The right AIs for the rights of citizens

Francesca Lagioia, Marco Lippi, Paolo Torroni
Bologna, Alma AI Seminar, June 17th, 2020
PART II

The CLAUDETTE system

(CLAUseDETecTEr)

Automation of personal data and consumer law enforcement using AI
How to empower consumers?

• **Protection** against unwanted monitoring (GDPR)

• **Support** in detecting unfair use of AI

• **Control** commercial practice fairness

“An opposing exercise of power is the principal solvent of economic power, the basic defense against its exercise in economic affairs”. Ken Galbraith

In the AI era an effective countervailing power needs to be supported by AI.
Automatically detect potentially unfair clauses in Terms of Services and Privacy Policies

• Consumers agree but don’t read

• NGOs have competence to control but lack resources

• Business keeps using unlawful clauses
Terms of Service (ToS): The Training Set

The ToS Corpus

WHERE DID WE START?

... 50 ToS (manually annotated)...

7,090 sentences, 787 of which (11.1%) were labeled as positive, thus containing a potentially unfair clause.

WHERE ARE WE NOW?

... 100 ToS (manually annotated)...

claudette.eui.eu
Part 1: Unfair Contract Terms Law and Practice

Directive 93/13 art 3.1:

A contractual term which has not been individually negotiated shall be regarded as unfair if, contrary to the requirement of good faith, it causes a significant imbalance in the parties' rights and obligations arising under the contract, to the detriment of the consumer.

Bottom-line: there are some types of clauses that traders are prohibited from using in the contracts.

claudette.eui.eu
8 unfairness categories
(Art. 3 of Directive 93/13)

<table>
<thead>
<tr>
<th>Type of clause</th>
<th>Symbol</th>
<th># clauses (50 Tos)</th>
<th>#documents (50 Tos)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arbitration</td>
<td>&lt;a&gt;</td>
<td>44</td>
<td>28</td>
</tr>
<tr>
<td>Unilateral change</td>
<td>&lt;ch&gt;</td>
<td>188</td>
<td>49</td>
</tr>
<tr>
<td>Content removal</td>
<td>&lt;c&gt;</td>
<td>118</td>
<td>45</td>
</tr>
<tr>
<td>Jurisdiction</td>
<td>&lt;j&gt;</td>
<td>68</td>
<td>40</td>
</tr>
<tr>
<td>Choice of law</td>
<td>&lt;law&gt;</td>
<td>70</td>
<td>47</td>
</tr>
<tr>
<td>Limitation of liability</td>
<td>&lt;ltd&gt;</td>
<td>296</td>
<td>49</td>
</tr>
<tr>
<td>Unilateral termination</td>
<td>&lt;ter&gt;</td>
<td>236</td>
<td>48</td>
</tr>
<tr>
<td>Consent by using</td>
<td>&lt;use&gt;</td>
<td>117</td>
<td>48</td>
</tr>
</tbody>
</table>

1) clearly fair; 2) potentially unfair; 3) clearly unfair
Consent by using Clause

If a clause states that the consumer is bound by the terms of service simply by visiting the website or by downloading the app, or by using the service: potentially unfair

A potentially unfair consent by using clause (Airbnb):

<use2>By accessing or using the Airbnb Platform, you agree to comply with and be bound by these Terms of Service.</use2>

A potentially unfair consent by using clause (Facebook):

<use2>By using or accessing the Facebook Services, you agree to this Statement, as updated from time to time in accordance with Section 13 below.</use2>
Jurisdiction Clause

Where a dispute will be adjudicated?

If giving consumers a right to bring disputes in their place of residence: clearly fair

If stating that any judicial proceeding takes a residence away (i.e. in a different city, different country): clearly unfair

A clearly unfair jurisdiction clause (Dropbox):

<j3> You and Dropbox agree that any judicial proceeding to resolve claims relating to these Terms or the Services will be brought in the federal or state courts of San Francisco County, California, subject to the mandatory arbitration provisions below. Both you and Dropbox consent to venue and personal jurisdiction in such courts.</j3>
An online server

claudette.eui.eu
**An online server**

CLAUDETTE

An Automated Detector of Potentially Unfair Clauses

Claudette found 1 potentially unfair clause (displayed in **bold**) out of 1 sentences.

<table>
<thead>
<tr>
<th>Potentially unfair clause #1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>We may stop (permanently or temporarily) providing the Services or any features within the Services to you or to users generally.</strong></td>
</tr>
</tbody>
</table>

Unfairness categories: **Unilateral Termination**

---

The clause is potentially unfair for **Unilateral Termination** since the contract or access can be terminated where the user fails to adhere to its terms, or community standards, or the spirit of the ToS or community terms, including inappropriate behaviour, using cheats or other disallowed practices to improve their situation in the service, deriving disallowed profits from the service, or interfering with other users' enjoyment of the service or otherwise puts them at risk, or is investigated under any suspicion of misconduct. (score = 0.992)

---

The clause is potentially unfair for **Unilateral Termination** since the contract or access may be terminated where the user has been engaging in illegal or unlawful activity, including fraudulent behaviour, abusive, misusive or otherwise harmful behaviour, or for reasons of safety or fraud prevention (score = 0.770)

---

The clause is potentially unfair for **Unilateral Termination** since the contract or access may be terminated for any reason, without cause or leaves room for other reasons which are not specified. (score = 0.638)
CLAUDETTE meets GDPR
The Golden Standard: Lawfulness Fairness Transparency

Comprehensiveness of information
The policy should contain all the information required by articles 13 and 14 of the GDPR.

Clarity of expression
The policy should be framed in an understandable and precise language.

Substantive compliance
The policy should only allow for processings of personal data that are compliant with the GDPR.

Different Levels of Achievement: Optimal and Suboptimal (questionable or insufficient)

claudette.eui.eu
**Examples of failure**

Failure under the substantive dimension

**Epic games Privacy Policy (last updated on 24 May 2018)**
<cuse3> when you use our websites, games, game engines, and applications, you agree to our collection, use, disclosure, and transfer of information as described in this policy, so please review it carefully.</cuse3>

**Rationale**
The clause above is an unfair processing clause since it states that the data subject consents to the collection, use, disclosure and transfer of his/her information, and thus s/he is bound by the privacy policy, simply by using the Epic Games web-sites, games, game engines and applications.

Failure under the comprehensiveness dimension

**Facebook Privacy Policy (last updated on 19 April 2018)**
<dpo2>Contact the Data Protection Officer for Facebook Ireland Ltd.</dpo2>

**Rationale**
The clause above fails to be fully informative since it generically refers to the possibility of contacting the DPO but does not provide the DPO name and a postal address, only a link to an online form. Thus, it only reaches a low standard for the clarity and accessibility of the information.
WHERE DID WE START?

- 14 documents (100 now)
- 3,658 sentences
- 80,398 words
- 11.0% sentences contain unclear language
- 33.9% sentences contain potentially unlawful clauses
WEB-CRAWLER

Developed as a tool for automatic privacy policy monitoring
Two types of monitoring:
• Checking the date on the document
• Comparison of the content with the previously saved version
Earnings reports by e-mail
What we are working on ...

• Experimenting new method for privacy policies

• Multilingualism (The Claudette german version)

• Empowerment through transparency
  • Linguistic transparency
  • Provide explanations opening black box AI Systems
Selected Publications


[Image] claudette.eui.eu