# **Annotation Study of Japanese Judgments on Tort for** Legal Judgment Prediction with Rationales

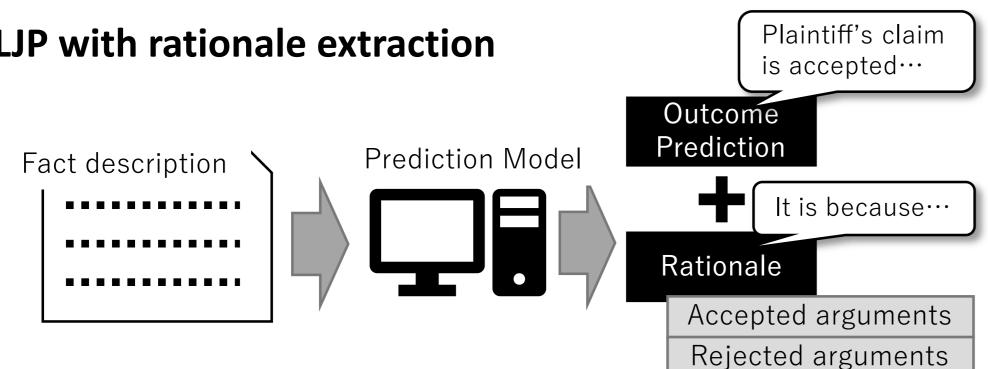
Hiroaki Yamada<sup>1</sup> Takenobu Tokuanaga<sup>1</sup> Ryutaro Ohara<sup>2</sup> Keisuke Takeshita<sup>3</sup> Mihoko Sumida<sup>3</sup> <sup>1</sup>Tokyo Institute of Technology, <sup>2</sup>Nakamura, Tsunoda & Matsumoto, <sup>3</sup>Hitotsubashi University

# **Our contributions**

- Annotation for Legal Judgment Prediction (LJP) with rationales.
- Extended annotation studies on torts by legal experts.
- The first and reliable annotation scheme for the Japanese judgment document.

# **1. Legal Judgment Prediction in Japanese**

Framework of LJP with rationale extraction



# **3. Annotation example**

第1 請求 被告は,原告に対し,別紙発信者情報目録記載の情報を開示せよ。	Annotations on "Span assoc	Annotations on "Span associating"				
第2 事案の概要 本件は、原告が、被告に対し、インターネット上の電子掲示板に投稿された別紙投稿記事目録記載の書込みにより社会的評価を低下させられ名誉 権を侵害されたと主張して、特定電気通信役務提供者の損害賠償責任の制限及び発信者情報の開示に関する法律(以下「法」という。)4条1項に基づ き、別紙発信者情報目録記載の情報の開示を求める事案である。 ① 令和元年7月12日午後11時46分17秒、インターネット上の電子掲示板「C」中の「D」「E」に作成された「F」という標題のスレ	本件投稿の対象が原 告であるとはいえない ことから、本件投稿が 原告の社会的評価を低 下させて原告の名誉を 毀損するものであると いうことはできない。	▲件投稿は、一般の × 閲覧者の普通の注意と 閲覧の仕方を基準とす ると、B製作所に勤務 する「X1」という人 物が、特定の個人から 金銭を借り入れたがそ				
ッド(以下「本件スレッド」という。)に, 「X1さん金返さないと」という書込み(別紙投稿記事目録記載のもの。以下「本件投稿」という。)が, IPアドレス「ooo. ooo. ooo. oo」を経由して投稿された。		の返済をしていないと の事実を摘示するもの である。				
<ul> <li>2 争点及び当事者の主張</li> <li>本件の争点は、「本件投稿について権利侵害の明白性が認められるか。」であり、これに関する当事者の主張は以下のとおりである。</li> <li>(1)原告の主張</li> <li>ア 同定可能性</li> <li>2 本件投稿は、一般の閲覧者の普通の注意と閲覧の仕方を基準とすると、B製作所に勤務する「X1」という人物が、特定の個人から金銭を借</li> <li>り入れたがその返済をしていないとの事実を摘示するものである。</li> </ul>	本件投稿の対象が原 告であるとはいえない ことから、本件投稿が 原告の社会的評価を低 下させて原告の名誉を 毀損するものであると いうことはできない。	■ B製作所に勤務する × 「X1」という姓の人 物は,原告とそのいと この2名のみである。				
<ul> <li>B製作所に勤務する「X1」という姓の人物は、原告とそのいとこの2名のみである。</li> <li>本件投稿の閲覧者のうち、原告を知っているが原告のいとこを知らない者は本件投稿の対象が原告と考えるであろうし、原告と原告のいとこを知る者が「X1」という記載から原告のことを思い浮かべることもあるはずである。</li> <li>したがって、本件投稿の対象と原告との間に同定可能性はある。</li> <li>(2) 被告の主張</li> <li>上記(1)原告の主張は、いずれも争う。</li> </ul>	本件投稿の対象が原 告であるとはいえない ことから、本件投稿が 原告の社会的評価を低 下させて原告の名誉を 毀損するものであると いうことはできない。	▲件投稿の閲覧者の × うち, 原告を知ってい るが原告のいとこを知 らない者は本件投稿の 対象が原告と考えるで あろう				
	•					

