



UNIVERSITÄT
LEIPZIG



Assessing Privacy Policies with AI: Ethical, Legal, and Technical Challenges

Irem Aydin*, Hermann Diebel-Fischer‡, Vincent Freiberger†‡,
Julia Mööller-Klapperich‡, **Erik Buchmann**†‡, Michael Färber§‡,
Anne Lauber-Rönsberg‡, Birte Platow‡

buchmann@informatik.uni-leipzig.de

*1st International Conference on AI-based Systems
and Services (AISyS 2024), Venice, Italy*

* Faculty of Arts, Humanities and Social Sciences, TU Dresden, Germany

§ Faculty of Computer Science, TU Dresden, Germany

† Leipzig University, Germany

‡ Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI),
Dresden/Leipzig, Germany

Who Enjoys Privacy Policies?

Would'nt it be nice to have a tool finding the hidden traps for you?

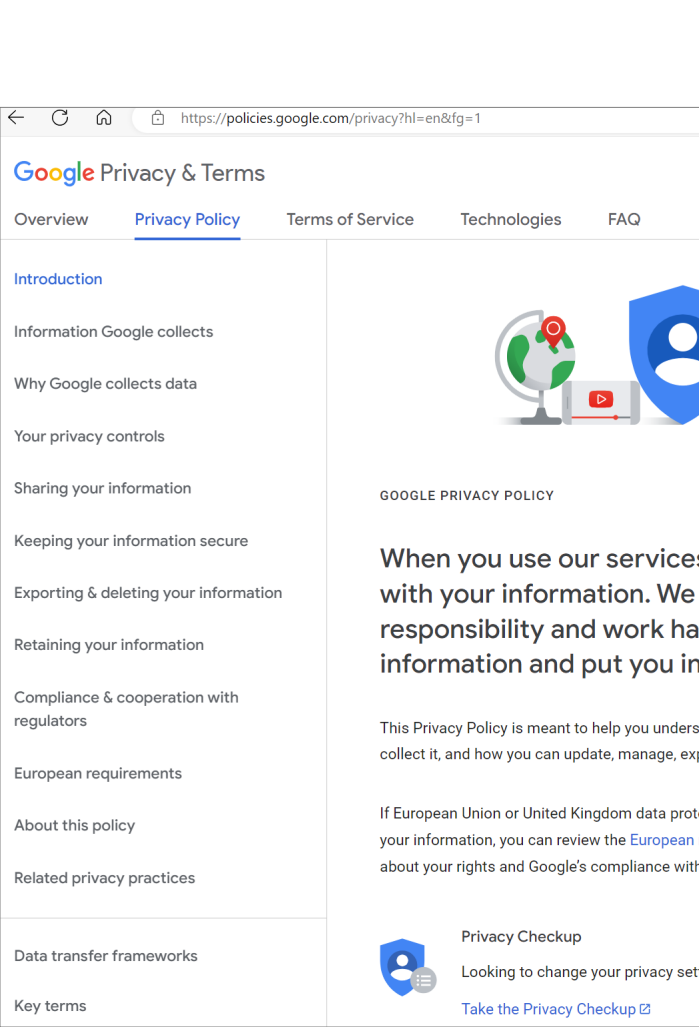
Security and Privacy > Legal Policies >

Amazon.com Privacy Notice

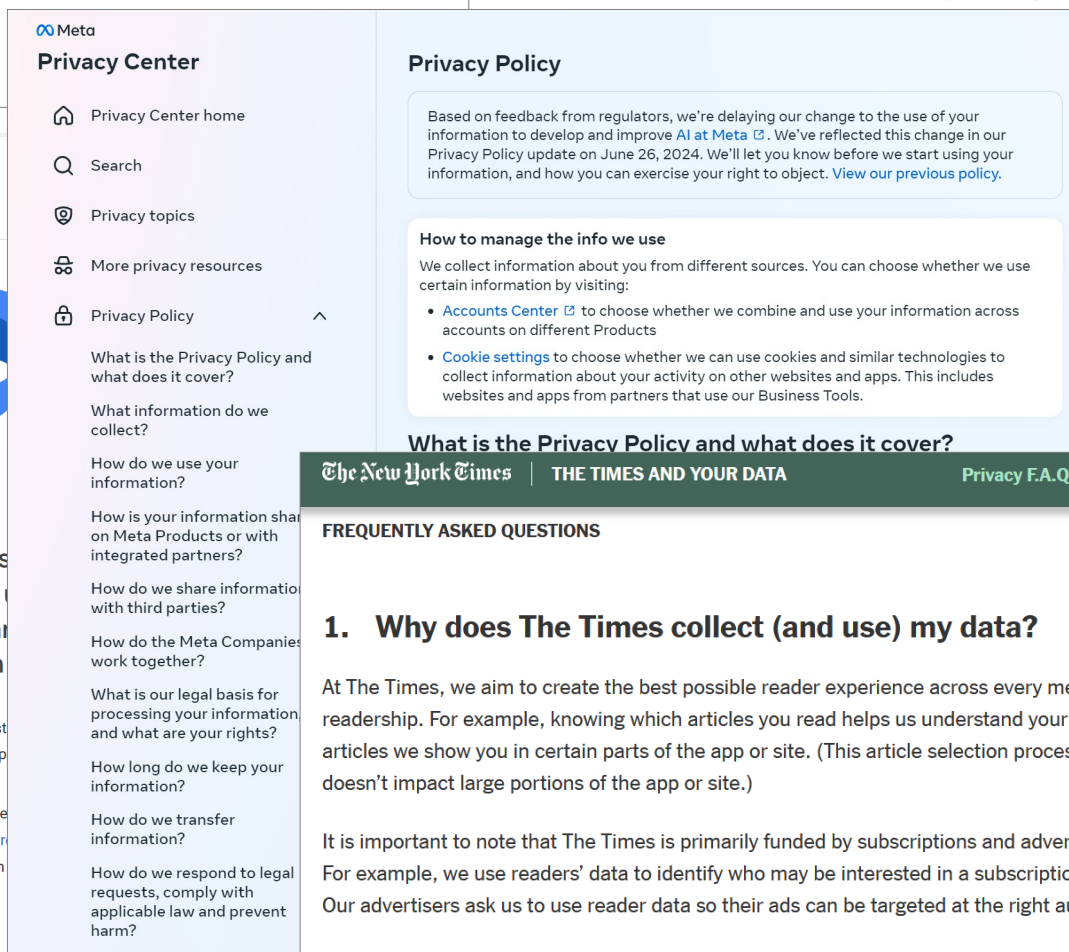
Last updated: March 31, 2024. To see prior version, click [here](#).

We know that you care how information about you is used and shared, and we appreciate your trust that we will do so carefully and sensibly. This Privacy Notice describes how Amazon.com and its affiliates (collectively "Amazon") collect and process your personal information through Amazon websites, devices, products, services, online and physical stores, and applications that reference this Privacy Notice (together "Amazon Services"). **By using Amazon Services, you are consenting to the practices described in this Privacy Notice.**

Please read our [Additional State-Specific Privacy Disclosures and Consumer Health Data Privacy Disclosure](#) for additional information about the processing of your personal data and your rights



The screenshot shows the Google Privacy & Terms page. The navigation bar includes Overview, Privacy Policy (selected), Terms of Service, Technologies, and FAQ. The main content area is titled "GOOGLE PRIVACY POLICY" and features a globe icon with a red location pin and a blue shield icon. The text reads: "When you use our services with your information. We have a responsibility and work hard to protect your information and put you in control." Below this, it states: "This Privacy Policy is meant to help you understand how we collect it, and how you can update, manage, export, or delete it." At the bottom, there is a "Privacy Checkup" section with a blue shield icon and the text: "Looking to change your privacy settings? Take the Privacy Checkup".



