ANALYSIS OF

PRIVACY AND AI AS A TOOL FOR ONLINE SERVICES PERSONALIZATION

2023-2024

Made by: Arthur Vanbossuyt, Carla Pérez, Joana Esteves, Tomás Figueiredo and Yuzi Chegn

Introduction

The **primary objective of this analysis is to explore individuals' attitudes and behaviors regarding the sharing of personal data in the context of AI (Artificial Intelligence) as a tool for personalized services.** We aim to understand to what extent people are **willing to share their personal information for customization purposes** and how their **trust in companies and current regulation on data privacy a ects this willingness**. This project seeks to shed light on a relevant and current topic, thereby contributing to the understanding of the dynamics surrounding data privacy and artificial intelligence in society.

Key Goals

1. **Habits related to online privacy.**
2. **Level of concern about the privacy of personal data.**
3. **Trust in companies to manage personal data.**
4. **Trust in the use of AI as a tool for online personalization.**
5. **Sense of safety with the current measures to protect personal data.**
6. **Willingness to share data in exchange for better personalization experience.**

Dataset of the analysis

Our dataset originates from responses gathered in a previous survey, reflecting a varied demographic profile as follows:

Age: A **diverse age range** spanning from 12 to 68 years, with an average age of approximately 24.87 years, that highlights a predominant participation of younger individuals.

Gender: The participant distribution reveals that 39.2% identify as male, while a majority of 58.2% identify as female. **Relatively balanced**.

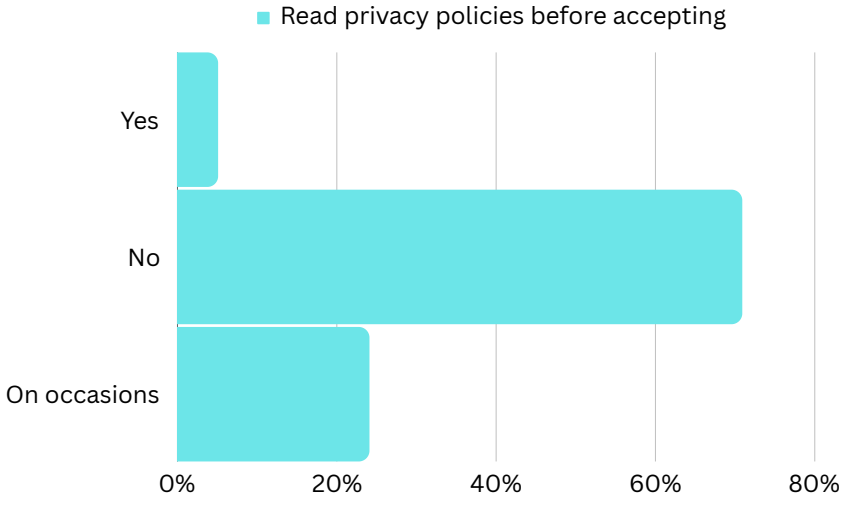
Continents: Europe is the most represented (41.8%), followed by North America (32.9%), South America (12.7%), Asia (11.4%) and Africa (1.3%). This diversity showcases a **global representation**.

Education: The **majority (83.5%) holds a university education**, 8.9% completed High school, and 3.8% completed Junior HIgh school. Notably, also includes a small amount of participants mentioning Doctorate (Ph.D.).

1 - Habits related to online privacy

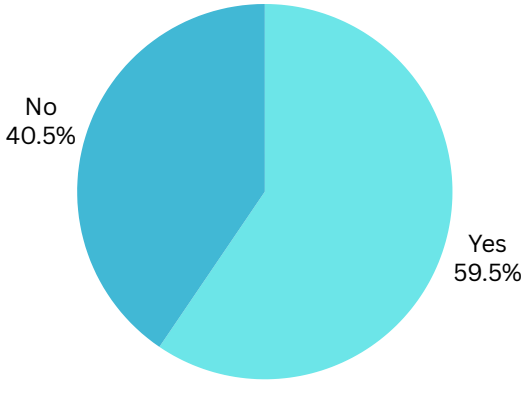
1.1. Privacy policies

A significant **majority of individuals admit to not read privacy policies**, opting instead to accept them without prior understanding of the content. Conversely, **only a very small percentage consistently engages in the practice of reading them thoroughly**. This finding emphasizes the need for companies to reconsider the accessibility and clarity of their policies, promoting user understanding and informed consent.