## Aim and Scope

- ML-based LJP is an emerging topic (Katz et al., 2017; Xiao et al., 2018; Chalkidis et al., 2021;)
- However, there is no Japanese dataset for LJP; even no annotation study so far.
- $\rightarrow$  Towards constructing large-scale dataset for the Japanese LJP, we develop an annotation scheme.

# **Target documents**

- The first instances of the Civil Code cases in Japan.
  - Especially, **Torts** from disclosure of identification information of the sender cases, defamation cases, privacy infringement cases

# **2.** Our annotation scheme

### **Document screening**

- A simple task to filter out non-tort judgments.  $\bullet$
- We ask annotators to read through the judicial decision section in a judgment  $\bullet$ and check if the court considers an issue of torts and makes any decisions on it. If annotators find a judgment has nothing to do with torts, they are instructed to flag the judgment and stop annotating it.

- 第3 当裁判所の判断
- 1 争点(権利侵害の明白性)について

前記前提事実,証拠(甲4,5)及び弁論の全趣旨によれば,本件投稿が,本件スレッド(インターネットサイト「C」中の「D」「E」上に 「X 1 さん金返さないと」という内容のものであること、原告が岡山県倉敷市所在のB製作所に勤務 作成されたIF」と 勤務する「X1」という姓の人物は原告とそのいとこの2名であることか認められる

B製作所に「X1」という姓の人物が原告のほかにもう1名いることを知っている閲覧者は,本件投稿の「X1さ ん」が原告とそのいとこのいずれを指すのかを判別できず,また,B製作所に「X1」という姓の人物が原告のほかに何名いるのかを知らない閲覧者。 とするとそのような誤信をするであろうとはいえないから、これをもって一般の閲覧者の



このように、本件投稿の対象が原告であるとはいえないことから、本件投稿が原告の社会的評価を低下させて原告の名誉を毀損するもので ということはできない。 8

D Text (English versions are our translation.)	Span type	$ (\omega) \vee \nabla$	@D	@AC	Associated CD Spans
令和元年7月12日午後11時46分17秒,インターネット上の電子掲示板「C」中の 「D」「E」に作成された「F」という標題のスレッド(以下「本件スレッド」という。)に, 「X1さん金返さないと」という書込み(別紙投稿記事目録記載のもの。以下「本件投稿」と いう。)が,ⅠPアドレス「○○○.○○○.○○○.○○○」を経由して投稿された。	UF				8
At 11:46:17 PM, July 12, 2019, a posting "Mr X1, you should pay back the money (the attached list of submitted articles)" was made in the thread titled "F", which was created in "D" and "E" on the bulletin board system "C" on the Internet, via IP address ***.***.***.					
本件投稿は,一般の閲覧者の普通の注意と閲覧の仕方を基準とすると,B製作所に勤務する 「X1」という人物が,特定の個人から金銭を借り入れたがその返済をしていないとの事実を 摘示するものである。	- FC	PI.		т	8
This posting, based on a viewer of ordinary prudence and his way of viewing , indicate the fact that a person named "X1," who works at factory B, borrowed money from a certain individual but has not repaid it.				·	•
3 B製作所に勤務する「X1」という姓の人物は、原告とそのいとこの2名のみである。	- FC	PI.		Т	8
<ul> <li>There are only two persons with the surname "X1" who work at factory B: the plaintiff and his cousin.</li> <li>本件投稿の閲覧者のうち,原告を知っているが原告のいとこを知らない者は本件投稿の対象が</li> </ul>					
4 原告と考えるであろう				F	0
The viewers of this posting, who know the plaintiff but do not know the plaintiff's cousin, would regard the plaintiff as the subject of the posting.	- FC	Pl.		F	8
原告と原告のいとこを知る者が「X1」という記載から原告のことを思い浮かべることもある					
う Some of people, who knows the plaintiff and the plaintiff's cousin, can recall the plaintiff from the mention of "X1".	- FC	PI.		F	8
6本件投稿の対象と原告との間に同定可能性はある。	FC	וח		Г	8
It is possible to identify the subject of this posting as the plaintiff.		PI.		Г	0
∂ 上記(1)原告の主張は,いずれも争う。	- FC	Def.		т	8
We dispute all of the above (1) allegations of the plaintiff.					•
本件投稿の対象が原告であるとはいえないことから,本件投稿が原告の社会的評価を低下させ て原告の名誉を毀損するものであるということはできない。					
Given that the subject of this posting is not really the plaintiff, we cannot say that the posting is defamatory to the plaintiff by diminishing the plaintiff's social reputation.			F		

