Repairing Trust in Domain Name Disputes Practices: Insights from a Quarter-Century's Worth of Squabbles

Vinny Adjibi, Athanasios Avgetidis, Manos Antonakakis, Alberto Dainotti, Michael Bailey and Fabian Monrose
Georgia Institute of Technology
{vinny.adjibi, avgetidis, manos, dainotti, mbailey, fmonrose}@gatech.edu

***Abstract*—The Uniform Domain Name Dispute Resolution Policy (UDRP) seeks to balance two competing goals: empowering trademark holders to swiftly address cybersquatting cases targeting their brands and protecting domain registrants from aggressive legal tactics by overreaching trademark holders. Since its inception, the UDRP has become the de facto dispute resolution mechanism for over two thousand domain extensions, a substantial increase from the original three. However, despite its successes, critics argue that the policy enables practices that undermine trust and fairness. Unfortunately, meaningful reform efforts have stalled due to the absence of large-scale structured data, limiting empirical evaluations and leaving foundational questions unanswered for more than two decades.**

To address this long-standing gap, we trained models to extract structured data from 90,153 UDRP dispute proceedings, enabling the most comprehensive empirical analysis of the policy to date. Our findings shed light on several issues, showing evidence of forum shopping in almost one-third of all the disputes, potential conflicts of interest in 43 cases, and delays (by many parties) that fall well outside the expected response times—all of which impact the perceived fairness and efficiency of UDRP. Beyond eroding trust, those issues create serious security challenges: 2,751 malicious domains remained under malicious actors' control for up to four months after a panel ordered their transfer. Overall, our findings underscore the need for policy reform to help restore trust and improve transparency in the Internet's de facto standard for countering trademark infringement. Based on our discoveries, we recommend introducing greater automation, strengthening oversight, and enforcing clearer compliance rules to ensure that the UDRP remains a reliable tool for trademark-based name disputes—especially as the Internet continues to expand with new generic top-level domains and the digital environment becomes increasingly hostile to users.

I. Introduction

Domain names are the primary way to identify and interact with online services, making them essential to Internet operation. This attracts malicious actors who take advantage of the inexpensive and unvetted registration processes to buy domain names that support various malicious campaigns targeting individual users and reputable businesses alike. To combat those threats, the security community often relies on technical solutions such as automated detection, threat intelligence feeds, and domain reputation systems, which, despite their usefulness [1], are neither comprehensive nor bulletproof [2]. Those tools are particularly ineffective against highly deceptive forms of abuse, like spear phishing and counterfeiting, where human judgment is often required to establish wrongdoing.

To address those complex types of domain name abuse, companies often rely on the Uniform Domain Name Dispute Resolution Policy (UDRP), which offers a *faster, less expensive, and more accessible* process than traditional litigation and coordinated takedown efforts by law enforcement [3]. Over the past 25 years, trademark holders successfully utilized the UDRP's streamlined dispute process to contest the registration of over 165,681 domains, of which only 12,253 were considered malicious by at least one VirusTotal engine before a dispute was filed about them. At the same time, domain registrants successfully defended their rights to 9,991 contested domain names and preserved their ownership rights. Because of this zero-sum nature of the disputes, preserving trust in the policy is essential to its proper functioning.

Meanwhile, various stakeholders have repeatedly raised concerns about factors that undermine trust in the policy. These issues generally fall under three key areas: *fairness, effectiveness, and efficiency* [4], [5]. Concerns over fairness focus on complainants' presumed tendency to choose the provider most likely to rule in their favor (i.e., *forum shopping*) and claims of *panelist bias* toward trademark holders. Studies suggest that some providers gain market share by ruling more often in favor of complainants, raising doubts about impartiality [6]. In terms of effectiveness, inconsistent interpretations of UDRP criteria have led to *unpredictable outcomes*, making it harder for trademark holders to trust the process as identical fact patterns often lead to conflicting rulings [7]. Efficiency concerns center on delays: although ICANN suggests that most cases are resolved within 55 days [8], recent research shows that nearly half exceed this timeframe [9], allowing abuse to persist

Network and Distributed System Security (NDSS) Symposium 2026
23 - 27 February 2026 , San Diego, CA, USA
ISBN 979-8-9919276-8-0
https://dx.doi.org/10.14722/ndss.2026.230174
www.ndss-symposium.org

longer than expected. These issues have been repeatedly raised in public comments, legal scholarship, and policy reviews [8], [10]–[14], yet ICANN (and by extension, its supporting organizations) has not enacted meaningful reforms. With the upcoming release of new generic top-level domains in 2026, the need for data-driven policy analyses is clear.

One reason why the Internet's governing body for the domain name system may not have addressed longstanding concerns about the UDRP is the lack of consensus among its stakeholders. Opinions often diverge on both the seriousness of the issues and the appropriateness of proposed solutions. Much of this disagreement stems from reliance on *anecdotal* rather than *empirical* evidence, partly because the technology (e.g., transformer architectures) needed to reliably extract data from free-form text in UDRP proceedings has only recently become available. Even when data-driven studies have been conducted on limited data, they have produced conflicting results. For example, while some researchers argued that forum shopping was widespread [6], [12], others contended this was not so [15]. Disagreements over UDRP persist to this day, raising doubts on whether ICANN can meaningfully revise the policy to satisfy all stakeholders without an "evidence-based recommendation [16]" strategy. Our work aims to support that effort. Specifically, drawing on data from 90,153 disputes decided between December 1999 and August 2024, we empirically assess how the behavior of various stakeholders may have affected the trustworthiness of the UDRP and propose steps for remediation. In doing so, we make the following contributions:

1) A data curation methodology based on Natural Language Processing (NLP) techniques for accurately extracting pertinent entity information from PDF or HTML files of domain name dispute proceedings in English.
2) A conservative technique for identifying how quickly successfully contested domain names are transferred to complainants, finding that, on average, registrars implemented decisions within 28 business days post-decision. However, we also found that slow transfers allowed malicious activity to continue on 2,791 domains for up to four months, perpetuating real security threats.
3) An in-depth assessment of the prevalence of forum shopping, showing that up to a third of disputes might have been affected, which could make the policy appear biased towards trademark holders, since they are afforded the freedom of selecting which provider handles a dispute.
4) Extensive measurements of the efficiency of the various stages of domain name disputes, finding that while most parties executed their responsibilities within the expected timeframe, some often took much longer than expected. Those delays incurred by providers, panelists, or domain name registrars lead to unexpectedly long disputes, undermining trust and potentially extending abuse.
5) A systematic analysis of potential panelists' conflicts of interest. While events leading to a perception of bias are rare, there is a need to establish policies that clearly define what constitutes a conflict of interest.

Overall, our longitudinal analysis suggests that although the policy has been broadly successful in meeting some of its goals, providers, panelists, complainants, and registrars have all exhibited behaviors that could cast doubts about the perception of the policy as a fast and fair mechanism to settle domain name abuse. We hope our discoveries and recommendations are viewed as useful contributions to the policy debate and spur discussions for a more focused examination of solutions that promote a better balance of trust and accountability.

II. Background

The UDRP is a widely used process for resolving domain name disputes globally. Over the years, companies have relied on the policy to address egregious instances of trademark infringement targeted at their brands. For instance, a large cybersecurity firm used it to stop a malicious campaign that utilized a domain name incorporating their trademarks to send personalized emails to unsuspecting users, purporting to be an officer at the company offering job opportunities [17]. In another instance, a third party used a domain name copying the trademark of a famous fashion outlet to sell counterfeit products [18]. A further example relates to an individual who registered a domain name and harassed the brand owner to buy it back at the exorbitant cost of $5.5 million [19]. In each of those cases, the policy helped trademark holders disrupt ongoing abuse of their brands.

Besides stopping bad-faith domain name uses, the policy also serves as a shield against the bullying practices of some trademark holders. For example, a domain name owner who acquired a domain name for over $1 million won a dispute brought against them by a company that had made multiple unsuccessful attempts to buy the domain [20]. There have also been hundreds of cases where the respondents' registrations predated the complainants' trademark rights, yet the latter filed a UDRP complaint after failing to acquire the domain in question (e.g., cassy[.]com in 2025 and lawcloud[.]com in 2023). Fortunately, the panels for those disputes decided that the domain should remain with the respondent. In an ideal scenario, the UDRP would achieve both objectives: to disrupt trademark infringement and protect legitimate third-party owners securing names for their products and services.

To guarantee its continued performance, the policy relies on interactions between various actors, as shown in Fig. 1. The process begins with a trademark holder (the complainant) gathering evidence to make a case as to why one or more third-party domain names *(i)* are identical or confusingly similar to their trademark(s), and *(ii)* were registered with no rights or legitimate interests, and *(iii)* were registered and being used in bad faith. Once all the requisite details are obtained, the complainant chooses one of the ICANN-approved dispute resolution service providers with whom to file their complaint. At present, only five providers can settle disputes.

1) The World Intellectual Property Organization (WIPO), an international organization headquartered in Switzerland;
2) The National Arbitration FORUM (FORUM), headquartered in the United States;

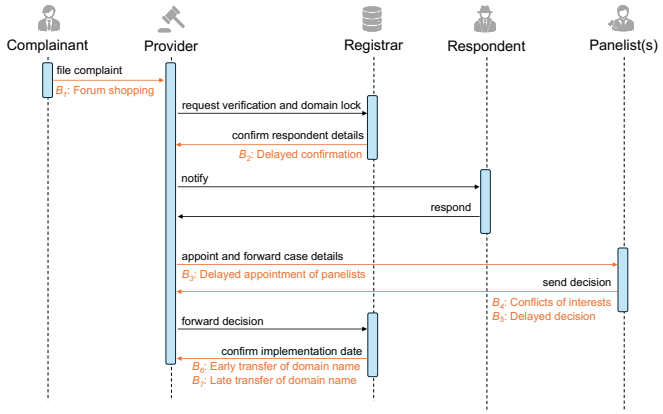

Fig. 1. Stages of a domain name dispute under the Uniform Domain Name Dispute Resolution Policy annotated with behaviors (B_i) that can undermine trustworthiness.

3) The Asian Domain Name Dispute Resolution Centre (ADNDRC), headquartered in mainland China with offices in Malaysia, Korea, and Hong Kong;
4) The Czech Arbitration Court Arbitration Center for Internet Disputes (CAC), based in the Czech Republic;
5) The Canadian International Internet Dispute Resolution Centre (CIIDRC), headquartered in Canada.

Upon receiving the complaint and confirming its compliance with the rules, the provider directs a request to the registrar of the contentious domain name to put the domain name on lock, and confirm the identity of the registrant within two business days. The lock ensures that the domain's ownership information does not change during the proceedings, and the confirmation of details guarantees that the registrant (i.e., the respondent) receives all subsequent communications about the case. After receiving the registrar's confirmation, the provider notifies the respondent, who is given 20 days to respond to the complainant's contentions. At the end of that period or after the respondent replies, the provider appoints a panel of one or three panelists. According to the rules, the complainant initially chooses the panel size. When the complainant opts for a one-member panel, the respondent can increase the panel's size from one to three. After appointing the panel, the provider forwards the case's files to the panelists, who have 14 days to decide. Once a decision is reached and shared with the provider, the registrar is asked to implement the decision (i.e., cancel or transfer the domain to the complainant in successful cases or simply remove the lock otherwise).

Although the overall process is more nuanced, this description effectively describes the interplay between the various actors whose behaviors can undermine the trustworthiness of the entire process. We refer interested readers to the policy document [21], the official rules [22], and the supplemental rules of WIPO [23], FORUM [24], ADNDRC [25], CAC [26], and CIIDRC [27] for more details.

III. Related Work

Given its prominence, UDRP has been the subject of study for many years, especially in legal scholarship. For instance, a recent study of the operational differences between providers called the policy an overly successful measure against cybersquatting [9]. Historically, this view has also been shared by the security community, endorsing the policy's potential to curb domain name squatting [28]–[30]. Some view it as an effective way to protect a company's online brands, albeit more costly than proactive defensive registrations [31]–[33]. Others suggest that its fast resolution mechanism can help rein in squatting in TLDs that adopt it [28], [34], [35].