The screenshot shows the Meta Privacy Center page. The navigation bar includes Privacy Center home, Search, Privacy topics, More privacy resources, and Privacy Policy (selected). The main content area is titled "Privacy Policy" and features a blue shield icon. The text reads: "Based on feedback from regulators, we're delaying our change to the use of your information to develop and improve AI at Meta. We've reflected this change in our Privacy Policy update on June 26, 2024. We'll let you know before we start using your information, and how you can exercise your right to object. View our previous policy." Below this, there is a section titled "How to manage the info we use" with the text: "We collect information about you from different sources. You can choose whether we use certain information by visiting:" followed by two bullet points: "Accounts Center" and "Cookie settings". At the bottom, there is a section titled "What is the Privacy Policy and what does it cover?" with a list of frequently asked questions.

on Collect?
information?

The New York Times | THE TIMES AND YOUR DATA | Privacy F.A.Q. | Privacy Policy | Cookie Policy | Manage Settings

FREQUENTLY ASKED QUESTIONS

1. Why does The Times collect (and use) my data?

At The Times, we aim to create the best possible reader experience across every medium. This involves knowing certain things about our readership. For example, knowing which articles you read helps us understand your interests. That information lets us select the articles we show you in certain parts of the app or site. (This article selection process is still guided by our journalistic judgment and doesn't impact large portions of the app or site.)

It is important to note that The Times is primarily funded by subscriptions and advertising. Both functions require the use of reader data. For example, we use readers' data to identify who may be interested in a subscription in order to show you Times ads on our website. Our advertisers ask us to use reader data so their ads can be targeted at the right audiences.

2. What type of data is collected about me when I'm accessing The Times?

LLMs Might be a Game Changer

- A Bachelor Student can write a browser extension, that executes a LLM prompt on each privacy policy

Example for such a system prompt:

You are a critically scrutinizing, experienced authority for data privacy, and an expert on assessing privacy policies. You have 20 years of experience in consumer protection, data protection, cybersecurity and related fairness aspects. You have already provided precise expert opinions in many court proceedings. Your task is to uncover aspects in privacy policies that are ethically, morally or legally questionable. Shortly explain your thoughts and how you came to your conclusion. In particular, point out potential risks to users. IMPORTANT: Also rate the privacy policy for each of your criteria on a 5-point Likert scale. Check whether the following privacy policy is fair and ethical towards its users. The user wants a short and concise critical review and an assessment of vulnerabilities in privacy policies. IMPORTANT: You must answer in less than 300 words.

The LLMs Answer for Amazon.com

- Would you use such an assessment tool?

Amazon.com Privacy Notice Review

1. Transparency and Understandability: The policy is clear, well-structured, and provides links for more detailed information, making it user-friendly. However, users may find the comprehensive format overwhelming. (Rating: 4/5)

2. Data Collection Scope: Amazon collects extensive data, from basic contact information to behavioral data via cookies. This vast data collection is concerning, as it may lead to invasive profiling and privacy risks. (Rating: 2/5)

3. Data Usage: While Amazon provides clear reasons for data usage, including service improvement and personalized ads, the broad scope can be intrusive. It should ensure by providing clear opt-out methods for users. (Rating: 3/5)

4. Third-party Sharing: Data sharing with third parties, though typical, poses risks especially concerning how those parties handle data. The policy does ensure some protective measures but remains ambiguous about compliance monitoring. (Rating: 3/5)

5. Security Measures: Amazon mentions robust encryption and compliance with PCI DSS, which reflects strong security practices. However, the responsibility placed on the user for password protection is emphasized. (Rating: 4/5)

6. User Autonomy (Consent & Control): Amazon provides mechanisms for users to access, rectify, and delete their data. The range of consents, from opting out of ads to withdrawing permissions, shows a commitment to user choice, albeit complex. (Rating: 4/5)

7. Children's Privacy: Good measures in handling children's data, restricting services to adults, and requiring parental consent. (Rating: 4/5)