### **Span Extraction**

- Annotators are asked to identify the text span describing the court's conclusion on torts from the judicial decision section. And then, they also extract rationales from the sections of the parties' claims and facts.
- Steps:
  - **1. Identifying the court's conclusions and rationales** as text spans.
  - 2. Assigning attributes
    - @Accepted Claim: If the claim is accepted by judges or not?
    - @Who: Whose claim?
    - @Decision: If the torts are affirmed by judges or not? •

#### Court Decisions (CD)

- CD describes a judges' decision on tort. Annotators must find Court Decisions spans from the section of the judicial decision. One span identifies one tort.
- Attributes: @Decision

#### Factual Claims (FC)

- FC describes important claims from parties, which are relevant to judgment on torts. Annotators must find Factual Claims spans from the sections other than the judicial decision section. The Factual Claims contain factual allegations, an assertion of opposing facts against them, and rebuttals against one's factual allegations.
- Attributes: @Accepted Claim, @Who

# 4. Annotation study results

#### Annotators

• Study 1: law school graduates and lawyers

0.549

0.647

- We revised and improved annotation guidelines and scheme through Study 1 to 3.
- Study 3: lawyers / law school graduate / undergraduate

• Study 2: undergraduates in law and law school students

Statistics		# of Spans (macro-avg. over rators)						misc			
JUALISTICS		FC	NC	MC	UF	CD	Total	Rators	Docs	Avg. chars	
	Study 1	88.2	2.7	14.7	35.2	39.5	180.2	6	10	13507.9	
	Study 2	44.0	0	8.0	16.6	15.2	83.8	5	5	14890.6	
	Study 3	252.4	6.8	N/A	70.2	48.0	377.4	5	25	10763.6	

#### **Reliability metrics**

- *Fleiss's* κ (Fleiss, 1971) and *agreement ratio* for document screening task
- *Krippendorff's*  $\alpha_{II}$  (Krippendorff, 1995) for span extraction tasks and associating

#### **Claims of Norms (NC)**

• NC describes abstract legal arguments regarding torts. Annotators must find NC spans from the sections other than the judicial decision section. This type of span often consists of references to past precedents, in particular the Japanese Supreme Court judgments. • Attributes: @Accepted Claim, @Who

#### **Jndisputed Facts (UF)**

• UF spans describes facts that play important roles in judging torts. The facts covered with these spans are undisputed by any parties. The annotators find the spans from the sections other than the judicial decision section in principle

• Attributes: N/A

### Associating spans

Annotators are instructed to associate all Factual Claims, Major Claims, Claims or Norms, and Undisputed Facts spans with their corresponding CD spans.

#### spans task

Study 2

Study 3

Chan avtract		Document scree	ening	<b>Document Screening</b>				
Target Spans: Attribute types:	FC, NC, MC @Accepted Claim	S FC, NC, MC @Who	CD @Decision	expectedly shows high agreement as the task simpler and		Study 1	Agreement Ratio 0.83	Fleiss Kappa -0.0
Study 1 Study 1 (5 docs)	0.521 0.563	0.526 0.587	0.629 0.428	<sup>9</sup> easier.	S	Study 2 Study 3	1.00 0.96	N 0.
Study 2 Study 3	0.605 0.629	0.516 0.641	0.438 0.608		spans		gh $lpha_U$ of spantion task wa	
The studies achieved good agreement, particularly for the span extraction and the attributes task, suggesting that our annotation scheme and training materials were successful. Span extraction – Span types				Study 1 Study 1 (5 docs) Study 2 Study 3	$\begin{array}{c} \alpha_U \\ 0.301 \\ 0.209 \\ 0.321 \\ 0.430 \end{array}$	rs, ugh ests vorked		
Span extract	tion – Span typ							
Span extract	Factual Cla		of Norms	Undisputed Facts	Court I	Decisions	Overall	

()

0.069

0.337

0.520

0.457

0.607

0.498

0.654

#### Acknowledgement

We appreciate Prof. Souichirou Kozuka at Gakushuin University and Prof. Kazuhiko Yamamoto at Hitotsubashi University for their helpful comments. The judgment documents data for this study was provided by LIC Co., Ltd. This work was supported by JST RISTEX Grant Number JPMJRX19H3 and JST ACT-X Grant Number JPMJAX20AM.