Surprisingly, only a small segment of the voluminous related literature has been dedicated to *data-driven* analysis of how the actions of the parties involved could erode trust in the overall process. That is despite the fact that, much like with certificate revocation procedures, where trusted parties' carelessness can lead to prolonged security threats that undermine trust in the revocation process [36], [37], the proper observance of the policy's rules by all involved parties is tantamount to maintaining trust in the policy. Instead, the vast majority of work has theorized on the potential for misaligned incentives [14]. For instance, immediately after its adoption, Geist [12] and Mueller [38] argued that the policy's rules promoted forum shopping, defined as a complainant's deliberate choice of a provider to increase their chances of winning a dispute [12]. Despite Donahey et al.'s rebuttal of the allegations of forum shopping [15], follow-up studies of the disputes found that complainants were filing more disputes with providers that decided in favor of complainants more often, a situation that could undermine the perception of fairness [6], [39], [40].

One reason for the dearth of empirical studies may be that the technology (e.g., transformer architectures) needed to reliably extract pertinent data from the dispute proceedings has not been readily available until recently. In the limited empirical studies that exist, prior work examined the correlation between providers' market share and complainants' win rates to find occurrences of forum shopping. While that approach led to positive results in the early 2000s when win rates varied widely between providers [38], its most recent applications led to inconclusive results [9], [41], [42]. As we show later, by undertaking a more granular assessment focused on the *legal counsels representing complainants*, we find that forum shopping of this type—where counsels constantly alternate between providers based on factors that are highly correlated with historical win rates—is indeed common. Our analysis provides the first-ever numerical assessment of the prevalence of forum shopping in disputes, and our findings suggest a need for policy changes to minimize the perceived appearance of conflicts and improve transparency.

Lastly, some studies have measured the average time it takes to receive a decision after filing a UDRP dispute [8], [9]. By contrast, we zoom in on the delays induced by registrars, providers, and panelists. By analyzing the discrepancy between the allotted and the elapsed time at various stages of the

dispute process, we showed that ICANN should exercise better oversight and provide more incentives to make dispute timelines more predictable.

Taken as a whole, our work stands apart in that we not only empirically assess the issue of forum shopping and concerns about bias of some panelists towards trademark holders [12], [41], but also inquire into registrars' compliance and other issues that can undermine the policy's trustworthiness.

IV. Data set

A. Data Collection

ICANN requires that all providers maintain a list of cases filed with them on their website. The listings include the case number, the parties involved, the domain names in dispute, the publication date, and the outcome. Additional details about a case, such as the complaint's submission date and the panelists involved, are contained within a document shared as a PDF or HTML file. Providers are only required to share proceedings for cases decided by panelists; hence, no proceedings are available for cases pending a decision, withdrawn, or settled. The significant fraction of cases without proceedings reported for WIPO and FORUM in Table I (19.70% and 10.11%, respectively) results from the fact that those providers recorded the most withdrawn or settled disputes.

TABLE I
Number of Disputes by Provider (Dec 1, 1999 - Aug 18, 2024)

Provider	Disputes	Proceedings (%)	English Proceedings (%)
WIPO	70,112	57,003 (81.30%)	**50,192 (71.59%)**
FORUM	38,314	34,442 (89.89%)	**34,306 (89.54%)**
CAC	4,497	4,491 (99.87%)	**4,477 (99.56%)**
ADNDRC	3,231	3,231 (100.00%)	**1,013 (31.35%)**
CIIDRC	187	184 (98.40%)	**162 (86.63%)**
All	**116,341**	**99,357 (85.40%)**	**90,153 (77.49%)**

We downloaded the proceedings and extracted their textual content using the pypdf [43] Python library for the PDF files and a headless Firefox Selenium driver for the HTML files [44]. We failed to extract the text content for a few cases due to broken links (83) or PDFs with scanned content (75). For the 99,149 proceedings we extracted the text from, we used Facebook's state-of-the-art Fasttext language model to identify the text's language [45]. While most of the proceedings were in English (90.93%), many others, mainly from WIPO and ADNDRC, were written in other languages, with Chinese (2,518), Spanish (2,271), and French (1,716) being the most frequent. Given our intention to analyze the decisions in depth, we focused on the English cases since English was the only language common to all the authors and the most represented in the proceedings. In doing so, we omit 9.07% of cases, which could affect the generalizability of our findings, especially regarding ADNDRC. Overall, our filtering left us with 90,153 disputes related to 175,269 disputed domain names.

B. Data Curation

To extract details from the freeform texts, we trained a named-entity recognition (NER) model using a RoBERTa transformer [46] through the spaCy's NER interface [47]. This approach was chosen because prior work showed that training a NER model from scratch using this combination leads to accurate and fast models [48]. The chosen pipeline is also flexible and well-documented, facilitating the training, testing, and usage of our custom NER model. Our training data comprised between 10 and 100 randomly sampled cases per provider, with the exact sample size computed using a min-max scaling approach on the number of cases they handled. Likewise, each provider had between 3 and 25 sample decisions in our testing data set. We manually annotated the training and testing samples using Doccano [49]. We used 229 annotated documents to train our model, which surpasses the set size of 180 used elsewhere for training new pipelines [50]. Our testing data set contained 61 documents.

As summarized in Table II, the average F-scores ranged from 76.65 to 95.08% for 12 of the 15 named entities we trained the model to extract. The performances were less optimal for the remaining 3 named entities because the information they capture is not consistently reported in the proceedings. Nonetheless, our model achieved an overall average F-score of 79.44% with adequate performances per provider, as reported in Table V in the appendix. This allowed us to extract pertinent details from the dispute proceedings to study past decisions at a deeper level and on a larger scale than prior work that used commercial data [9]. For instance, unlike the commercial data, our curated data set contains the filing date for FORUM's 30K+ disputes, which our model extracts with high accuracy. Furthermore, nine of the details our model can extract accurately are not provided in the commercial data (i.e., $e_2 - e_5, e_8 - e_9, e_{11}, e_{13}, e_{15}$).

TABLE II
Performance of the NER Model per Entity Type

#	Entity	F-Score	#	Entity	F-Score
e_1	Filing date	95.08%	e_9	Complainant's counsel	80.00%
e_2	Registrar request date	89.86%	e_{10}	Respondents	76.65%
e_3	Registrar response date	87.04%	e_{11}	Respondent's address	73.33%
e_4	Commencement date	86.24%	e_{12}	Respondent default	80.41%
e_5	Panel appointment date	77.52%	e_{13}	Registrar	82.91%
e_6	Publication date	93.33%	e_{14}	Panelist	83.80%
e_7	Complainant	80.23%	e_{15}	Trademark	74.06%
e_8	Complainant's address	66.67%			

Given the freeform nature of the text, it is not uncommon to encounter different names referring to the same entity (e.g., TJX and TJX Companies, Inc., Waitomo Adventures Ltd and Waitomo Adventures Limited, Debrett Gordon Lyons and Debrett G. Lyons). We addressed inconsistencies in the data by combining text normalization (e.g., removing spaces, punctuation, and corporate endings) with reference data (e.g., a historical list of accredited registrars). Furthermore, to arrive at a consistent set of addresses, we used a popular API (from Nominatim [51]) to retrieve the country from each extracted address, and leveraged a large language model (GPT-4o) to

address known failure cases when using the API [52]. That overall process allowed us to extract a rich set of consistent data from the proceedings. To our knowledge, this is the most comprehensively curated URDP data set to date.

To promote further research, we publicly shared our curated data set and accompanying code.

V. General statistics: By the Numbers

To provide context for readers unfamiliar with the policy, we first present high-level insights from the analyzed proceedings.

A. Parties Involved

In our curated data set, there were 33,036 distinct complainants for an average of 3.15 cases filed per complainant. Most complainants (73.41%) only filed a single complaint, and (99.42%) filed at most a hundred disputes, which amounted to 76.35% of all the cases.

While only a handful of complainants (0.29%) filed more than a hundred cases, their cases account for 23.22% of the entire dispute data. The distributions are in line with other studies covering the policy's performance [9]. A deeper assessment using historical data from the Fortune Magazine revealed that 64 of the 95 companies with more than a hundred cases were listed on one of the published Fortune 500, Global 500, or Fortune Europe 500 lists. In fact, the 10 most active complainants are all regular members of the Global 500 ranking. Those companies' popularity might explain why they are often targets of trademark infringement, leading them to file more disputes than others.

Excluding respondents with redacted names, we identified 67,071 respondents, most of whom were involved in only a single dispute. Of the remaining respondents, most had fewer than ten disputes. We found 446 (0.68%) respondents involved in at least 10 disputes, totaling 12,394 (13.75%) disputes. As shown in Fig. 2, 62.56% (279) of these high-frequency respondents lost every dispute brought against them (0% win rate), while 73.77% (329) won less than 5%, and 96.41% (430) won under half.

The 16 remaining respondents involved in more than 10 disputes had win rates between 60% and 100%. After inspecting a random sample of 10 cases filed against those well-performing respondents, we found that three-member panels decided the case in all but one dispute. Again, those observations are consistent with prior work showing that respondents are more likely to win a case when it is decided by three-member panels and the respondent participates [9], [53]. A cursory examination of the data suggests that respondents who are legally represented in the proceedings do appear to have better chances of winning, but unfortunately, the number of documents with legally represented respondents in our training data was too low to conduct that analysis. We leave that as an exercise for future work.

B. Factors Correlated with Winning Outcomes

Using the 90,153 English cases our study focuses on, we found that complainants win most cases. This aligns with

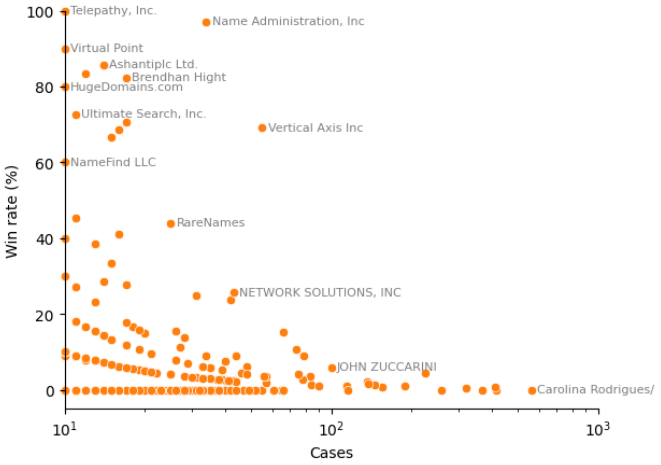

Fig. 2. No. of cases and winning rates for respondents involved in at least ten disputes. Some respondents won many of the disputes brought against them.

ICANN's recent analysis [8, p. 62] covering the 2013 - 2020 observation period in which they estimated a win rate of 93%. For that same period, we estimate a similarly high win rate of 91.87% overall. Those high win rates might not come as a surprise since the policy was designed to address the most flagrant types of cybersquatting, with other disputes left to courts [54]. The win rate is especially high for cases arbitrated by one panelist, with win rates for complainants above 95.00% for most providers when the registrant defaults.

Consistent with Cogburn et al.'s results [9], we found that when the registrant participates in the proceedings, the win rate drops but remains above 70% for all providers. The only exception is CIIDRC, with a win rate of 63.2% for complainants, as shown in Fig. 3. Those results demonstrate that with one panelist, complainants win most of their disputes even when the registrant participates in the proceedings.