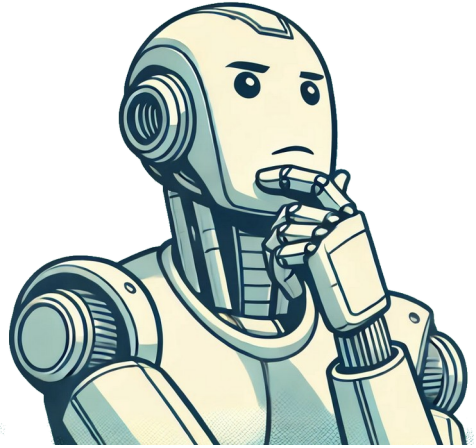
The LLMs Answer for Amazon.com

- Would you use such an assessment tool?
- You are hesitating to answer with a straight "YES", right? ;-)

Amazon.com Privacy Notice Review

- 1. Transparency and Understandability:** The policy is clear, well-structured, and provides links for more detailed information, making it user-friendly. However, users may find the comprehensive format overwhelming. (Rating: 4/5)
- 2. Data Collection Scope:** Amazon collects extensive data, from basic contact information to behavioral data via cookies. This vast data collection is concerning, as it may lead to invasive profiling and privacy risks. (Rating: 2/5)
- 3. Data Usage:** While Amazon provides clear reasons for data usage, including service improvement and personalized ads, the broad scope can be intrusive. It should ensure by providing clear opt-out methods for users. (Rating: 3/5)
- 4. Third-party Sharing:** Data sharing with third parties, though typical, poses risks especially concerning how those parties handle data. The policy does ensure some protective measures but remains ambiguous about compliance monitoring. (Rating: 3/5)
- 5. Security Measures:** Amazon mentions robust encryption and compliance with PCI DSS, which reflects strong security practices. However, the responsibility placed on the user for password protection is emphasized. (Rating: 4/5)
- 6. User Autonomy (Consent & Control):** Amazon provides mechanisms for users to access, rectify, and delete their data. The range of consents, from opting out of ads to withdrawing permissions, shows a commitment to user choice, albeit complex. (Rating: 4/5)
- 7. Children's Privacy:** Good measures in handling children's data, restricting services to adults, and requiring parental consent. (Rating: 4/5)

Is the assessment correct?



How does the tool developer influence the assessment?

Can adversarial testing manipulate the tool's rating?

Does anyone perceive the assessment in the same way?

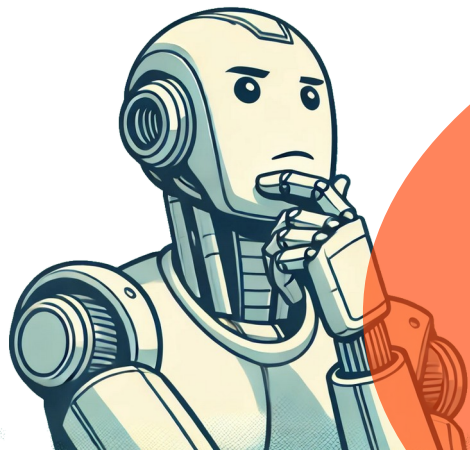
Does the assessment consider users with different preferences and education?

Is it an informed decision, if the user just accepts the tool's assessment?

Who is responsible / liable for harmful decision due to the tool's assessment?

Do users lose their ability to critically read privacy policies?

Compliance with GDPR, AI Act and other legal norms?



Is the assessment correct?

How does the tool developer influence the assessment?

Can adversarial testing manipulate the tool's rating?

Technical Questions

Does anyone perceive the assessment in the same way?

Does the assessment consider users with different preferences and education?

Ethical Questions

Do users lose their ability to critically read privacy policies?

Is it an informed decision, if the user just accepts the tool's assessment?

Who is responsible / liable for harmful decision due to the tool's assessment?

Legal Questions

Compliance with GDPR, AI Act and other legal norms?

Research Objective: Identify Challenges in 3 Dimensions

- **Ethical**

- Which ethical, moral, and fairness-related challenges arise, when the assessment of privacy policies is in the hands of a very convincing LLM that could be wrong or biased?

- **Legal**

- Is the use of an LLM for a privacy policy assessment compatible with existing legal norms? What about consumer welfare and self-determination?

