What’s Feminist Data Protection?

Dr. h.c. Marit Hansen
State Data Protection Commissioner of Schleswig-Holstein, Germany
CAU, Kiel, 8 March 2024
Overview

1. Why data protection?
2. Feminist data protection
3. Gendered digital harms – examples
4. Challenges for enforcement & for system design
5. Conclusion
6. Reading List

Imbalance in power ⇒ data protection necessary

Important: Perspective of the individual

More than security of personal data

Source: beludise via Pixabay
Overview

1. Why data protection?
2. Feminist data protection
3. Gendered digital harms – examples
4. Challenges for enforcement & for system design
5. Conclusion
6. Reading List

**Perspective: Alice & Bob**

Alice

- a, g, p
- $A = g^a \mod p$
- $K = B^a \mod p$

Bob

- b
- $B = g^b \mod p$
- $K = A^b \mod p$

IT security: The adversary is Eve (or Mallory).

Data processing: interference with fundamental rights

**Data Protection: The adversary is Bob!**
(At least: Bob is one of them.)
Overview

1. Why data protection?

2. Feminist data protection

3. Gendered digital harms – examples

4. Challenges for enforcement & for system design

5. Conclusion

6. Reading List

What's Feminist Data Protection?

"So it's not about women?"
Overview

1. Why data protection?
2. Feminist data protection
3. Gendered digital harms – examples
4. Challenges for enforcement & for system design
5. Conclusion
6. Reading List

Example 1: (High-)tech domestic violence

https://pbs.twimg.com/media/Ds7fJIPWsAA5t7G?format=jpg&name=large (2019)
Overview

1. Why data protection?

2. Feminist data protection

3. Gendered digital harms – examples

4. Challenges for enforcement & for system design

5. Conclusion

6. Reading List

Example 2: FemTech industry, e.g. period apps

Missed period? The significance of period-tracking applications in a post-Roe America

Bridget G. Kelly a, Maniza Habib b

a Director of Research for Sexual and Reproductive Health and Rights, Population Institute
b Research Associate, Population Institute, Washington, DC, USA

Keywords: sexual and reproductive health and rights, period-tracking apps, bodily autonomy


Example 3: Surveillance advertising

An investigation by netzpolitik.org is set to change this fundamental imbalance between the adtech industry and internet users. In June, we published a series of articles shining a light on the collection, trade and use of personal data in the global adtech industry. We analyzed an inventory file from a US-based data marketplace called Xandr. The file contains more than 650,000 so-called audience segments. These are used by advertising companies to categorize and target billions of people.

The scope and detail of this data collection is staggering. There is hardly a human characteristic that advertisers do not want to exploit for their purposes. Want to reach people in Denmark who have bought a Toyota? No problem. Italians with financial problems? No problem. Minors in Austria? Hardcore Christians in Portugal? Pregnant women in Poland? Fragile seniors in France? Queers in Spain? No problem.

Ingo Dachwitz: Surveillance advertising in Europe: The adtech industry tracks most of what you do on the Internet. This file shows just how much. 6 July 2023 (German version: 8 June 2023), https://netzpolitik.org/2023/surveillance-advertising-in-europe-the-adtech-industry-tracks-most-of-what-you-do-on-the-internet-this-file-shows-just-how-much/
Example 3: Surveillance advertising

Cancer, Depression and Eating Disorders

Hundreds of segment labels point to highly sensitive data such as health information. Advertisers can choose from categories such as breast cancer, bladder cancer and depression. Many segments also refer to reproductive health, period tracking, menopause or heavy buyers of pregnancy test kits. Some segment names even refer to visitors to individual clinics. Here are some examples from US supplier LiveRamp:

