SEMINAR ON

PRIVACY AND BIG DATA

Human Aspects and Perception of Privacy in relation to Personalization

Sanchit Alekh
MSc. Software Systems Engineering
sanchit.alekh@rwth-aachen.de

Organization of this Talk

1. **Chapter 1**
   - Recollection
   - Motivation

2. **Chapter 2**
   - Attitudes and Behaviour
   - Factors affecting Privacy Perception

3. **Chapter 3**
   - Transparency and Control
   - Privacy Nudges
   - Tailor-made Privacy Decision Support

4. **Chapter 4**
   - Discussion and Conclusion
   - References

CHAPTER 1

- Recollection
- Motivation

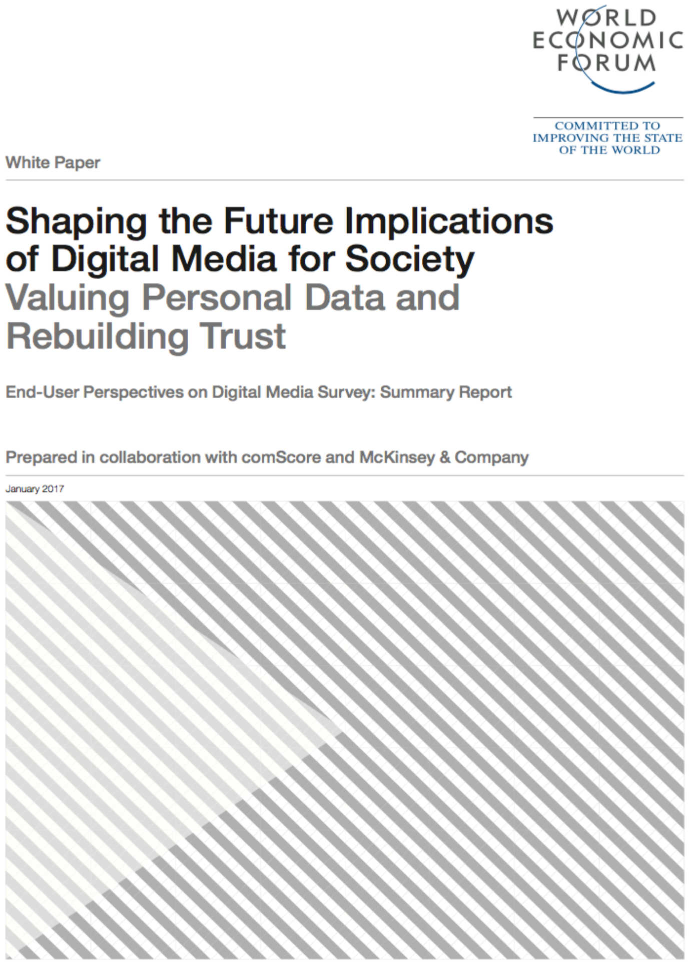
Recollection:

During the outline talks, we saw:

- Online personalization systems threaten user privacy in several ways
- User reactions towards privacy (or lack thereof) is highly context-dependent
- There is a dichotomy between privacy attitudes and actual behaviour (Privacy Paradox)
- Users indulge in a cost-benefit analysis before externalizing personal information

Motivation: Outline

- In January 2017, World Economic Forum published a white-paper
- It identifies privacy as one of the foremost global risks in the next decade
- Surveys conducted in Brazil, China, Egypt, Germany, South Africa and USA outline people’s perceptions of privacy
- 57% of the global respondents believe that privacy controls provided on websites are inadequate



WORLD ECONOMIC FORUM

COMMITTED TO IMPROVING THE STATE OF THE WORLD

White Paper

Shaping the Future Implications of Digital Media for Society

Valuing Personal Data and Rebuilding Trust

End-User Perspectives on Digital Media Survey: Summary Report

Prepared in collaboration with comScore and McKinsey & Company

January 2017

Source: http://www3.weforum.org/docs/WEF_End_User_Perspective_on_Digital_Media_Survey_Summary_2017.pdf