1.2 Measures to restrict data sharing

The **majority of the individuals surveyed do take measures to restrict their online data sharing**. This suggests **awareness** and **concern for digital privacy**. This significant percentage reflects a considerable portion of the individuals actively engaging in protective behaviors online.



Most widely used measures

1. Restrict access to navegation Cookies (46,8%).
2. Use of VPN (27,8%).
3. Privacy definition (27,8%).

These measures suggest that **users value control over their online privacy** and **seek tools that empower them to manage it e ectively**. For online services, aligning with these user preferences involves transparent communication, integration of privacy features, and continuous e orts to enhance user awareness and education.

1.3 Demographic Insights

Privacy policies - impact of age and gender

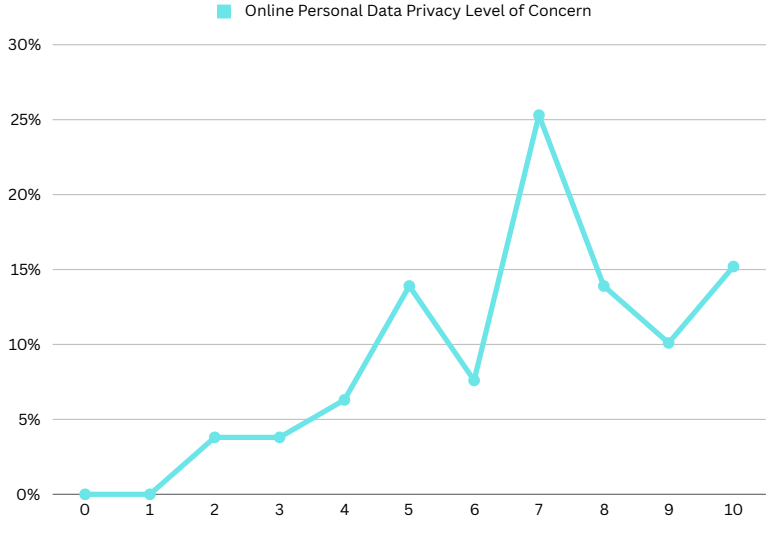
- **Younger individuals** (<30) **don't tend to read policies** or only **read them in certain occasions**.
- **Women** have a tendency to **always read the privacy policies**.
- **Men** have a tendency to **never read the privacy policies**.

Take measures to restrict data sharing - impact of gender

- **Men** tend to **take more measures** to restrict data sharing.
- **Women** tend **not to take any measures** to restrict data sharing.

2 - Level of concern about the privacy of personal data

The survey findings show that the most common level of concern is in a 7 on scale from 0-10, which represents a significant portion of the respondents. Additionally, a **majority of responses fall between the levels of 7 and 10**. This suggests that individuals are quite **mindful of privacy issues**, especially when it comes to online personalization driven by AI. The range of responses indicates that **people don't share a uniform view on privacy**. It's important to recognize and adapt to the diverse levels of user concern to create strategies that match individual preferences and comfort levels regarding online data privacy.



2.1. Justification for individuals concerns

There is some degree of correlation between people's privacy concerns and previous problems they have regarding this issue. This can indicate that for some **individuals that have concerns on data privacy, their past experience have played a role in this opinion**, and can be seen as an explanation for their worries.

For technical details on the correlation of these variables see technical annex.

2.2. Implications derived from privacy concerns regarding personal data

There is a **correlation between individuals who express higher levels of concern in relation to data privacy and those who stopped using an online service for this reason**. This is a strong indicator that people are actually worried, and not only claim to be worried, but have taken actual measures according to their concerns to protect their personal data.This serves as a negative indicator for online services, suggesting that individuals' **concerns may lead to a potential discontinuation of service usage.**

For technical details on the correlation of these variables see technical annex.