- LiveRamp Data Store | 8082 | 16237485 | HealthRankings > BPD
- LiveRamp Data Store | 8082 | 16237395 | HealthRankings > BPH
- LiveRamp Data Store | 8082 | 16237478 | HealthRankings > Breast Cancer
- LiveRamp Data Store | 8082 | 24900788 | HealthRankings > Breast Cancer Caregivers
- LiveRamp Data Store | 8082 | 16237416 | HealthRankings > Cholesterol
- LiveRamp Data Store | 8082 | 16237450 | HealthRankings > Cough/Cold
- LiveRamp Data Store | 8082 | 16237432 | HealthRankings > Diabetes
- LiveRamp Data Store | 8082 | 16237508 | HealthRankings > Diabetes Type II
- LiveRamp Data Store | 8082 | 16237498 | HealthRankings > Eating Disorder
Overview

1. Why data protection?

2. Feminist data protection

3. Gendered digital harms – examples

4. Challenges for enforcement & for system design

5. Conclusion

6. Reading List

Example 3: Surveillance advertising

The second largest group is demographics. Advertisers can select not only by gender or age, but also, for example, parents of teenagers, single mothers with small children or people who are about to get divorced. Lifestyle information is often included, such as „conservative retirees“, „urban elites“ or even „multicultural families“. Mothers seem to be a particularly interesting group; there are segments for „soccer moms“, „big city moms“, „busy moms“ or even „moms who shop like crazy“.

Ingo Dachwitz: Surveillance advertising in Europe: The adtech industry tracks most of what you do on the Internet. This file shows just how much. 6 July 2023 (German version: 8 June 2023), https://netzpolitik.org/2023/surveillance-advertising-in-europe-the-adtech-industry-tracks-most-of-what-you-do-on-the-internet-this-file-shows-just-how-much/
Example 4: Toeslagenaffaire

Dutch scandal serves as a warning for Europe over risks of using algorithms

The Dutch tax authority ruined thousands of lives after using an algorithm to spot suspected benefits fraud — and critics say there is little stopping it from happening again.

Chermaine Leysner’s life changed in 2012, when she received a letter from the Dutch tax authority demanding she pay back her child care allowance going back to 2008. Leysner, then a student studying social work, had three children under the age of 6. The tax bill was over €100,000.

“I thought, ‘Don’t worry, this is a big mistake.’ But it wasn’t a mistake. It was the start of something big,” she said.

The ordeal took nine years of Leysner’s life. The stress caused by the tax bill and her mother’s cancer diagnosis drove Leysner into depression and burnout. She ended up separating from her children’s father.

**Example 4: Toeslagenaffaire**

Dutch scandal serves as a warning for Europe over risks of using algorithms

The Dutch tax authority ruined thousands of lives after using an algorithm to spot suspected benefits fraud — and critics say there is little stopping it from happening again.

Leysner is one of the tens of thousands of victims of what the Dutch have dubbed the “toeslagenaffaire,” or the child care benefits scandal.

In 2019 it was revealed that the Dutch tax authorities had used a self-learning algorithm to create risk profiles in an effort to spot child care benefits fraud.

Authorities penalized families over a mere suspicion of fraud based on the system’s risk indicators. Tens of thousands of families — often with lower incomes or belonging to ethnic minorities — were pushed into poverty because of exorbitant debts to the tax agency. Some victims committed suicide. More than a thousand children were taken into foster care.

Example 4: Toeslagenaffaire

Dutch scandal serves as a warning for Europe over risks of using algorithms

The Dutch tax authority ruined thousands of lives after using an algorithm to spot suspected benefits fraud. And it's happening again.

European Parliament

In order to create risk profiles of individuals applying for childcare benefits, the Dutch Tax and Customs Administration used algorithms in which foreign sounding names and dual nationality were used as indicators of potential fraud. As a result, thousands of (racialised) low- and middle-income families were subjected to scrutiny, falsely accused of fraud and asked to pay back benefits which they had obtained completely legally. Thus, the algorithms led to racial profiling.

The consequences were very difficult for many families: thousands of families went into debt and many ended up in poverty because they were asked to pay back large sums of money (in the tens of thousands of euros), without being allowed the possibility of payment arrangements. Some lost their homes or their jobs. So far, more than 1 000 children have been taken out of their homes and placed in state custody as a result of the accusations[2]. These numbers are estimated to be higher.