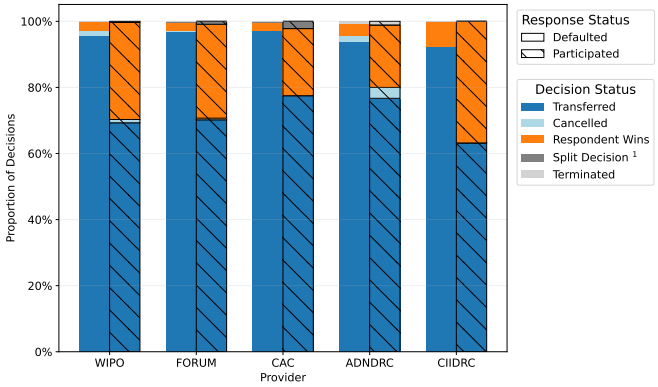

Fig. 3. Decision outcomes by provider and respondent participation for cases decided by a one-member panel

[1]Split decisions occur in cases involving multiple domain names where some are transferred and some denied.

However, cases arbitrated by three panelists decrease the complainant's chances of winning the case. We found that while the win rate remains above 70% for most providers when the registrant defaults, it quickly drops below 50% when the registrant participates in the proceedings, as shown in Fig. 4. This drop in win rate makes sense and aligns with arguments from legal scholars [12], [55] that state that cases decided by three-member panels are distinctly different than those decided by single-member panels because the former is only appointed if one of the parties requests it. In fact, our data analysis supports Christie's conjecture [55] that parties select a three-member panel when the outcome seems uncertain: complainants do so when they anticipate a close decision, and respondents do so when they see a viable defense.

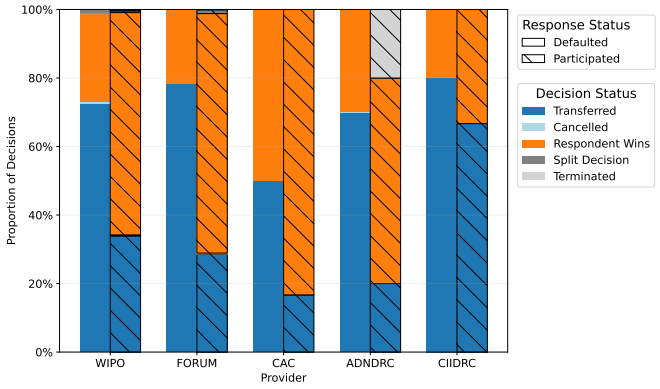

Fig. 4. Decision outcomes by provider and respondent participation for cases decided by a three-members panel

C. The Disputed Domain Names

The disputes we analyzed involved 175,269 distinct domain names registered in 463 TLDs. As we show in Fig. 5, most of the disputed domain names were registered in the legacy gTLDs (i.e., com, net, and org). That is not surprising, given the popularity of those TLDs among all domain name registrants. In fact, by conducting a Pearson correlation analysis, we found that most TLDs appeared in the domain name disputes just as frequently as people registered domain names in them ($r = 0.99, p < 0.01$). There were a few outliers, mainly among the new gTLDs from the 2012 round of applications, that were disputed more or less frequently than expected. Korczynski et al. made similar observations while comparing, for each TLD, the number of domains registered under it that appeared in phishing lists and the size of its zone file [56].

Using our NER model, we extracted the complainants' trademarks from the proceedings and subsequently classified 150,206 (85.70%) disputed domain names based on the transformation type they belonged to. Because trademarks often contain special characters or spaces that cannot be used in domain names, we converted the extracted trademarks into effective second-level domains (e2LD) by applying ICANN's

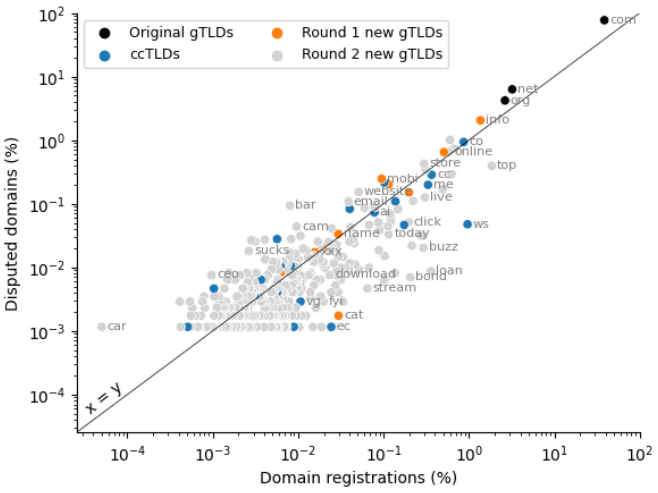

Fig. 5. Disputed domain names and number of registered domain names per top-level domain name. The rate of disputes related to a TLD has a strong linear correlation to the size of its zone file.

identical match rule for the new gTLDs' trademark clearinghouse (TMCH) [57]. The conversion consists of either removing or replacing all DNS impermissible characters with a dash and optionally replacing the characters "@" and "&" with their textual representation (i.e., *at* and *and* in English, respectively). Using those e2LDs, we combined Adjibi et al.'s generation models [33] with substring matching to determine the underlying transformations. As we show in Fig. 6, most of the disputed domain names could have been generated by combining complainants' trademarks with other terms. When registered and used maliciously, such domain names are considered instances of combosquatting [58], a type of abuse challenging to defend against proactively due to the unbounded space of possibilities. That said, among the remaining transformations, the simple identical match was the most frequent, affecting more distinctive marks than generic ones (i.e., marks that use dictionary words or common names). Identical matches are followed by various typo variants of the trademarks. This ordering of the transformations echoes an analysis of a sample of WIPO's disputes, which shows that, in 2019, more than 60% of disputed domain names combined a trademark with a term, less than 30% were identical matches, and fewer than 10% were typos [59].

D. Financial Costs

Using UDRP providers' current schedule of fees, we estimated that complainants spent approximately $12,365,343 on 7,188 disputes without legal representation, averaging $1,720 per dispute. For disputes where complainants were legally represented, we added a flat legal fee of $3,000 (based on the median price charged by six law firms [60]–[65]) to the mandatory filing fees. This calculation revealed that legally-represented complainants spent about $381,001,629

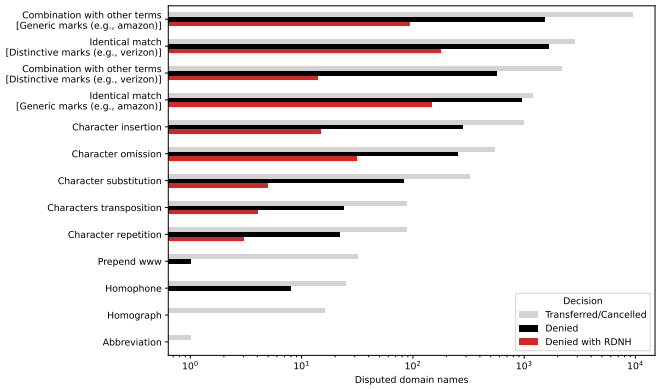

Fig. 6. Common transformations applied to trademarks. Most disputed domains incorporate the complainant's trademark entirely, often combining it with other terms.

on 82,962 disputes, averaging $4,592 per dispute.[2] To put these numbers into context, the American Intellectual Property Law Association found that in 2023, a trademark opposition case cost $168,000 on average [67]. More closely related are the attorney fees incurred when parties seek recourse under the Anticybersquatting Consumer Protection Act following unfavorable UDRP decisions. A brief review of cases decided by U.S. federal judges reveals that legal fees around $100,000 are common, with some cases far exceeding that amount. For example, in 2016, a district court awarded a plaintiff $236,752 in attorneys' fees[3].

While UDRP certainly offers a cheaper alternative for trademark owners, at the current median cost for registering a domain, the $95,134,681 (24% of the total) spent on the 21,385 disputes involving *identical* trademark matches is more than 200 times the cost of preemptively registering those domain names—roughly $447,689 in total. Even when including annual renewal fees, trademark owners would have spent at least an order of magnitude less if they had proactively registered the identical matches of their trademarks. This finding suggests that claimants, especially those with distinctive and fanciful marks, should consider proactively registering some domains to lower the cost of protecting their brands online while not stifling opportunities for legitimate domain name uses.

Regrettably, many trademark holders have been caught trying to weaponize the UDRP to ruthlessly disrupt legitimate uses of domain names that they claimed infringed on trademarks, even when there was no abuse. In UDRP, such baseless disputes are often sentenced with a **reverse domain name hijacking** (RDNH) ruling that attests to the bad-faith of the claimant. In the past, there have been 596 RDNH decisions, of which 490 (82.21%) were related to identical matches. Our analysis of case annotations [68] reveals that the most common reason panelists deemed complainants abusive was because they failed to provide evidence of bad-faith registration (351 cases). The second most common reason was that the complainants acquired trademark rights long after the respondent had registered the domain (290 cases). Panelists also ruled that claimants submitted misleading or incomplete claims (210 cases), pursued disputes despite knowing the respondent had legitimate rights (204 cases), or attempted to seize a domain through a dispute after failing to buy it directly (169 cases). These deceptive tactics highlight how some legitimate domain owners are forced to defend themselves against overreaching complainants. Even if only half of the victims of RDNH were legally represented, those respondents would have each spent on average $1,950 in legal fees [60], [63], yielding a minimum sum of $581,100. Thus, for RDNH cases *alone*, the cost to respondents for fighting the abuse could be almost a hundred times more than what it costs for a regular domain registration.

We caution the reader that while the fact that the number of RDNH verdicts is low might indicate that the policy is rarely abused in this way, the reality is more complex [69]. First, there are no guidelines on when a reverse domain name hijacking finding is appropriate, so panelists often shy away from making such rulings [70]. Second, because the policy does not allow a wrongfully accused registrant to recover monetary damages, there is little incentive to go through the time and effort of arguing for an RDNH judgment in their response, or for requesting a panel ruling on it. Third, many cases of such cyberbullying may never reach the dispute stage at all, as registrants might simply be persuaded to transfer domains after receiving cease-and-desist letters outside of the official dispute process [71].

VI. An Examination of Behaviors That Can Undermine Fairness and Efficiency

We examine how interactions between parties can enable behaviors undermining the UDRP's fairness and efficiency. Across the dispute process, we identify seven such behaviors (Fig. 1) that are enabled by gaps in the policy. Assessing the prevalence of these behaviors is critical to guide future reform.

As presented in Table III, each analysis relied on a subset of the entities extracted from the proceedings. To guarantee the accuracy of our findings, we only used data related to providers from whose proceedings we could extract the relevant details with a median F-score higher than 75%, which we represent in the table by a ✓ symbol. For the 90,153 disputes studied and presented longitudinally in Fig. 7, we report on our findings affecting *fairness* in Section VI-A and those related to *efficiency* in Section VI-B. Later on, in Section VII, we provide recommendations on mitigating or suppressing the dispositions that facilitate the occurrence of the discussed behaviors.

A. Behaviors Undermining Fairness

Two of the behaviors we studied could affect the policy's perception as a fair dispute mechanism. Those include the practice of forum shopping, whose prevalence was until now

[2]These figures offer a conservative estimate given that a prominent brand protection company estimated that, by the end of 2009, companies spent nearly $200 million for disputes [66]. Extrapolating that amount to today's dispute volume (of over 3.29x) yields roughly $857 million.

[3]David Dent v. Lotto Sport Italia SpA, U.S. District Court, D. Ariz., 2021.

TABLE III
RELEVANT ENTITIES USED FOR ASSESSING EACH BEHAVIOR WITH THE MEDIAN F-SCORES PER PROVIDER. THE ✓ SYMBOL MARKS THE PROVIDERS THAT WE USED FOR EACH BEHAVIOR'S ANALYSIS.