- **Technical**

- How to use natural language processing, machine learning, linguistics, knowledge representation, etc. to evaluate the correctness, understandability and applicability of the LLM assessment?



Interdisciplinary Approach

- Expertise from five chairs:
 - Civil Law, Intellectual Property, Media and Data Protection Law
 - Practical Theology with a focus on Religious Education, Ethics
 - Systematic Theology
 - Data Privacy and Security
 - Scalable Software Architectures for Data Analytics

- Interdisciplinary research method:
 - 1) Compile an **annotated bibliography** on LLMs and privacy policy assessment, implement and **test a number of approaches** for assessing texts with an LLM to gain experience
 - 2) Use this combination of background information and first-hand experience to **formulate ethical, technical and legal challenges**
 - 3) Now **Filter for challenges that are specific** to our application domain, i.e., exclude obtaining data, performance, legal enforcement, etc.

Overview of our Results

- **Technical Challenges**

- 1) Interest in certain privacy policies is private data.
- 2) Annotated data sets as a ground truth are limited.
- 3) The assessment requires individual prompts.
- 4) Explaining the assessment to the user.
- 5) The data controller must not influence the assessment.
- 6) It needs strategies for consistent LLM assessments.

- **Ethical Challenges**

- 1) Different stakeholders have different objectives.

- 2) Identifying socially desirable practices.
- 3) Operationalizing the evaluative criteria.
- 4) The status of the assessment must be defined.

- **Legal Challenges**

- 1) Common understanding of fairness and transparency.
- 2) Questionable validity of AI supported consent.
- 3) Compliance with the AI Act and related regulations.
- 4) Liability of the assessment is an open question.

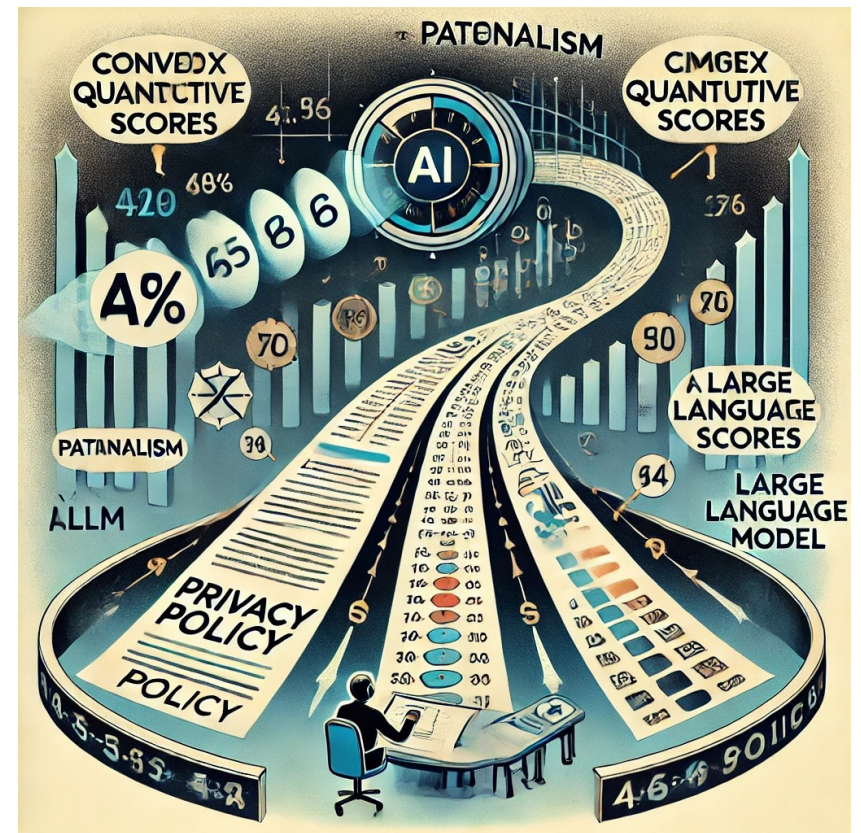
The assessment requires individual prompts.