2.3. Individuals that express higher worry on data privacy Vs. online privacy habits

Individuals who express more worry usually take more privacy measures to restrict data sharing. A direct correlation exists between individuals' levels of concern about data privacy and their online privacy habits. Those who express heightened worries about their privacy tend to adopt more protective measures, indicating that their **behaviors are influenced by their concerns** about personal data security.

For technical details on the correlation of these variables see technical annex.

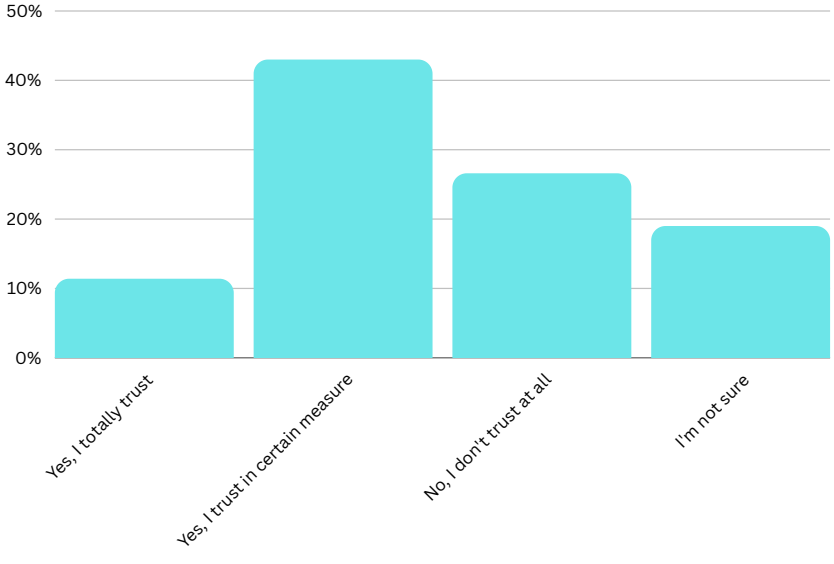
2.4 Demographic Insights

Level of concern about the privacy of personal data - impact of age

- The **older** the respondent the **biggest** the level of worry.

3 - Trust in companies to manage personal data

Overall, **individuals trust partially on companies to manage their data**, but there is a significant amount of individuals who don't trust completely. These findings highlight a nuanced landscape of trust, with a **sizable portion expressing reservations or uncertainty** in their relationship with companies regarding data handling. This underscores a clear **need for companies to enhance their practices and communication to build and maintain user trust** in handling personal information.



3.1. Main concerns

1. Share data with third parties without consent (43%).

2. Lack of transparecy (40%).

3. Online tracking (36.7%).

The survey results indicate that addressing concerns about unauthorized data sharing, transparency, and online tracking is **essential for enhancing digital privacy and user confidence**.

3.1. Trust in companies Vs. Privacy concerns

There is a significant correlation between the level of privacy concern and the level of trust in companies and platforms for data management. **The higher the privacy concerns the lowest the level of trust in companies for data management.**This correlation highlights the importance of **addressing privacy concerns to build and maintain trust** in personal data management. Companies and platforms need to **prioritize transparent and secure data practices** to align with the expectations and apprehensions of users regarding their privacy.

For technical details on the correlation of these variables see technical annex.