B_i	Entities	WIPO	FORUM	CAC	ADNDRC	CIIDRC
B_1	$e_1, e_8, e_9, e_{11}, e_{14}$	✓ 89.8%	✓ 97.3%	✓ 78.6%	47.6%	40.0%
B_2	e_2, e_3, e_{13}	✓ 94.1%	✓ 93.6%	-	66.7%	50.0%
B_3	e_4, e_5	✓ 93.0%	✓ 95.0%	-	54.0%	25.0%
B_4	e_1, e_9, e_{14}	✓ 94.1%	✓ 97.3%	✓ 100.0%	✓ 88.9%	45.5%
B_5	e_5, e_6, e_{14}	✓ 92.3%	✓ 94.7%	✓ 100.0%	✓ 88.9%	45.5%
B_6	e_1, e_6, e_{13}	✓ 84.8%	✓ 97.3%	✓ 100.0%	✓ 95.7%	✓ 80.0%
B_7	e_1, e_6, e_{13}	✓ 84.8%	✓ 97.3%	✓ 100.0%	✓ 95.7%	✓ 80.0%

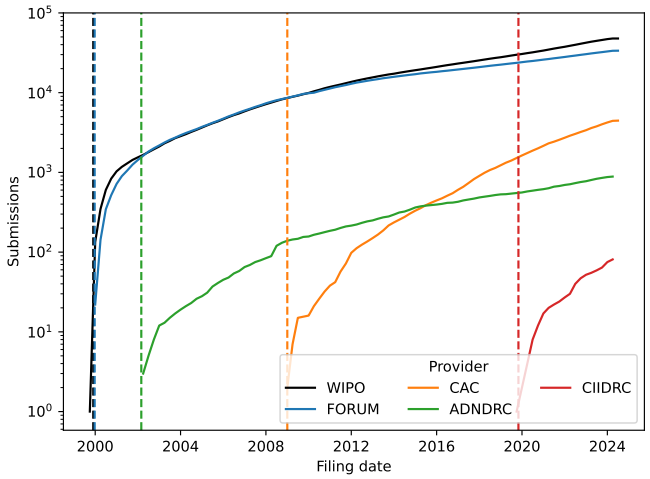

Fig. 7. Distribution of disputes by provider. The dashed colored lines represent the date ICANN accredited the corresponding provider. Best viewed in color.

anecdotal, and the panelists' conflicts of interest, which have often been alluded to but never been systematically studied.

1) Forum Shopping [B_1]:

> *Highlight*: Almost **one-third of the disputes** indicate practices suggestive of forum shopping.

In UDRP, any entity may initiate an administrative proceeding by submitting a complaint in accordance with the policy and the rules to *any* provider approved by ICANN. Critics argue that this creates an incentive for providers to favor complainants, as trademark holders may select providers based on their perceived likelihood of winning. This practice, known as *forum shopping*—where a complainant deliberately chooses a provider to increase their chances of winning [12]—raises concerns that systematic bias toward trademark holders could undermine credibility and fairness.

We examined forum shopping by analyzing how much past win rates might have influenced legal representatives' choice of a provider for client disputes. Legal representatives, rather than individual complainants, are the focus of our study since they engage in disputes more frequently. For example, Nameshield, the most active legal representative in our data, has handled almost three times as many cases as the most active complainant, Philip Morris S.A. (2,631 vs 930). Given that these firms specialize in domain name disputes, their success directly impacts their financial and professional standing. As a result, legal representatives are the most relevant group when studying forum shopping.

In our data, we found 10,357 legal representatives involved in 79,385 (92.52%) disputes. After identifying uninterrupted periods during which they consistently used a specific UDRP-accredited provider, we grouped the legal representatives into four groups, graphically represented in Fig. 8. There were: 1) ***loyalists*** who filed all their disputes with one provider; 2) ***one-timers*** who filed all but one of their disputes with one provider; 3) ***brand switchers*** who transitioned entirely from one provider to another; and 4) ***opportunists*** who constantly alternated between two or more providers for their disputes.

Because brand switchers and opportunists were the only groups that had filed many disputes with multiple providers, we restricted our search for evidence of forum shopping to those two categories. While doing so, we conservatively defined what constitutes evidence of forum shopping to account for the fact that legal representatives might have diversified their panel selection for legitimate reasons, such as a provider's geographical proximity to the contending parties or the provider's processing speed.

For legal representatives labelled as brand switchers, we consider them to have engaged in forum shopping if the last dispute they filed before switching to a new provider resulted in an unfavorable decision. Our data shows that 62 out of 218 brand switchers match that behavior. Those companies represented complainants in 355 disputes related to 643 domain names. For 44 (70.97%) of the companies, the negative decision preceding their switch was the first unfavorable ruling they received after filing between 1 and 26 disputes with their initial provider.

For companies in the opportunist category, we identified forum shopping by examining the correlation between their historical win rates with a provider and the likelihood of selecting that provider for a dispute. To estimate this likelihood, we trained logistic regression models to understand the decision-making processes of legal representatives. This approach, commonly used by economists to study consumer choices [72], is appropriate because it captures the assumption that rational companies will invariably select the provider that best fits their needs—whether that means winning the dispute or receiving a faster decision. The same method was used two decades ago to study [6] how individual complainants—*as opposed to legal representatives*—chose which provider to file their disputes with when UDRP was in its infancy.

To train our models in a practical setting, we used features that encompass information available to complainants before they file a dispute. Our features included the number of domain names in dispute, the parties' geographical locations, and the number of panelists requested. Taking inspiration from the econometric model [6, p37-48] used by Kesan and Gallo in their assessment of performance as a factor in choosing a provider in the early days of UDRP, we also included the

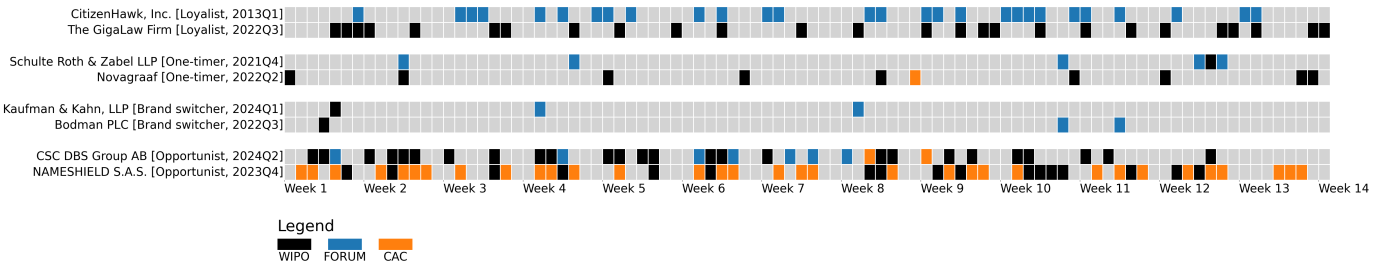

Fig. 8. Dispute submission pattern of select legal representatives over a quarter, showing the main categories of behaviors. The plot shows that while some representatives mainly stick with the same provider, others constantly switch between providers, often within a day.

company's win rate with each provider in the past quarter, and the natural logarithm of the average number of days it took to settle the company's disputes.

For that evaluation, we trained our models on the earliest 80% of an opportunist's submissions and tested them on the latest 20%. This assessment validated our approach, achieving a median F1 score of 89.04% for 256 opportunists involved in at least ten disputes. Using the models' estimates for the probabilities, we ran a Pearson correlation test between the historical win rate of an opportunist with a provider and the probability that they select that provider for a dispute. At a 95% confidence interval, we found that for 169 of 291 opportunists, the likelihood that they selected a provider positively correlated with their historical success rate with that provider ($\bar{r} = 0.35$). Those companies represented complainants in 26,622 disputes, a third of all the disputes in which complainants were represented (see Table IV).

TABLE IV
NO. OF DISPUTES WITH LEGAL REPRESENTATION, SHOWING THAT A THIRD OF THE DISPUTES INVOLVED FORUM SHOPPING.

Category	Likely forum shoppers	Others	All
Loyalists	-	34,548 (43.52%)	34,548 (43.52%)
One-timers	-	4,806 (6.05%)	4,806 (6.05%)
Brand switchers	355 (0.45%)	1,446 (1.82%)	1,801 (2.27%)
Opportunists	26,622 (33.53%)	11,608 (14.62%)	38,230 (48.16%)
All	**26,977 (33.98%)**	**52,408 (66.02%)**	**79,385 (100.00%)**

Those findings provide the first quantification of the scale of forum shopping in disputes while confirming some of the earlier results from Kesan and Gallo [6] who found that various factors such as dispute duration and past win rate influenced complainants' selection of a provider. Our results indicate that forum shopping may have affected 26,977 disputes—33.98% of disputes involving legal representatives and 29.92% of all disputes analyzed. To rule out any bias from early UDRP events, we focused our analysis on disputes settled after January 2020 (prior to the final Phase 1 Working Group meeting on October 29, 2020). Even then, we found 161 potential forum shoppers involved in 10,378 out of 30,465 disputes, confirming that forum shopping remains a persistent issue in the disputes' proceedings.

2) Potential Conflicts of Interests [B_4]:

> *Highlight*: **Only four of the most active panelists had acted as both panelists and counsels for a given complainant, with an average of three years between roles**. Those results show little evidence of panelist bias yet highlight the need for clearer definitions of conflict-of-interest guidelines.

To preserve public confidence in UDRP decisions, the policy requires that panelists disclose any circumstances that could reasonably raise doubts about their impartiality or independence before deciding a dispute. Despite those measures, some critics have asserted that panelists, especially intellectual property (IP) lawyers, were biased towards trademark holders [73]. To assess the presence of this bias, we manually searched for the qualifications of the 20 most active panelists for each provider, among whom we found 46 IP lawyers. Eighteen of the identified IP lawyers were affiliated with firms that legally represented complainants in 2,925 disputes, as explained in Appendix A-B. In most cases, the complainants whose cases a panelist decided were different from those represented by the panelist's law firm. However, in a few instances, those two groups overlapped. We observed this in nine disputes filed by eight complainants and decided by panelists whose law firm later represented the same complainants in 39 disputes. On average, it took three years between the last time a panelist decided on a complainant's dispute and the first dispute in which their law firm represented the same complainant. Meanwhile, the fastest switch happened slightly after a year (384 days), and only two occurred within two years.

Taken together, those results indicate that panelists' role switches are rare and typically occur with significant time gaps. However, because their mere occurrence can raise doubts about impartiality, clarifying what disqualifies a panelist from deciding a dispute would help preserve trust in the policy.

B. Behaviors Undermining Efficiency

While prior work has shown that many disputes last longer than expected [9], there has been no empirical evaluation of the stage at which delays occur during the disputes and why they happen. In this section, we compared the observed timelines

of each stage of the disputes to the expected ones, highlighting and explaining the observed delays. Our analysis also provides the first assessment of the time it takes registrars to apply transfer decisions after the dispute process has ended.

1) Delayed Confirmation by Registrars [B_2]:

> *Highlight*: While most domain registrars responded to providers' requests for verification of respondent details within the prescribed two-business-day timeframe, several registrars took significantly longer—in some cases, **up to 12 additional business days**.

The policy's rules state that within two (2) business days of receiving the provider's verification request, the registrar shall provide the registration data for each of the specified domain names in the verification request. Obviously, if registrars unduly delay decisions by not answering the providers' requests in a timely manner, that practice would hurt efficiency and prolong the period of malicious activity.

For registrars involved in disputes, we measured how often they responded to provider requests within the required timeframe. Because providers and registrars are often located in countries with misaligned workdays and holidays, we considered responses received within three business days acceptable. This follows the common practice of public institutions granting individuals one additional day when a deadline falls on a public holiday. Figure 9—restricted to registrars handling at least 10 disputes for simplicity—shows that while most registrars respond on time, a smaller group consistently delays their responses. Over the study period, the median delay was between one and twelve additional business days. Similarly, for requests received from January 2020 onward, 927 responses (7.14%) were late by one to twelve business days, aligning with the overall statistics.