Example 5: “Gender penalty” in AI systems for fraud detection

Of the 30,000 welfare recipients in Rotterdam, roughly the top 10 percent, or everyone ranked above 27,000 on the list, are at risk of being investigated.

Imagine a typical 30-year-old male born in Rotterdam who receives social welfare to make ends meet. He ranks 16,815 on the list, with over 10,000 people between him and a potential investigation into benefits fraud.

https://www.wired.com/story/welfare-state-algorithms/
Overview

1. Why data protection?
2. Feminist data protection
3. Gendered digital harms – examples
4. Challenges for enforcement & for system design
5. Conclusion
6. Reading List

Example 5: “Gender penalty” in AI systems for fraud detection

Let's compare his likelihood of being investigated to Sara's by adjusting key attributes about him until, to the algorithm, he becomes Sara.

If our typical Rotterdam male were female, she'd move up 4,542 spots closer to a possible investigation.

https://www.wired.com/story/welfare-state-algorithms/
Overview

1. Why data protection?
2. Feminist data protection
3. Gendered digital harms – examples
4. Challenges for enforcement & for system design
5. Conclusion
6. Reading List

Example 5: “Gender penalty” in AI systems for fraud detection

Financially Struggling Single Mother

First, we wanted to understand the impact of being both a woman and a parent. We know that independently, both women and parents scored higher. To illustrate what we did, let's begin with a simple test and calculate the average score of all the combinations between gender and parenthood.

<table>
<thead>
<tr>
<th>Has no children</th>
<th>Man</th>
<th>Woman</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.461</td>
<td>0.483</td>
</tr>
<tr>
<td>2 primary school children</td>
<td>0.507</td>
<td>0.525</td>
</tr>
</tbody>
</table>

We can see in the table above the [gender penalty] is 0.018 for women with children and 0.022 for women with no children.

Example 5: “Gender penalty” in AI systems for fraud detection

Overview

1. Why data protection?
2. Feminist data protection
3. Gendered digital harms – examples
4. Challenges for enforcement & for system design
5. Conclusion
6. Reading List

Intersection of Our Archetypes
The archetypes above already reveal how a small number of characteristics — mostly out of people's control — can dramatically raise a welfare recipient's score. In reality, most people's lives are lived across a much broader cross section of variables. We therefore wanted to look at the intersection of both of our archetypes: a migrant mother. We constructed two copies of the training data, one where everyone was a migrant mother, and one where everyone was a single Dutch man.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Average Risk Score</th>
<th>Percentage High Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woman Parent Non-Dutch speaker Resident in Delfshaven Struggling Financially High estimate of labor reintegration</td>
<td>0.603</td>
<td>26.45%</td>
</tr>
<tr>
<td>Man No children Does not live in Delfshaven Dutch speaker Not struggling financially Low estimate of labor reintegration</td>
<td>0.452</td>
<td>4.97%</td>
</tr>
</tbody>
</table>

The migrant mother archetype above is more than five times as likely to be flagged for investigation as the archetype with the opposite characteristics. More than 26 percent of people in the training data where everyone is a migrant mother is above the high risk threshold. If we use the approximated real world risk threshold, which is lower than the training data threshold, nearly 50 percent of the dataset is high risk.

What's Feminist Data Protection?
Overview

1. Why data protection?
2. Feminist data protection
3. Gendered digital harms – examples
4. Challenges for enforcement & for system design
5. Conclusion
6. Reading List

Challenges for enforcement & system design

<table>
<thead>
<tr>
<th>Problem</th>
<th>Ex. 1 (High-tech abuse)</th>
<th>Ex. 2 (FemTech)</th>
<th>Ex. 3 (Targeted ads)</th>
<th>Ex. 4+5 (Autom. decisions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does transparency help?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possibility for self-protection?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Who may help?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urgency?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of system design?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Overview