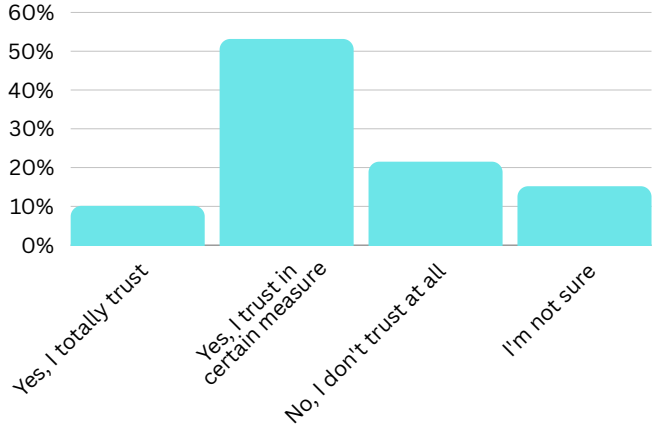
3.2 Demographic Insights

Trust in companies - impact of gender and continent

- **Women** tend to **trust partially** on companies for dealing with their data.
- **Men** tend to **trust completely** on companies for dealing with their data.
- **Europe** tends to **trust partially** on companies for dealing with their data.
- **Asia** and **South America** tend to **trust completely** on companies for dealing with their data.
- **North America** tends to **not trust at all** on companies for dealing with their data.

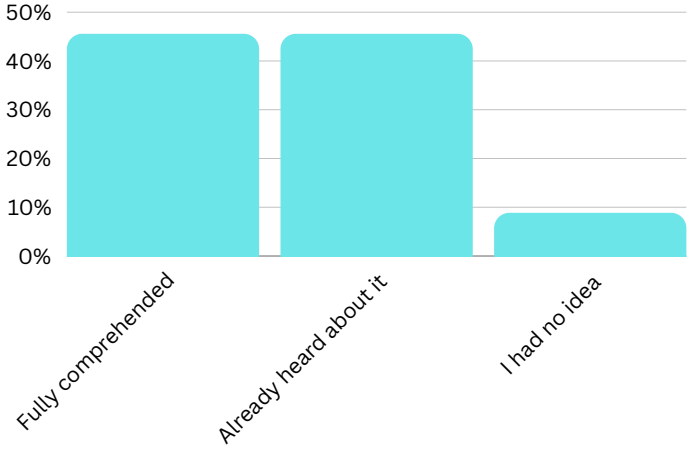
4 - Trust in use of AI as a tool for online personalization

Most of the individuals trust in the use of AI as a tool for online personalization to a certain extent. But there is still a **relevante percentage who do not trust at all or are not sure** in the use of this tool by online services. This finding emphasizes the need for **increased transparency, ethical practices, and user education** to bridge the gap and foster greater confidence among users in embracing AI technologies for personalized online experiences.



4.1. Awareness on the use of AI

Our results highlight a **notable trend where the vast majority of participants, either fully comprehend or have at least some knowledge** in the impact of AI on personalizing the online experience. This awareness suggests that a considerable portion of the public acknowledges the role of AI in managing their data. On the flip side, there is till a small percentage, who needs to be educated on AI's involvement in personalization.



4.2. Main concerns

1. Losing control of the online experience.

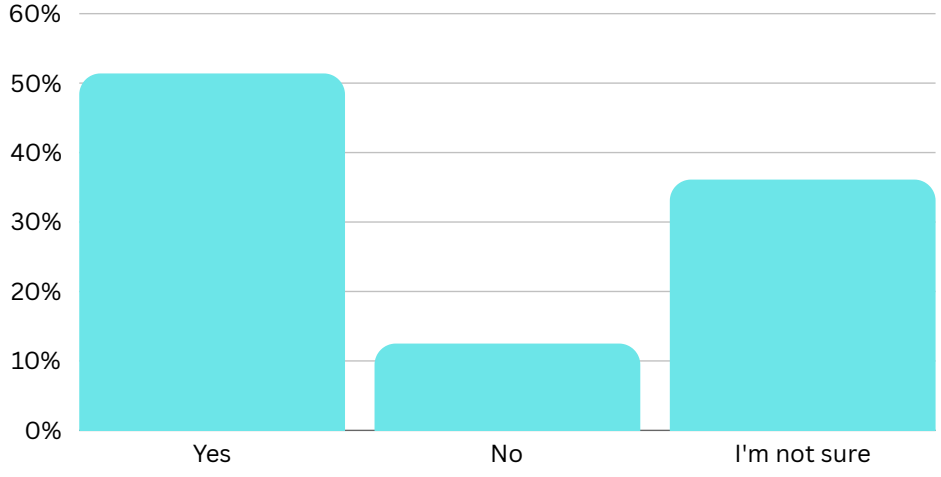
2. Failures in interaction with AI systems.