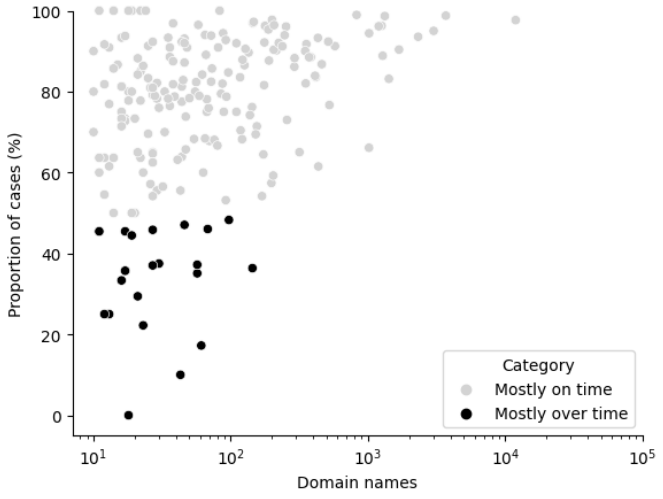

Fig. 9. Proportion of cases for which the registrars responded to the providers within two business days. We show that most registrars often respond in time, but others take longer than expected.

By assessing the Spearman-ranked correlation between the number of weekly requests a registrar received and the number of late confirmations, we observed a positive correlation for 24 registrars at the 99% confidence interval ($\bar{r} = 0.44$). These registrars handled 28,686 disputes (37.81%) related to 59,288 domain names. This finding suggests that the current systems registrars use to respond to providers might be inadequate for the volume of disputes they manage.

2) Delayed Appointment of Panelists [B_3]:

> *Highlight*: **Up to 88.53% of panels were appointed late**, with delays ranging between 1 and 48 days.

For one and three-member panels, appointments should be made within 25 and 35 days, respectively. Appointments made later risk eroding public trust and extending the period of malicious activity on the disputed domain names.

By estimating the days elapsed between the commencement date and the panel appointment date of WIPO and FORUM's proceedings, we found that both providers were often late. By the set deadline, WIPO only completed 11.47% and 16.17% of panel appointments of one and three panelists, respectively. FORUM did significantly better, completing 51.48% and 73.30% of panel appointments on time, respectively, for one- and three-member panels. After discarding outliers using the interquartile range method [74], we estimated that the number of additional days used by WIPO ranged from 1 to 48, with a median of 7 and 15 days, respectively, for one and three-member panels. On the other hand, FORUM used between 1 and 30 additional days, averaging 3 and 7 days, respectively, for one and three-member panels.

While those results might suggest that WIPO is slower than FORUM, we found that both providers appointed panelists at a similar pace. This suggests that WIPO's delays are a result of the number of disputes it handles; indeed, at a 99% confidence interval, the number of disputes for which WIPO was late in appointing panelists was strongly correlated ($r = 0.99$) with the number of disputes commenced in that month.

3) Delayed Decisions by Panelists [B_5]:

> *Highlight*: While most panelists rendered their decision within the prescribed limits, some repeatedly submitted decisions beyond the 14-day window, taking **up to 31 additional days**.

One of the touted strengths of the UDRP is its speed in rendering decisions. Because delays by panelists can prolong disputes and enable continued malicious activity, the policy mandates that, unless exceptional circumstances apply, panels must forward their decision to the provider within fourteen days of their appointment.

We estimated how often each panelist submitted their decision by the set deadline, considering decisions submitted within 15 days acceptable for the same reasons described in §VI-B1. To maintain clarity, we presented the results of that

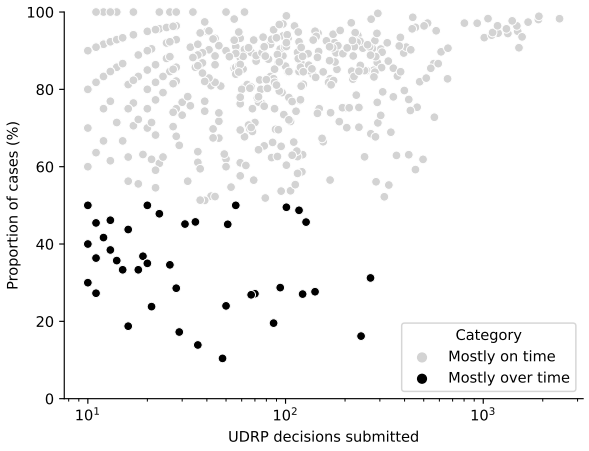

Fig. 10. Proportion of cases for which appointed panelists submitted their decision within 14 calendar days. We show that most panelists submit their decisions on time.

analysis for panelists who decided on at least ten disputes in Fig. 10. The plot shows that most panelists submitted their decisions on time, but a few others often made late submissions. Excluding outliers discarded using the interquartile range technique [74], the findings show that when panelists were late, they took between 1 and 31 additional days to decide on a case. The results are practically unchanged if we narrow the observation window to a more recent timeframe (between January and August 2024).

By running a Spearman-ranked correlation analysis, we found at the 99% confidence interval that for 95 panelists who decided on 30,045 cases (39.64%), the more panel appointments they received in a given two-week period, the more decisions they submitted late ($\bar{r} = 0.34$). This suggests that the overload of panelists might justify some of the delays they caused in the disputes.

4) Speed of Domain Names Transfers [$B_6 - B_7$]:

> *Highlight*: Registrars transferred successfully contested domains in a median of 24 business days, and only 172 domains were transferred prematurely—before the mandatory 10-business-day grace period. Conversely, **slow processing allowed 2,751 domains flagged as malicious by VirusTotal to remain active for up to 4 months after a transfer decision was issued**.

Registrars involved in a dispute must notify the provider of when they will implement the decision within three business days of receiving it. While the policy does not opine on the maximum number of days registrars can take before implementing a decision, it requires a pause of ten business days to allow either party to contest the decision in court. This guarantees a degree of fairness to the process, and the non-observance of that prescription can erode public trust in the policy. Additionally, the policy mandates that providers share the implementation date publicly to promote transparency. By failing to share that information, providers make it difficult to assess the policy's effectiveness fully.

Unfortunately, the dispute proceedings we analyzed did not record the effective transfer dates for domains receiving a transfer decision[4]. As a proxy for that information, we identified early and late transfers (B_6 and B_7) by inspecting the daily zone file data from the Centralized Zone Data Service [75] for 29,724 disputed domains across 344 TLDs from October 2020 to December 2024. Our analysis shows that 99.45% of these domains do not change name servers during the dispute process—a likely result of registrar-imposed locks and/or a reluctance to alter evidence by the respondents. We approximate the transfer date as the first day after the decision's publication when a domain's name servers change from those set when the dispute was filed. Although a name server change does not always indicate a transfer, given that the domains that received a transfer decision rarely changed name servers during the dispute, we argue this heuristic is reasonable. Additionally, we discarded transfer delta outliers by using the interquartile range method [74], resulting in a data set of 25,936 disputed domains across 12,979 disputes. We find the mean (μ), median ($\tilde{x}$), and standard deviation (σ) transfer deltas in business days to be: $\mu = 28.01$, $\tilde{x} = 24$ and $\sigma = 15.17$. As an extra and more conservative validation step, we measured the transfer delta only for 10,514 domain names (40.53%) that were transferred to name servers of online brand protection service providers and found that the mean, median, and standard deviation ($\mu = 28.68$, $\tilde{x} = 24$, $\sigma = 14.86$) are very close to the complete set of studied domain names—providing further evidence that our approximation is robust.

Given registrars must wait for 10 business days before applying the transfers [21], we consider an early transfer (B_6) as one that occurred sooner than that threshold. While there are no rules governing how promptly a transfer should be performed after the 10 business day grace period, we consider a late transfer (B_7) as a change in name servers that deviates significantly from the median transfer delta of popular registrars. Under that lens, we found 339 disputed domains (1.14%) from 172 disputes with an *early* transfer date, suggesting that registrars rarely act too quickly. Among the disputes with an early transfer date, case D2023-4776 was the most prominent, involving 33 domains registered with registrar "Domain Best Limited" and subsequently transferred to MarkMonitor's name servers only seven business days post-decision. That registrar had no other cases of early transfer.

Figure 11 reports the median dispute transfer delta (by registrar) after a decision was issued. We observe that for the studied registrars, it takes a median of 24 business days for

[4]In response to our requests for data pertaining to the implementation date, WIPO, FORUM, CAC, and CIIDRC all responded that they do not keep track of when decisions are implemented because that falls beyond the scope of their duties as providers. FORUM did mention that registrars provide the implementation date in about 25% of the cases, but stated it does not store that information. ADNDRC did not reply to our email.

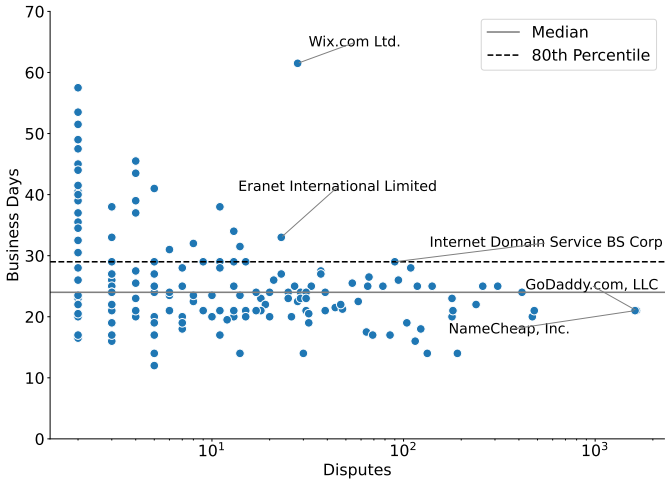

Fig. 11. Median dispute domain transfer delta in business days per registrar for registrars with more than one dispute. We show that registrars took a median of 28 business days to transfer domains.

the name servers of a disputed domain to change after the dispute. Large registrars with a high volume of cases, such as NameCheap and GoDaddy, have a transfer delta lower than the median, meaning that they enable faster transfers of the domain names to the claimants. Interestingly, registrars with mid to high numbers of registrations [76], such as Wix.com, Eranet International, and Internet Domain Service BS Corp., feature slower median transfers compared to the rest of the registrars, surpassing in many cases the 80^{th} percentile. Slower transfers, especially by popular registrars, could undermine the perception of fairness because claimants may feel delayed in controlling domain names they have rightfully won.

Worse yet, delays allow malicious activity to persist long after decisions. Specifically, we found that out of 5,678 domains flagged as malicious by at least one VirusTotal engine before receiving a transfer decision, registrars transferred just 2,966 within a median of 24 business days. They took up to 80 business days—four months—to transfer the rest. On average, five VirusTotal engines flagged these domains, suggesting that sluggish transfers could have prolonged their abusive use. These findings underscore the need for stricter oversight of registrars' UDRP enforcement.

VII. Discussion and Recommendations

Trust involves some form of expectation about outcomes. This is especially true in dispute resolution systems where there is an expectation that those entrusted with specific roles will perform competently and uphold shared standards. Because all parties in the ecosystem shape the overall reputation of the arbitration, the processes that are followed must instill confidence in the general public. In particular, given that one of the core tenets of UDRP is procedural fairness, trust hinges on the belief that the system works as expected, so any undue delays or hiccups in decision-making could sour that perception and chip away at the policy's integrity.

> **Operational Changes:** We recommend that ICANN consider the merits of several policy changes that could help mitigate loss of trust. Chief among them are policies that help curb forum shopping (see §VI-A1). While there is no easy fix, ICANN could revisit the controversial proposal that registrars choose the provider that handles disputes for their customers [38].

Irrespective of the merits of that proposal, any solution thereof must accommodate cases where a single respondent registers multiple domain names with different registrars—a scenario that occurred in 1,549 (2.04%) disputes in our data. Alternatively, disputes could be randomly assigned to providers [6], or to maintain competition without favoring larger providers, supplemental rules could be standardized, and providers' performance continuously evaluated using predefined objective metrics. These metrics would generate a score to assign new disputes to the most suitable provider.

The data also supports the suggestion that all disputes should be handled by three-member panels [12]. Because three-member disputes allow both parties to suggest three panelists for consideration [22], this change might enhance the perceived fairness of UDRP decisions while preserving the policy's incentives of being cheaper and faster than traditional litigation. Indeed, at the 99% confidence interval, we found that three-member panels took only one week longer than one-member panels. That said, such a change could face significant implementation challenges in practice as it imposes additional burdens on some stakeholders.

