fact that they had trouble remembering some dates. Moreover, the panel did not note any implausibility or major contradiction in their testimony. As for the appellant's mother, the panel found her testimony vague, and on several occasions, she went far beyond the scope of the questions asked, and the information provided was less relevant. The panel also noted that in doing this, she avoided giving a direct response to the question asked, and the questions had to be repeated to her many times, making her testimony less reliable than that of her sons.</p> <p>[13] The Minister's counsel submitted that the appellant's parents used the entrepreneur immigration program in order to enable their son, the appellant, to immigrate to Canada with them and to give him the benefit of an education at a recognized university at a lower cost. He submitted that that was what they intended when they filed their application and that they never really intended on settling in Canada, given the little time that they spent here and the fact that the appellant's father had his own company in Colombia.</p>	<p>IAD File No. / N° de dossier de la SAI : MB0-03574 Client ID No. / N° ID client : 5276-3336</p> <p>5</p> <p>[14] He added that they clearly knew the immigration system because their eldest son had already studied in Canada as a foreign student and they had to pay almost twice as much in tuition as residents did. The Minister's counsel submitted that since the appellant's humanitarian and compassionate considerations stem from the parents' somewhat questionable intentions, the appellant should not be able to use this as a reason for retaining his permanent residence.</p> <p>[15] He supported that position by submitting that for the sake of the integrity of the immigration program, this appeal should not be allowed, because it would send a bad message to the public—namely, that it is easy to study in Canada for cheap by abusing immigration programs.</p> <p>[16] The Minister's counsel did not deny that there are humanitarian and compassionate considerations, and he acknowledged that the appellant has a significant degree of establishment here. However, he was not of the opinion that they are sufficient.</p> <p>[17] The panel does not agree at all with the point of view of the Minister's counsel. Although it may be possible and even probable that some individuals use or have used immigration programs primarily to guarantee a better future for their children and to offer them an education at a low cost, the panel is far from satisfied that the evidence shows that it was planned in this case. Some evidence—namely, the testimony from the appellant's mother and the behaviour of the appellant's parents after obtaining permanent residence—certainly raises questions regarding the appellant's parents' willingness to settle long term in Canada, given the difficulties encountered. However, the conclusion of the Minister's counsel that the appellant's parents were involved in a complex immigration process and that they planned their actions in a timely manner merely to save on tuition fees—which would be equivalent to approximately \$10,000—is very tenuous, in the panel's opinion.</p> <p>[18] This link seems even weaker given that the appellant's father is an engineer with a university degree and a master's degree in Colombia and that his mother is also a university graduate. These are important and determinative factors⁴ in an immigration application as a member of the skilled worker class. It is more than likely that the appellant's parents could have obtained permanent residence that</p>
---	---	--

Structure via an Information Extraction Pipeline

Decision_ID	date_decision	date_hearing	Tribunal (1st instance/app eal)	Judge	Claimant sequence of events	Gender	Age	Citizenship	Dependant persons or not	One/multiple applicants
Credibility (binary)	Timeline of dates	Hard evidences provided	Explanation given by judge	Geographical indications	Law citations	Case citations	Report citations	Organization cited	Names cited	Legal procedure events