3. Misuse of data by AI.

For online services **addressing these issues is crucial to fostering a more user-friendly and secure digital landscape**, emphasizing the need for improved user control, seamless AI interactions, and stringent measures to prevent data misuse.

4.3. Usefulness of AI

The majority of individuals believe AI improves their online experience, this underscores the technology's perceived value in enhancing user satisfaction and tailoring digital interactions. On the other hand, a smaller percentage think AI doesn't enhance their experience, which suggests **some level of skepticism or dissatisfaction** with the current state of AI applications. Additionally, some are uncertain about AI's impact, indicating a **significant portion of respondents is unsure or unaware** of how AI influences their online experiences.



4.3.1. Reasons for thinking AI enhances online experience

1. Improved content (40%).
2. Understandment of preferences (40%).
3. Predictive understanding (10%).
4. Save time searching (5%).
5. Help when shopping (5%).

The findings highlight the **potential for AI to enhance user satisfaction through personalized content**, emphasizing its valuable role in shaping a more tailored and e cient online environment.

4.3.2. Reasons for thinking AI does not enhances online experience

1. Dislike or indi erence (50%).
2. Wrong suggestions (25%).
3. Rapidly changing preferences (12,5%).
4. Concerns on data privacy (12,5%).

These opinions underscore the importance of addressing user dissatisfaction, improving AI accuracy, and implementing robust privacy measures to **mitigate negative perceptions and enhance the overall online experience.**

4.4. Trust in AI Vs. Privacy concerns

We can say there is a correlation between privacy concerns and the trust in AI, **as privacy concerns increases, trust in AI decreases**. Individuals who completely trust AI have the lowest privacy concerns, while those with no trust at all have the highest concerns.

For technical details on the correlation of these variables see technical annex.

4.5. Trust in AI Vs. Trust in companies

A **robust correlation exists between trust in companies and confidence in the utilization of AI as a tool.** Typically, individuals hold similar opinions on both subjects, indicating a **cohesive perspective** among respondents.

For technical details on the correlation of these variables see technical annex.

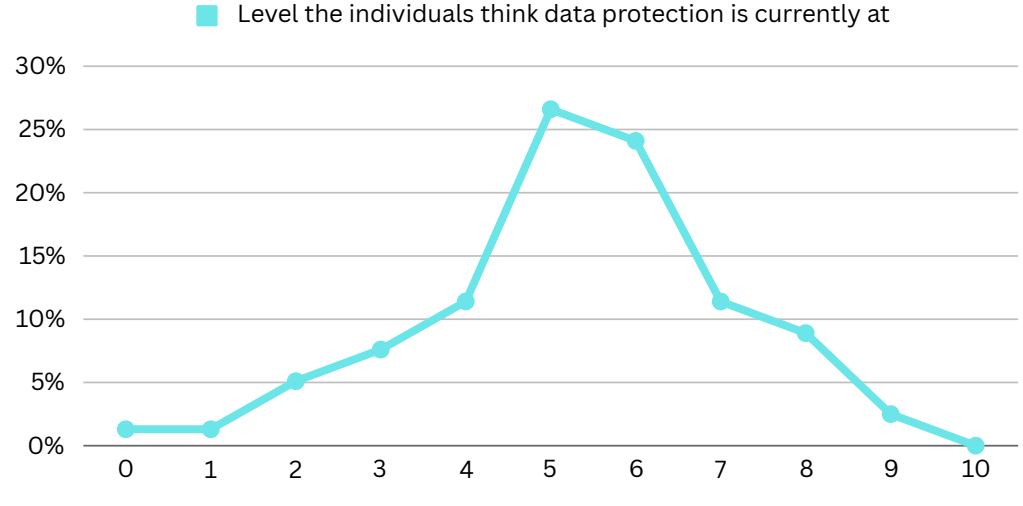
4.6 Demographic Insights

Trust in Artificial Intelligence - impact of gender and continent

- **Women** tend to **not trust at all** or **not be sure** in the use of AI as tool for personaization.
- **Men** tend to **trust completely** or **in certain mesure** in the use of AI as tool for personaization.
- **Europe** tends to **not be sure** if they trust on the use of AI as tool for personaization.
- **South America** tends to **trust completely** on the use of AI as tool for personaization.
- **Asia** tends to **trust in certain measure** on the use of AI as tool for personaization.
- **North America** is **divided** between trusting completely or not at all on the use of AI as tool for personalization