In regard to efficiency (see §VI-B), our findings suggest that the introduction of automation could make disputes consistently faster. For instance, ICANN could develop a secure interface between providers and registrars so that providers can access the information they need to commence the proceedings quickly, regardless of the number of domains in dispute. Similarly, providers could automate their panel assignment process while adjusting the size of their panelist roster to meet the growing demands of the domain name dispute market. The Financial Industry Regulatory Authority (FINRA)—the largest securities dispute resolution forum in the United States for adjudicating disputes between investors and stockbrokers—performs such an automated panel assignment using its "Neutral List Selection" tool [77]. In instances where delays still occur despite automation, it is essential to document the underlying reasons for the delays to guide potential policy adjustments. All those changes should be supplemented by better oversight. Specific to registrar compliance, it is worth acknowledging that ICANN has terminated some non-compliant registrars in the past (e.g., for not providing accurate registrant information in a timely manner or for not implementing panelists' decisions) [78], but many other functional registrars still consistently delay the proceedings. Undoubtedly, long delays not only impact the efficiency of disputes but cast doubts about predictability.

New Incentives: Given the observed delays, policies should also be explored that better align incentives with efficiency. For instance, intermediaries with low delays could be financially rewarded (e.g., changes or kickbacks relative to fee structures) or, similarly, for registrars with disproportionally high delays, their ICANN accreditation could be re-evaluated based on failure to comply with longstanding policies.

The use of incentives in other parts of the DNS ecosystem is not new. Probably the most well-known are the financial incentives that a number of country-code TLD registries offered to operators for implementing DNSSEC. Several rollouts [79], [80] credited the incentives for their success, and in the Netherlands, for example, nearly 60% of all .nl domains are now DNSSEC-enabled. That said, any incentives must be carefully designed to address systemic delays without undermining the integrity of the policy.

To mitigate concerns about biases, ICANN should provide clear guidance on what conditions disqualify a panelist from deciding a case. Such conflict-of-interest policies are well-established best practices for maintaining public trust. For instance, after programmatically filtering for known conflicts among its arbitrators, FINRA shares detailed information about each arbitrator's award history and known conflicts of interest with all parties before assigning a neutral panel to a case [77]. Likewise, UDRP providers can improve the perceived neutrality of their panel assignments by maintaining and sharing historical records about their panelists via mandatory arbitrator disclosure reports, as done in FINRA disputes[5].

Compliance: New compliance rules on enforcing decisions should be put into effect. Providers, with the help of registrars, should disclose the date when disputed domain names receiving a transfer decision were actually transferred. Alternatively, regulatory bodies such as ICANN or independent research groups can monitor the date of effective transfer of domain names on an *ongoing* basis to guarantee registrars' compliance.

The disclosure of more detailed information would allow for better insights into the inner workings of the policy and allow the governing organizations to explore mechanisms for incentivizing registrars to both respect the 10-day post-decision hold for a transfer to come into effect, as well as to promptly transfer the domain names after a reasonable period of time. Overall, our analysis suggest that updating areas where the policy falls short is critical to ensuring predictability and consistency in its interpretation and outcomes—factors that directly impact the public's trust in the URDP as a whole.

[5]A sample disclosure report is available at https://www.finra.org/sites/default/files/ArbMed/p122952.pdf. FINRA's policies require members to regularly disclose circumstances that might preclude them from rendering an objective and impartial determination in a proceeding.

VIII. Limitations

According to social scientists, trust in dispute processes like UDRP is undermined when non-compliance occurs repeatedly or when adjudicators render questionable decisions. In this study of the endogenous threats to UDRP's trustworthiness, we only investigated the former cause, leaving the more nuanced and complex examination of the consistency and fairness of panelists' interpretations of UDRP's rules for future work.

To ensure a rigorous analysis, we limited our study to proceedings published in English and extracted key information using a custom named entity recognition (NER) model. This decision introduces two primary limitations. First, by focusing exclusively on English-language proceedings, our study is not fully comprehensive, as it omits a small number of disputes published in other languages by certain providers. Second, due to the inconsistent formatting of UDRP proceedings, our NER model performed less reliably for certain entities—particularly those infrequently mentioned in the documents. To mitigate the effects of potential annotation errors, we validated and de-duplicated extracted information using reference lists, proprietary APIs, and pattern-matching techniques. While these measures significantly improve the reliability of our dataset, we acknowledge that some errors may persist. By drawing attention to the difficulties of applying state-of-the-art language models to inconsistently formatted UDRP proceedings, we hope that stakeholders take notice and move towards a standardized format that facilitates robust and scalable analyses. A recent expert-driven report authored by practitioners and scholars with deep experience in UDRP disputes made a similar recommendation [81]. To help fill the void until such recommendations become a reality, we are publicly releasing our annotated data and trained models. We hope this encourages other researchers to refine our models and design new multilingual extraction models that support broader analysis of past and future UDRP disputes.

Regarding our analysis of forum shopping, we did not consider whether providers' marketing strategies could have influenced the choices made by counsels. For instance, both WIPO and FORUM maintain regular newsletters that discuss recent cases and high-level statistics of their UDRP practices [82], [83]. Some providers have hosted seminars targeted towards "trademark practitioners and potential filing parties" [84]. Thus, the extent to which the confounding factor of "forum selling" [85] relates to forum shopping remains a direction for future work.

IX. Conclusion

The UDRP has long served as a fast, standardized mechanism for addressing trademark abuse involving domain names. Yet, even as the domain name system expands and new threats emerge, the policy has remained largely unchanged.

Our empirical analysis reveals that this stagnation has enabled behaviors (e.g., forum shopping and persistent procedural delays) that risk undermining the policy's credibility relative to its ability to fairly serve both trademark holders and

legitimate domain owners. Beyond offering actionable recommendations for modernizing domain name dispute resolution, our work also underscores the importance of expanding the security community's focus. For instance, as abuse increasingly occurs outside traditional channels—particularly on social media [34]—researchers can help guide policy by understanding where current mechanisms fall short and evaluating alternative mechanisms for enforcement and protection. Only by adapting these systems to the realities of today's Internet can we build a more resilient digital infrastructure that upholds fairness, transparency, and trust for all stakeholders in an ever-evolving online landscape.

X. Ethics Considerations

Because our study reveals concerning behaviors in domain name disputes, we recognize the potential for unintended reputational or financial impact. To mitigate this risk, we avoided naming specific individuals or companies when describing questionable behaviors, focusing instead on presenting robust empirical evidence for the broader community. In select cases, we redacted the names of the entities involved.

XI. Acknowledgements

The authors thank the anonymous reviewers for their detailed feedback that helped improve the paper. We express our profound gratitude to Professor Kathryn Kleiman for her insightful comments on an earlier draft of the paper and for assisting in sharing our results with UDRP stakeholders. We also thank Matthew Thomas and Tillson Galloway for their editorial comments and suggestions.

References

[1] D. Akhawe and A. P. Felt, "Alice in warningland: A large-scale field study of browser security warning effectiveness," in *22nd USENIX Security Symposium (USENIX Security 13)*, Aug. 2013, pp. 257–272.

[2] N. Chachra, D. McCoy, S. Savage, M. Geoffrey, and Voelker, "Empirically characterizing domain abuse and the revenue impact of blacklisting," in *Proceedings of the Workshop on the Economics of Information Security (WEIS)*, vol. 4, 2014, pp. 1–13.

[3] R. Shewchuk, "The UDRP and the ACPA: What are they, and which should be used," *Asper Review of International Business & Trade Law*, vol. 2, pp. 85–106, 2002.

[4] N. Ebner and J. Zeleznikow, "Fairness, trust and security in online dispute resolution," *Hamline Journal of Public Law & Policy*, vol. 36, no. 6, pp. 143–160, 2015.

[5] J. Lewis, "Trust as social reality," *Social Forces*, 1985.

[6] J. P. Kesan and A. A. Gallo, "The market for private dispute resolution services – an empirical re-assessment of ICANN-UDRP performance," *Michigan Telecommunications & Technology Law Review*, vol. 285, no. 11, pp. 285–380, Jul. 2004.

[7] A. Allemann, "Same complainant, same respondent. two different UDRP outcomes." Apr. 2025. [Online]. Available: https://domainnamewire.com/2025/04/10/same-complainant-same-respondent-two-different-udrp-outcomes/

[8] ICANN, "Revised uniform domain name dispute resolution policy (UDRP) report," 2022. [Online]. Available: https://gnso.icann.org/sites/default/files/file/field-file-attach/udrp-status-report-redlined-13jul22-en.pdf

[9] D. L. Cogburn, T. A. Ochieng, and H. M. Wong, "Towards an understanding of global 'private ordering' in ICANN: Text mining 23 years of uniform domain-name dispute resolution policy (UDRP) decisions," *Journal of Cyber Policy*, vol. 8, no. 2, pp. 186–217, Dec. 2023.

[10] ICANN, "Staff manager's issues report on UDRP review," Aug. 2003. [Online]. Available: https://archive.icann.org/en/gnso/issue-reports/udrp-review-report-01aug03.htm

[11] ——, "Final GNSO issue report on the current state of the uniform domain name dispute resolution policy," Oct. 2011. [Online]. Available: https://gnso.icann.org/sites/default/files/filefield_27051/udrp-final-issue-report-03oct11-en.pdf

[12] M. A. Geist, "Fair.com?: An examination of the allegations of systemic unfairness in the ICANN UDRP," *Social Science Research Network*, pp. 1–35, Aug. 2001.

[13] M. Mueller, "ICANN and internet governance: Sorting through the debris of "self-regulation"'," *Digital Policy, Regulation and Governance*, vol. 1, no. 6, pp. 497–520, 1999.

[14] M. Zalnieriute, "Beyond the governance gap in international domain name law: Bringing the UDRP in line with internationally recognized human rights," *Stanford Journal of International Law*, vol. 56, no. 1, pp. 79–122, 2020.

[15] M. S. Donahey, T. Z. Morosoli, and L. L. P. Maser, "The UDRP: Fundamentally fair, but far from perfect," *Electronic Commerce and Law Reports*, vol. 6, no. 34, 2001.

[16] R. Tushnet, K. Kleiman, M. Stoltz, C. H. Farley, P. Aufderheide, and M. Karanicolas, "Comments to the initial report of the WIPO-ICA UDRP review project team," Online, Jun. 2025.

[17] M. Schwimmer, "F5 Networks, Inc. v. ryan," Feb. 2017. [Online]. Available: https://www.wipo.int/amc/en/domains/decisions/text/2016/d2016-2475.html

[18] K. V. Fink, "Guess? IP Holder L.P. and Guess?, Inc. v. lee sugar," Jun. 2014. [Online]. Available: https://www.adrforum.com/DomainDecisions/1557429.htm

[19] H. Samaras, "Six Continents Hotels, Inc. v. Interbase, Inc." Apr. 2003. [Online]. Available: https://www.wipo.int/amc/en/domains/decisions/html/2002/d2002-1045.html

[20] W. A. Rothnie, "STABLE APP LLC v. Unlimint Holding EU Ltd/ Mr Admin - Cardpay Ltd," Aug. 2023. [Online]. Available: https://www.wipo.int/amc/en/domains/decisions/pdf/2023/d2023-2594.pdf

[21] ICANN, "Uniform domain name dispute resolution policy," Feb. 2024.