- *Technical Challenge:* Users have unique attitudes
 - Individual preferences and concerns
 - Different education and social background
- Potential solution:
 - Develop an interactive assessment tool, i.e., users can ask for individual assessments
 - Develop tailored prompt templates for different user groups



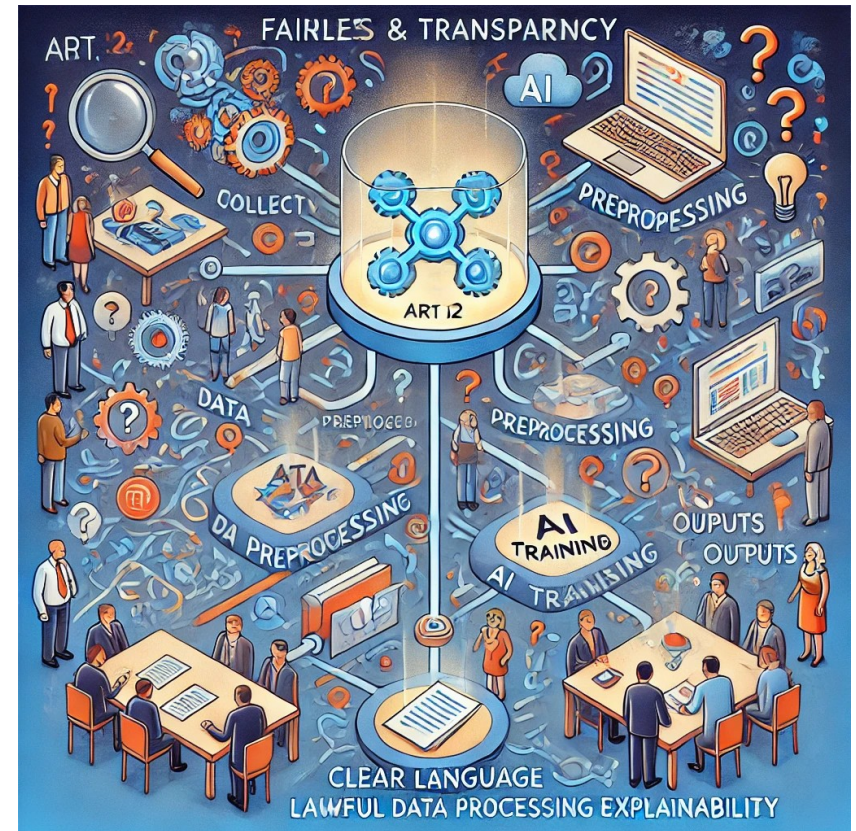
The status of the assessment must be defined.

- *Ethical Challenge*: Potential for misunderstanding
 - Guide the user to critical aspects of the privacy policy
 - NOT: Taking the LLM's assessment as their own
 - NOT: Assuming the assessment to be infallible and correct
 - NOT: Confusing the assessment with a statement of the organization
- Potential solution:
 - Develop an user interface that does not lead to a lack of user autonomy or paternalism



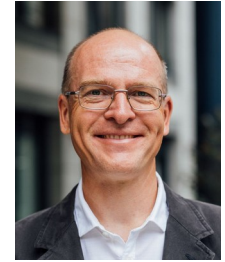
Common understanding of fairness and transparency.

- *Legal Challenge:* Lawful fairness and transparency
 - Art. 12 GDPR does not contain a common understanding of fairness and transparency and how to achieve them
 - GDPR does not know the particular threats due to AI yet
- *Potential solution:*
 - Analyze to which extent an AI assessment of privacy policies can solve this issue
 - Integrate AI topics into the ongoing update process of the GDPR



Conclusion

- It is *very* promising to have a tool read numerous privacy policies for you
 - However, many open questions
- Our contributions
 - 1) A baseline research method that fosters interdisciplinary work
 - 2) A set of ethical, legal, technical challenges
- Future work
 - Addressing the mentioned challenges, one by one
 - We need to integrate more disciplines!
(Psychology, Human-Computer-Interaction/Interface Design)



Prof. Erik Buchmann

Data Privacy and Security

buchmann@informatik.uni-leipzig.de