1. Why data protection?
2. Feminist data protection
3. Gendered digital harms – examples
4. Challenges for enforcement & for system design
5. Conclusion
6. Reading List

Challenges for enforcement & system design

<table>
<thead>
<tr>
<th></th>
<th>Ex. 1 (High-tech abuse)</th>
<th>Ex. 2 (FemTech)</th>
<th>Ex. 3 (Targeted ads)</th>
<th>Ex. 4+5 (Autom. decisions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem</td>
<td>Harassment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty</td>
<td>Household</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does transparency help?</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possibility for self-protection?</td>
<td>+/-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Who may help?</td>
<td>Police</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urgency?</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of system design?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Challenges for enforcement & system design

<table>
<thead>
<tr>
<th></th>
<th>Ex. 1 (High-tech abuse)</th>
<th>Ex. 2 (FemTech)</th>
<th>Ex. 3 (Targeted ads)</th>
<th>Ex. 4+5 (Autom. decisions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem</td>
<td>Harassment</td>
<td>Prosecution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty</td>
<td>Household</td>
<td>Changing rules</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does transparency help?</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possibility for self-protection?</td>
<td>+/-</td>
<td>+/-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Who may help?</td>
<td>Police</td>
<td>Developers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urgency?</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of system design?</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>
Overview

1. Why data protection?
2. Feminist data protection
3. Gendered digital harms – examples
4. Challenges for enforcement & for system design
5. Conclusion
6. Reading List

Challenges for enforcement & system design

<table>
<thead>
<tr>
<th>Ex. 1 (High-tech abuse)</th>
<th>Ex. 2 (FemTech)</th>
<th>Ex. 3 (Targeted ads)</th>
<th>Ex. 4+5 (Autom. decisions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem</td>
<td>Harassment</td>
<td>Prosecution</td>
<td>Manipulation</td>
</tr>
<tr>
<td>Difficulty</td>
<td>Household</td>
<td>Changing rules</td>
<td>Ecosystem of ad networks</td>
</tr>
<tr>
<td>Does transparency help?</td>
<td>–</td>
<td>–</td>
<td>(–)</td>
</tr>
<tr>
<td>Possibility for self-protection?</td>
<td>+/–</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>Who may help?</td>
<td>Police</td>
<td>Developers</td>
<td>Data Protection Authorities</td>
</tr>
<tr>
<td>Urgency?</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Change of system design?</td>
<td></td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>
### Challenges for enforcement & system design

<table>
<thead>
<tr>
<th></th>
<th>Ex. 1 (High-tech abuse)</th>
<th>Ex. 2 (FemTech)</th>
<th>Ex. 3 (Targeted ads)</th>
<th>Ex. 4+5 (Autom. decisions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem</td>
<td>Harassment</td>
<td>Prosecution</td>
<td>Manipulation</td>
<td>Unfair treatment</td>
</tr>
<tr>
<td>Difficulty</td>
<td>Household</td>
<td>Changing rules</td>
<td>Ecosystem of ad networks</td>
<td>Biased data, reinforcement of discrimination</td>
</tr>
<tr>
<td>Does transparency help?</td>
<td>–</td>
<td>–</td>
<td>(–)</td>
<td>+/-</td>
</tr>
<tr>
<td>Possibility for self-protection?</td>
<td>+/-</td>
<td>+/-</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Who may help?</td>
<td>Police</td>
<td>Developers</td>
<td>Data Protection Authorities</td>
<td>Data Protection Authorities</td>
</tr>
<tr>
<td>Urgency?</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Change of system design?</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>
Conclusion

• Fair system design
  ▪ ... needed to overcome gendered surveillance and harms
  ▪ ... has to be actively demanded

• (Feminist) data protection
  ▪ ... focuses on informational power
  ▪ ... does not offer quick solutions
  ▪ ... best suitable for “fighters” and “endurance athletes”, not so much for single members of vulnerable groups
  ▪ ... is only one piece of the full picture.
Reading List