[22] ——, "Rules for uniform domain name dispute resolution policy," Sep. 2013. [Online]. Available: https://www.icann.org/resources/pages/udrp-rules-2015-03-11-en

[23] WIPO, "The wipo "supplemental rules"," Oct. 2009. [Online]. Available: https://www.wipo.int/amc/en/domains/supplemental/eudrp

[24] FORUM, "Supplemental rules," Jul. 2019. [Online]. Available: https://www.adrforum.com/assets/resources/domain/udrp/FinalUDRPSupplementalRules_Eff2019.docx

[25] ADNDRC, "The supplemental rules," Aug. 2023. [Online]. Available: https://www.adndrc.org/storage/files/udrp/udrp_supplemental_rules_en_revised.pdf

[26] Czech Arbitration Court, "CAC's UDRP supplemental rules," Sep. 2013. [Online]. Available: https://udrp.adr.eu/website/cacs-udrp-supplemental-rules-of-the-czech-arbitration-court

[27] CIIDRC, "CIIDRC supplemental rules," May 2018. [Online]. Available: https://ciidrc.org/domain-name-disputes/ciidrc-supplemental-rules/

[28] P. Agten, P. F. Wouter, and N. Nikiforakis, "Seven months' worth of mistakes: A longitudinal study of typosquatting abuse," in *Proceedings of the Annual Network and Distributed Systems Security Symposium*, Feb. 2015, pp. 1–13.

[29] J. Szurdi, B. Kocso, G. Cseh, J. Spring, M. Felegyhazi, and C. Kanich, "The long "taile" of typosquatting domain names," in *USENIX Security Symposium*, Aug. 2014, pp. 191–206.

[30] J. Szurdi and N. Christin, "Email typosquatting," in *Proceedings of the 2017 Internet Measurement Conference*, 2017, pp. 419–431.

[31] T. Halvorson, K. Levchenko, S. Savage, and G. M. Voelker, "XXXtortion? inferring registration intent in the .xxx TLD," in *Proceedings of the International Conference on World Wide Web*, Apr. 2014, pp. 901–912.

[32] B. C. Benjamin, J. Bayer, S. Fernandez, A. Duda, and M. Korczynski, "Shielding brands: An in-depth analysis of defensive domain registration practices against cyber-squatting," in *Network Traffic Measurement Annaul Conference*, May 2024, pp. 1–11.

[33] B. Adjibi, A. Avgeditis, M. Antonakakis, M. Bailey, and F. Monrose, "The Guardians of name street: Studying the defensive registration practices of the Fortune 500," in *Proceedings of the Annual Network and Distributed System Security Symposium*, Feb. 2025.

[34] B. Acharya, D. Lazzaro, E. López-Morales, A. Oest, M. Saad, A. E. Cinà, L. Schönherr, and T. Holz, "The imitation game: Exploring brand

impersonation attacks on social media platforms," in *Proceedings of the USENIX Security Symposium*, Aug. 2024, pp. 4427–4444.

[35] Y. Nadji, M. Antonakakis, R. Perdisci, D. Dagon, and W. Lee, "Beheading hydras: Performing effective botnet takedowns," in *Proceedings of the ACM SIGSAC Conference on Computer & Communications Security*. Association for Computing Machinery, Nov. 2013, pp. 121–132.

[36] D. Kim, B. J. Kwon, K. Kozák, C. Gates, and T. Dumitras, "The broken shield: Measuring revocation effectiveness in the Windows code-signing PKI," in *USENIX Security Symposium*, Aug. 2018, pp. 851–868.

[37] D. Kim, H. Cho, Y. Kwon, A. Doupé, S. Son, G.-J. Ahn, and T. Dumitras, "Security analysis on practices of certificate authorities in the HTTPS phishing ecosystem," in *Proceedings of the ACM Asia Conference on Computer and Communications Security*, 2021, pp. 407–420.

[38] M. Mueller, "Rough justice: A statistical assessment of ICANN's uniform dispute resolution policy," *The Information Society*, vol. 17, no. 3, pp. 151–163, 2001.

[39] A. M. Froomkin, "ICANN's uniform dispute resolution policy: Causes and (partial) cures," *Brooklyn Law Review*, vol. 67, no. 3, pp. 605–718, 2002.

[40] J. Selby, "Competitive justice: The role of dispute resolution providers under ICANN's UDRP," *Macquarie Journal of Business Law*, vol. 1, pp. 23–58, 2004.

[41] L. Kong, "A brief review of UDRP: Achievements, challenges and recommendations," *China and WTO Review*, vol. 3, no. 1, pp. 145–162, 2017.

[42] A. F. Christie, J. Gloster, and S. Goddard, "An empirical analysis of 15 years of australian domain name disputes," *Australian Intellectual Property Journal*, vol. 30, no. 1, pp. 1–33, 2019.

[43] M. Fenniak, "Welcome to pypdf," 2006. [Online]. Available: https://pypdf.readthedocs.io/en/stable/

[44] B. Muthukadan, "Selenium with Python," 2011. [Online]. Available: https://selenium-python.readthedocs.io/

[45] A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, "Bag of tricks for efficient text classification," in *Proceedings of the European Chapter of the Association for Computational Linguistics*, Apr. 2017, pp. 427–431.

[46] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "RoBERTa: A robustly optimized BERT pretraining approach," 2019.

[47] I. Montani, M. Honnibal, M. Honnibal, A. Boyd, S. V. Landeghem, and H. Peters, "Explosion/spacy: v3.7.2: Fixes for apis and requirements," Oct. 2023. [Online]. Available: https://doi.org/10.5281/zenodo.10009823

[48] H. Shelar, G. Kaur, N. Heda, and A. Poorva, "Named entity recognition approaches and their comparison for custom NER model," *Science & Technology Libraries*, vol. 39, no. 3, pp. 324–337, 2020.

[49] H. Nakayama, T. Kubo, J. Kamura, Y. Taniguchi, and X. Liang, "doccano: Text annotation tool for human," 2018. [Online]. Available: https://github.com/doccano/doccano

[50] M. Srivastava and P. Greaney, "Resume parsing across multiple job domains using a BERT-based NER model," in *31st Irish Conference on Artificial Intelligence and Cognitive Science*, 2023, pp. 1–4.

[51] Nominatim developer community , "Nominatim manual." [Online]. Available: https://nominatim.org/release-docs/develop/

[52] K. Clemens, "Qualitative comparison of geocoding systems using OpenStreetMap data," *International Journal on Advances in Software*, vol. 8, no. 3-4, pp. 377–386, 2015.

[53] M. Geist, "Fair.com: An examination of the allegations of systemic unfairness in the ICANN UDRP," *Brooklyn Journal of International Law*, vol. 27, no. 3, pp. 903–938, Aug. 2002.

[54] P. Segal, "Attempts to solve the UDRP's trademark holder bias: A problem that remains unsolved despite the introduction of new top level domain names," *CARDOZO ONLINE J. CONFLICT RESOL.*, vol. 3, p. 1, 2001.

[55] A. F. Christie, "Online dispute resolution: The phenomenon of the UDRP," in *Research Handbook on Cross-Border Enforcement of Intellectual Property*, 2014, pp. 642–682.

[56] M. Korczynski, S. Tajalizadehkhoob, A. Noroozian, M. Wullink, C. Hesselman, and M. Van Eeten, "Reputation metrics design to improve intermediary incentives for security of TLDs," in *IEEE European Symposium on Security and Privacy*, Apr. 2017, pp. 579–594.

[57] ICANN, "Explanatory memorandum: Implementing the matching rules," Jul. 2016. [Online]. Available: https://newgtlds.icann.org/sites/default/files/matching-rules-14jul16-en.pdf

[58] P. Kintis, N. Miramirkhani, C. Lever, Y. Chen, R. Romero-Gómez, N. Pitropakis, N. Nikiforakis, and M. Antonakakis, "Hiding in plain sight: A longitudinal study of combosquatting abuse," in *Proceedings of the ACM Conference on Computer and Communications Security*, 2017, pp. 569–586.

[59] G. Bleda and T. Gray, "The UDRP in numbers," Oct. 2019. [Online]. Available: https://www.wipo.int/export/sites/www/amc/en/docs/2019_udrp_numbers.pdf

[60] The Internet Business Law Attorneys, "Legal fees." [Online]. Available: https://www.theinternet.lawyer/legalfees

[61] Kirton McConkie PC, "Domain name lawyer in Utah | Kirton McConkie attorneys in SLC: Kirton McConkie," Online. [Online]. Available: https://www.kirtonmcconkie.com/practices-Domain-Name-Disputes

[62] 101domain, "Reclaim your digital assets: UDRP disputes - the 101domain blog," Online, Dec. 2024. [Online]. Available: https://blog.101domain.com/udrp-disputes

[63] D. L. Brinza, "Domain name dispute fees," Online. [Online]. Available: https://www.canadianiplaw.com/domain-name-dispute-fees.html

[64] Cislo & Thomas LLP, "Typical fees," Online. [Online]. Available: https://cisloandthomas.com/overview/typical-fees/

[65] Western IP Law LLC, "Domain name disputes," Online. [Online]. Available: https://utahpatentlaw.com/domain-name-disputes/

[66] PRWeb, "Brand owners could have prevented $220 million in domain name recovery by spending $1 million," Aug. 2009. [Online]. Available: https://web.archive.org/web/20090926223849/https://www.prweb.com/releases/domain_management/brand_protection/prweb2778974.htm

[67] American Intellectual Property Law Association, *Report of the Economic Survey*. American Intellectual Property Law Association, 2023.

[68] RDNH.com, "What is RDNH?" Jul. 2014. [Online]. Available: https://www.rdnh.com/rdnh-provision-udrp/

[69] C. A. Schultz and C. A. Hofflander, "Reverse domain name hijacking and the uniform domain name dispute resolution policy: Systematic weaknesses, strategies for the respondent, and proposed policy reforms," *Cybaris*, vol. 4, no. 2, pp. 218–249, 2013.

[70] Z. Muscovitch, "The hidden perils of filing a baseless UDRP complaint," Dec. 2018. [Online]. Available: https://circleid.com/posts/20181213_the_hidden_perils_of_filing_a_baseless_udrp_complaint

[71] L. C. Grinvald, "Policing the cease-and-desist letter," *University of San Francisco Law Review*, vol. 49, pp. 411–468, 2015.

[72] K. E. Train, *Discrete choice methods with simulation*. Cambridge university press, 2009.

[73] NCSG, "Policy status report: Uniform domain name dispute resolution policy (UDRP) NCSG comments," Apr. 2022.

[74] H. P. Vinutha, B. Poornima, and B. M. Sagar, "Detection of outliers using interquartile range technique from intrusion dataset," in *Information and Decision Sciences*, S. C. Satapathy, J. M. R. S. Tavares, V. Bhateja, and J. R. Mohanty, Eds., Singapore, 2018, pp. 511–518.

[75] ICANN, "Centralized zone data service." [Online]. Available: https://czds.icann.org/home

[76] Domain Name Stat, "List of domain name registrars." [Online]. Available: https://domainnamestat.com/statistics/registrar/others

[77] N. G. Iannarone, "Finding light in arbitration's dark shadow," *Nevada Law Journal Forum*, vol. 4, pp. 1–12, 2020.

[78] ICANN, "Notices of breach, suspension, termination and non-renewal." [Online]. Available: https://www.icann.org/compliance/notices

[79] A. Löwinder, "DNSSEC deployment in sweden: How do we do it? ICANN50, 2014."

[80] R. Gieben, "DNSSEC in NL," *URL: http://miek.nl/publications/dnssecnl/secreg-report. pdf Jan*, 2004.

[81] WIPO-ICA Project Team, "Initial report of the WIPO-ICA UDRP review project team," Online, Apr. 2025. [Online]. Available: https://www.wipo.int/export/sites/www/amc/en/docs/wipoicareportapril2025.pdf

[82] WIPO, "WIPO ADR highlights - archive." [Online]. Available: https://www.wipo.int/en/web/newsletters/archives?newsletter=adr_highlights

[83] FORUM, "Domain daily decision emails." [Online]. Available: https://cm.adrforum.com/daily-emails

[84] WIPO, "WIPO advanced workshop on domain name dispute resolution: Update on precedent and practice," Oct. 2022. [Online]. Available: https://www.wipo.int/amc/en/events/workshops/2022/domainname/

[85] D. Klerman, "Forum selling and domain-name disputes," *University of Chicago, Law Journal*, vol. 48, pp. 561–584, 2016.

TABLE V
F-SCORES OF THE ENTITIES EXTRACTED FROM THE UDRP PROCEEDINGS USING WIPO CASE D2009-0190 AS AN EXAMPLE.