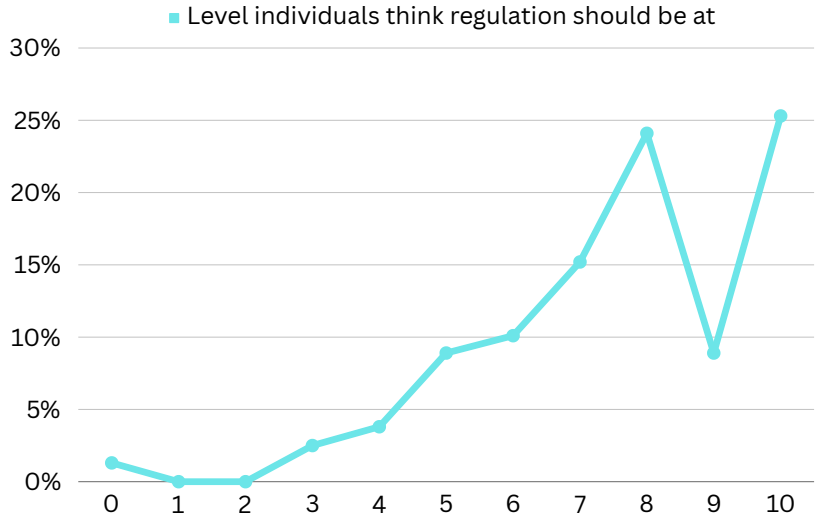
5 - Sense of safety with the current measures to protect personal data

People perceive the current data protection level as medium, indicating a moderate view overall. But there are considerable **divergent opinions**, which can be related to people's level of knowledge on the topic. The general perception is that data protection is on a good level, with a **slight inclination towards higher levels**.



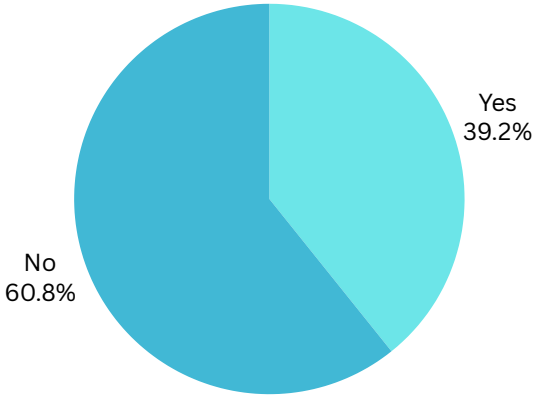
5.1. Desired Level of Regulation for Personal Data and its Manipulation in the Context of AI

The prevailing opinion among individuals is a **belief in the necessity of stringent regulations for the management of personal data**. There is **big degree of concern**, indicating a clear **preference for robust regulatory measures** in this domain. Divergent opinions suggest varying levels of apprehension, with some individuals expressing a more relaxed stance and others adopting a more serious perspective, potentially influenced by their past experiences on the subject.



5.2. Individuals knowledge on their rights related to data privacy

The **majority claim they are not informed about their rights regarding personal data.** This highlights a potential **gap in knowledge or communication regarding data protection rights and regulations**. To address this, there may be a need for increased education and awareness initiatives to empower individuals with a better understanding of their rights in the context of personal data.



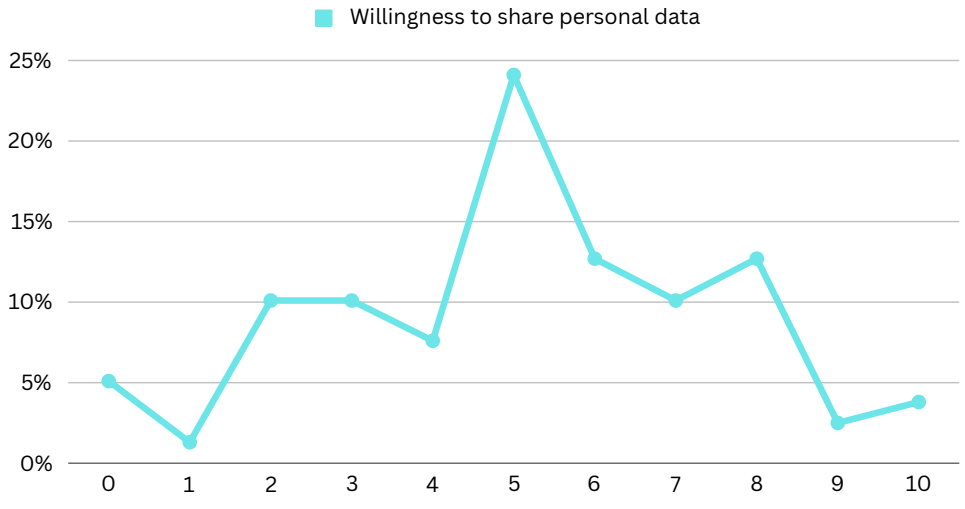
5.3 Demographic Insights

- Sense of safety with the current measures to protect personal data - impact of gender and education

- **Men** mainly think that the current protection of personal data is in a **good range of levels**.
- **Women** mostly agree that current protection of personal data is at **medium-lower level**.
- The opinion of the current level of protection of personal data seem to lower as the education increases.

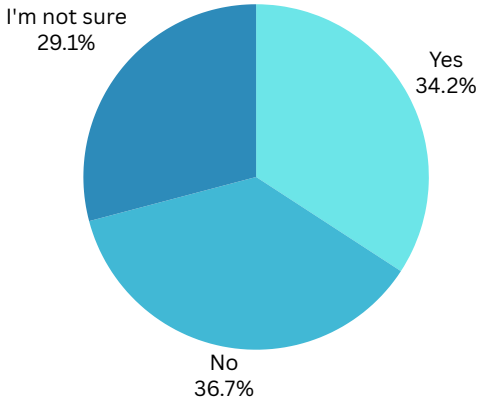
6 - Willingness to share data in exchange for a better personalization experience

People generally have a **balanced approach when it come to sharing data for personalization**, but there is a significant **discrepancy in opinions**. Furthermore, a good percentage of individuals still lean stowards lower levels of willingness, which indicates **some level of skepticism toward sharing data** even for a more personalized experience.



6.1. Individuals' opinion on personalization surpassing the potencial risks to data privacy

A **significant portion indicated that they find the benefits of a more personalized experience worth the associated risks.** Conversely, a **notable group expressed reservations**, believing that the advantages of personalization do not outweigh the potential drawbacks. A **substantial segment remained uncertain**, highlighting the **complexity and varied opinions** surrounding the perceived value of sharing personal data for personalized online experiences. There seems to be a lot of diverse opinion related to this topic.



This opinion is correlated with the amount of information they share, if the i**ndividuals agree that the experience surpasses the risks they tend to share more information**, while the contrary is also true.

For technical details on the correlation of these variables see technical annex.

6.2. Willingness to share data Vs. Trust in companies

There is notable correlation between the level of trust in companies and platforms for data management and the level of willingness to share personal data. **The higher the level of trust in these entities, the more inclined individuals are to share their data.** This highlights the role of trust in companies as a possible factor influencing individuals' openness to sharing their personal information.

For technical details on the correlation of these variables see technical annex.

6.3. Willingness to share data Vs. AI

The analysis suggests that **individuals who have higher trust in AI are more willing to share their personal data for a more personalized online experience**. Additionally, **those who perceive AI as useful exhibit a heightened propensity for data sharing.** This underscores the influence of trust and opinion in AI as a key determinant shaping individuals' readiness to share their personal information.

For technical details on the correlation of these variables see technical annex.