Entity Type	Excerpt from the Example	WIPO	FORUM	CAC	ADNDRC	CIIDRC	Overall
Filing date	"The Complaint was filed ... on February 12, 2009"	94.12%	97.30%	100.00%	95.65%	80.00%	95.08%
Registrar	"<tiffanyjewelry925.com> is registered with ... PublicDomainRegistry.com"	71.79%	93.05%	-	65.12%	50.00%	82.91%
Registrar request date	"On February 13, 2009, ... transmitted ... a request for registrar verification"	94.12%	-	-	87.50%	0.00%	89.86%
Registrar response date	"On February 19, 2009, ... transmitted ... its verification response"	95.83%	94.12%	-	66.67%	50.00%	87.04%
Commencement date	"... the proceedings commenced on February 19, 2009"	93.62%	94.44%	-	63.64%	50.00%	86.24%
Panel appointment date	"The Center appointed ... the sole panelist ... on March 17, 2009"	92.31%	94.74%	-	44.44%	0.00%	77.52%
Publication date	"Dated: March 30, 2009"	84.75%	100.00%	100.00%	100.00%	100.00%	93.33%
Complainant	"The Complainants are Tiffany (NJ) LLC ..."	75.00%	87.50%	80.00%	76.00%	92.31%	80.23%
Complainant's address	"Tiffany (NJ) LLC of Parsippany, New Jersey, United States of America"	76.60%	100.00%	-	45.45%	0.00%	66.67%
Complainant's counsel	"... (the "Complainants" or "Tiffany"), represented by Arnold & Porter LLP"	89.80%	100.00%	57.14%	44.44%	40.00%	80.00%
Respondent	"The Respondent is La Tondra Moultrie, TiffanyJewelry925.com"	69.84%	82.05%	66.67%	86.96%	61.54%	76.65%
Respondent's address	"La Tondra Moultrie, TiffanyJewelry925.com of Turku, Finland."	74.19%	94.12%	0.00%	47.62%	0.00%	73.33%
Respondent default	"The Respondent did not submit any response"	81.08%	92.06%	85.71%	63.41%	66.67%	80.41%
Panelist	"The Center appointed Nathalie Dreyfus as the sole panelist"	97.03%	76.36%	100.00%	88.89%	45.45%	83.80%
Trademark	"Complainants have rights in the trademark TIFFANY"	77.16%	68.39%	64.71%	78.14%	71.43%	74.06%
Overall		**82.03%**	**82.69%**	**74.36%**	**73.86%**	**60.87%**	**79.44%**

APPENDIX A
SUPPLEMENTAL INFORMATION

A. Data extraction

Table V provides the per-provider scores for the entities used in our analyses.

B. Firms

The anonymized data for the top panelists whose law firms also served as counsels for complainants is listed in Fig. 12.

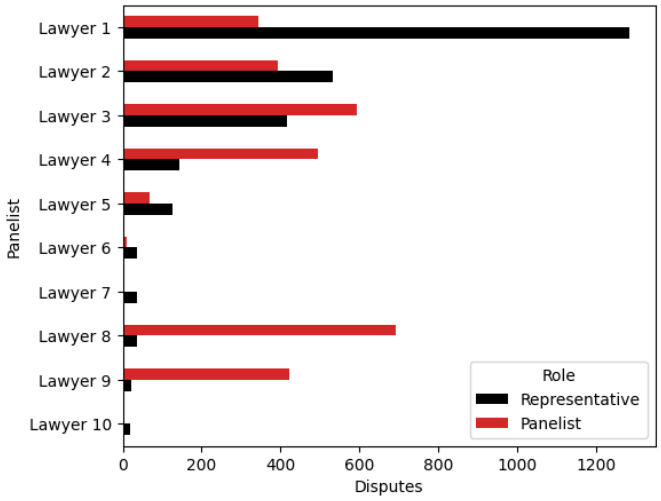

Fig. 12. Top 10 panelists whose law firms they work for also represented complainants in UDRP disputes.

APPENDIX B
ARTIFACT APPENDIX

The shared artifacts include the raw and curated UDRP data used in the study, the named entity recognition (NER) model used for the experiments, and the source code used to measure the prevalence of forum shopping and dispute delays ($B_1 - B_3, B_5$). Those resources support the core analyses presented in the study, excluding data related to potential panelists' conflicts of interest, as sharing such information offers limited value but could negatively affect some individuals or businesses. Additionally, due to the Centralized Zone Data Service (CZDS) terms prohibiting data-sharing, the current artifacts do not include the resources used to measure how fast registrars transferred domains to winning complainants after the disputes. Nevertheless, interested parties with access to CZDS data can use the shared curated dataset to replicate those experiments independently.

A. Description & Requirements

1) How to access: Download the source code and supporting data at https://doi.org/10.5281/zenodo.16954717, then uncompress the downloaded file.

2) Hardware dependencies: These artifacts can be executed on any commodity machine with at least 16 GB of RAM, but processing will be faster on a GPU-powered machine.

3) Software dependencies: To run these artifacts, Python version 3.7 or higher is required.

4) Benchmarks: None

B. Artifact Installation & Configuration

- Create a virtual environment in the project's root directory using venv, conda, or similar tools.
- Install the Python dependencies by running the command below:

```
pip install -r requirements.txt
```

- Follow the instructions on `spaCy`'s website to obtain a command to install the most appropriate package for your device. In the setup wizard, select the options as follows:
  - *Configuration*: Check both "virtual env" and "train models".
  - *Trained pipeline*: Check English.
  - *Select pipeline for*: Check "accuracy".
- Before installing `spaCy`, change the relevant command to force the download of version 3.7.6, which was used to

train the model used in our analysis. Do this by inserting "==3.7.6" before the closing quote of the command. For example, the final command might look like this.

```
pip install -U
↪  'spacy[transformers,lookups]==3.7.6'
```

1) Troubleshooting:

- MacBook users might fail to install the Python dependencies because of an error while compiling the package `tokenizers`. To fix this issue, install and configure the Rust compiler using the commands below.

```
curl https://sh.rustup.rs -sSf | sh
# add the binary to your path. For example
echo '. "$HOME/.cargo/env"' >>
↪  .venv/bin/activate
source .venv/bin/activate
# set Rust flags to avoid errors
export RUSTFLAGS="-A
↪  invalid_reference_casting"
```

- Users of Python 3.12 and above might fail to visualize annotated documents with `displacy` because of an error related to a deprecated API. To solve that problem, affected users can follow the workaround discussed here.

C. Experiment Workflow

Each of the following experiments can be executed independently.

D. Major Claims

- (C1): Using manually annotated UDRP proceedings, we trained an NER model that extracts relevant information from published UDRP proceedings with high accuracy. We demonstrate this with experiment E1, whose results are presented in Tables II, III, and V.
- (C2): Using our trained NER model, we extracted details from 90,153 UDRP proceedings and deduplicated the extracted values using reference lists, proprietary APIs, and pattern matching as described in Section IV-B. We demonstrate this with experiment E2.
- (C3): As described in Section VI-A1, we grouped legal representatives into various categories and inferred whether they engaged in forum shopping from their filing patterns. We demonstrate this with experiment E3.
- (C4): Using our curated data set, we found that providers, registrars, and panelists often delayed UDRP proceedings by many days. We demonstrate this with experiment E4 that reproduces Figures 9 and 10, and visually represents the statistics discussed in Section VI-B2.

E. Evaluation

1) Experiment (E1): [Model training and evaluation][15 human-minutes + 4.5 compute-hours on CPU / 40 compute-minutes on GPU]: This experiment trains and evaluates a NER model using previously annotated proceedings.

[Preparation]

1) Run all the cells in the Jupyter notebook located at `code/e1.data-validation.ipynb` to generate general statistics about the collected proceedings, including those shown in Table I. The notebook also provides statistics on the annotated proceedings and includes a visualization of a randomly selected annotated UDRP case for reference.
2) Using the command below, generate random data splits to train and evaluate the NER model, overall and per-provider. The number of proceedings per provider in the training or testing data sets can be configured via the supporting script's command line options.

```
python3 code/e1.sample-ner-datasets.py
↪  data/disputes-content.jsonl.gz
↪  data/annotated-udrp-corpus.jsonl.gz data/
```

[Execution]

1) Train and evaluate the model. Note that `spaCy` trains new pipelines on the CPU by default. The optional flag `-gpu-id` must be set to train the model using a GPU.

```
spacy train data/ner-config.cfg --paths.train
↪  data/train-ALL.spacy --paths.dev
↪  data/test-ALL.spacy --output model
↪  [--gpu-id 0]
```

2) After the training is complete, evaluate the model's performance per provider to generate a table like Table V. As before, set the `-gpu-id` option to run the evaluations on the GPU instead of the CPU.

```
spacy evaluate model/model-best
↪  data/test-WIPO.spacy --output
↪  results/models-performance/WIPO.json
spacy evaluate model/model-best
↪  data/test-FORUM.spacy --output
↪  results/models-performance/FORUM.json
spacy evaluate model/model-best
↪  data/test-CAC.spacy --output
↪  results/models-performance/CAC.json
spacy evaluate model/model-best
↪  data/test-ADNDRC.spacy --output
↪  results/models-performance/ADNDRC.json
spacy evaluate model/model-best
↪  data/test-CIIDRC.spacy --output
↪  results/models-performance/CIIDRC.json
```

[Results]

1) Run all the cells in the Jupyter notebook located at `code/e1.model-performance.ipynb` to generate tables reporting the model's F-scores for each entity. Note that while the exact values may vary slightly due to randomness inherent to model training and evaluation, the results should closely match those presented in Tables II, III, and V.
2) Setup the newly trained model for inference by creating and installing a dedicated Python package using the commands below.

```
spacy package model/model-best packages --name
↪  udrp_extractor --version 0.0.1
cd packages/en_udrp_extractor-0.0.1/dist/
pip install en_udrp_extractor-0.0.1.tar.gz
```

3) Next, follow the instructions in the Jupyter notebook located at `code/e1.extract-data.ipynb` to annotate a random proceeding with the new model.

2) Experiment (E2): [Data extraction] [5 human-minutes + 4 compute-hours on CPU / 40 compute-hours on GPU]: This experiment uses *a subset of the disputes* to demonstrate how to extract meaningful information from UDRP proceedings.

[Preparation]

1) Create a package for the model used for our analyses and install it using the commands below from the project's root directory:

```
spacy package resources/model-best/ packages/
↪  --name udrp_extractor_baseline --version
↪  0.0.1
cd packages/en_udrp_extractor_baseline-0.0.1/
pip install
↪  dist/en_udrp_extractor_baseline-0.0.1.tar.gz
```

[Execution] Run all the cells in the Jupyter notebook at `code/e2.extract-and-clean-details.ipynb`.

[Results] The parsed proceedings will be saved to the file `data/Jan2023-Aug2024-parsed-proceedings.jsonl.gz` that can be easily loaded with the Pandas library. The last execution cell shows a snippet of the parsed data.

3) Experiment (E3): [Assessment of forum shopping] [5 human-minutes + 30 compute-minutes]: This experiment implements the methodology we used to identify which legal representatives potentially engaged in forum shopping.

[Preparation] None

[Execution] Run all the cells in the Jupyter notebook located at `code/e3.analysis-forum-shopping.ipynb`.

[Results] The last two execution cells show the median F1-score of the logistic regression models introduced in Section VI-A1 and the proportion of disputes potentially affected by forum shopping. The results should be similar to those reported in the paper, though not necessarily identical.

4) Experiment (E4): [Measurement of procedural delays] [2 human-minutes + 5 compute-minutes]: This experiment measures the duration of each stage of UDRP disputes against the expected timeframes.

[Preparation] None

[Execution] Open the notebook located at `code/e4.analysis-efficiency.ipynb` and run its cells sequentially.

[Results] The program generates three figures: the first two should be similar to Figures 9 and 10, and the last one should show the proportion of cases in which each provider appointed panelists on time, as described in Section VI-B2.