6.4. Willingness to share data Vs. Level of concern

The lowest the level of concern of the individuals the highest the willingness to share personal date. A notable correlation between these two variables underscores the **significant role that the level of concern plays in influencing one's willingness to share information**. Online platforms shows focus on mitigating the user's level of concern, in order to provide a better personalization.

For technical details on the correlation of these variables see technical annex.

6.5. Willingness to share data Vs. Data privacy measures and rights

Individuals exhibit a higher willingness to share data when they perceive the level of regulations to be more robust. Furthermore, **being well-informed about these regulations correlates with a greater inclination to share data for a more personalized experience**. This emphasizes the influence of individuals' opinions and awareness of regulatory measures on their data-sharing behavior.

For technical details on the correlation of these variables see technical annex.

6.6 Demographic Insights

Willingness to share data in exchange for a better personalization - impact of gender and education

- **Men** have a tendency to be **more willing** to share personal data in exchange for a more personalized experience **than women**.
- The **higher the education levels** the **lower** the level of sharing.

Conclusions

Habits related to Online Privacy:

- Users often **neglect reading privacy policies**, indicating a need for improved accessibility and clarity.
- The **majority actively takes measures to restrict data sharing**, emphasizing a significant awareness and concern for digital privacy.

Level of Concern on Privacy:

- **Substantial concern suggests users are mindful of privacy issues**, and take this topic with seriousness, impacting their behaviors regarding personal data privacy.
- **Correlation between concerns, past experiences, and actions** like discontinuing services highlights tangible impacts on user behavior.

Trust in Companies:

- Nuanced trust levels, with a **significant portion expressing partial trust and reservations**.
- Addressing **concerns related to unauthorized data sharing and lack of transparency** is crucial for building and maintaining user trust.

Trust in AI for Personalization:

- **General positive trust**, but a **notable percentage remains skeptical** or unsure.
- Increased **transparency, ethical practices, and user education are essential to bridge gaps** and foster confidence in AI technologies.

Sense of Safety with Current Data Protection Measures:

- Perceived as **moderate**, with **divergent opinions** reflecting varying levels of knowledge.
- **Majority supports stringent regulations** for personal data management, indicating a clear **preference for robust regulatory measures**.

Willingness to Share Data for Personalization:

- **Diverse opinions**, with some finding benefits worth risks and others expressing a more secure opinion.
- **Trust in companies and AI as a tool, the level of concern, and awareness of regulatory measures play crucial roles** in shaping users' willingness to share personal information.

Key insights

User needs: This analysis underscores the **need for education on data privacy** and emphasizes the importance of **user-centric approaches** in the design and communication of online services.

Strategy: Addressing **user concerns**, enhancing **transparency**, and promoting **awareness** are critical considerations for navigating the evolving relationship between AI, personalization, and privacy successfully.

Trust-Building as a Priority: Online platforms should **prioritize building trust** and **fostering an informed user base** to create a more secure and personalized digital environment.

Technical annex

Correlation Statistical tests:

2.1. Justification for individuals concerns

t-test: valor-p = 0,03

2.2. Implications derived from privacy concerns regarding personal data t-test: valor-p < 0,001

2.3. Individuals that express higher worry on data privacy Vs. online privacy habits

t-test: valor-p = 0,003

3.1. Trust in companies Vs. Privacy concerns

ANOVA: valor-p = 0,02

4.4. Trust in AI Vs. Privacy concerns

ANOVA: valor-p = 0,01

4.5. Trust in AI Vs. Trust in companies

Qui-square: valor-p < 0,001

6.1. Individuals' opinion on personalization surpassing the potencial risks to data privacy

ANOVA: valor-p < 0,001

6.2. Willingness to share data Vs. Trust in companies

ANOVA: valor-p = 0,005

6.3. Willingness to share data Vs. AI

Trust in AI: ANOVA: valor-p < 0,001

Usefulness of AI: ANOVA: valor-p < 0,001

6.4. Willingness to share data Vs. Level of concern

Pearson correlation = - 0,458

6.5. Willingness to share data Vs. Data privacy measures and rights

Opinion on current mesures: Pearson correlation = 0,427

Knowledge on data privacy rights: T-test: valor-p < 0,001