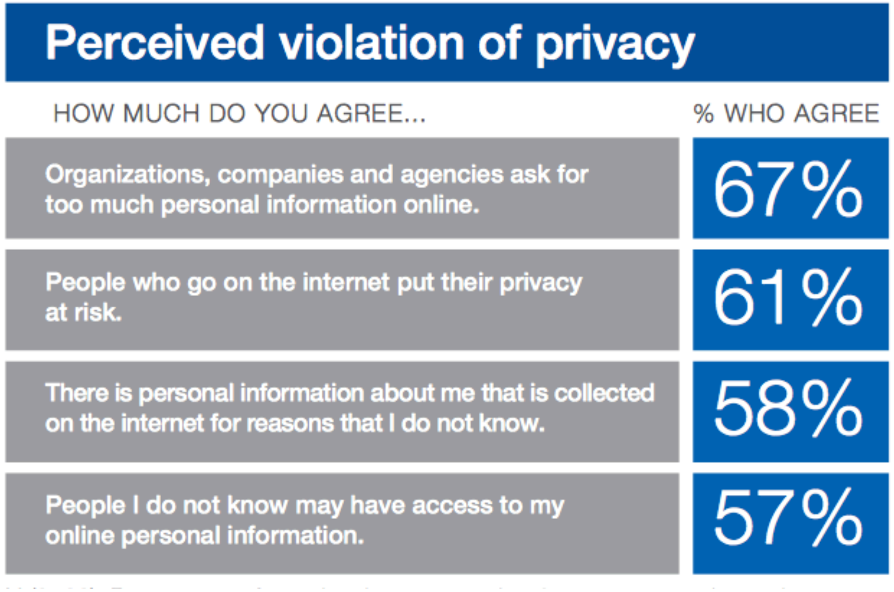
Motivation: Outline

- People are concerned about their privacy, and about how online behaviour affects their personal lives

- Need to understand privacy from a psychological and behavioural perspective, in addition to technical approaches (PETs)

- People in different countries have different expectations about data privacy

- But the primary issues are **TEA**: , Transparency, Empowerment, Accountability



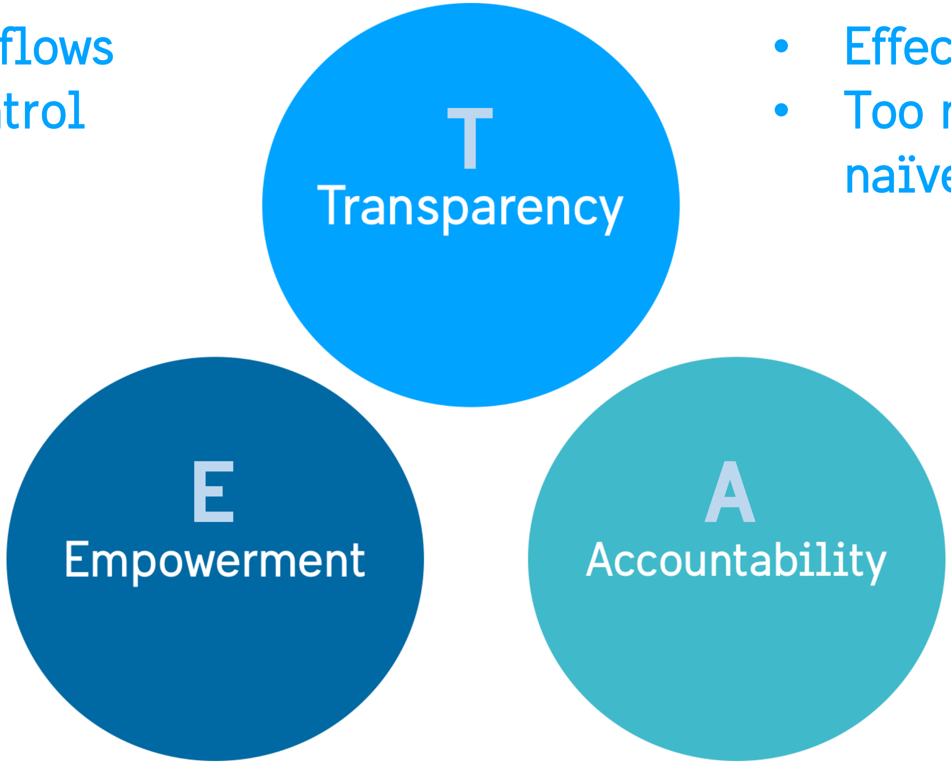
Perceived violation of privacy

HOW MUCH DO YOU AGREE...	% WHO AGREE
Organizations, companies and agencies ask for too much personal information online.	67%
People who go on the internet put their privacy at risk.	61%
There is personal information about me that is collected on the internet for reasons that I do not know.	58%
People I do not know may have access to my online personal information.	57%

Source: http://www3.weforum.org/docs/WEF_RethinkingPersonalData_ANewLens_Report_2014.pdf

Motivation: TEA Issues

- Impossible to track data flows
- Transparency without control creates tension



T
Transparency

E
Empowerment

A
Accountability

- Effective transparency is contextual
- Too much transparency overwhelms naïve users

- Lack of commercial or political incentives for empowerment
- Limited understanding of complex behaviours
- Diverse cultural and political norms

- Highly complex and dynamic data flows
- Inability to audit and enforce legal measures
- Lack of shared metrics/norms
- Inability to trace provenance and permissions

CHAPTER 2

- Attitudes and Behaviour
- Factors affecting Privacy Perception

Attitudes and Behaviours: Introduction

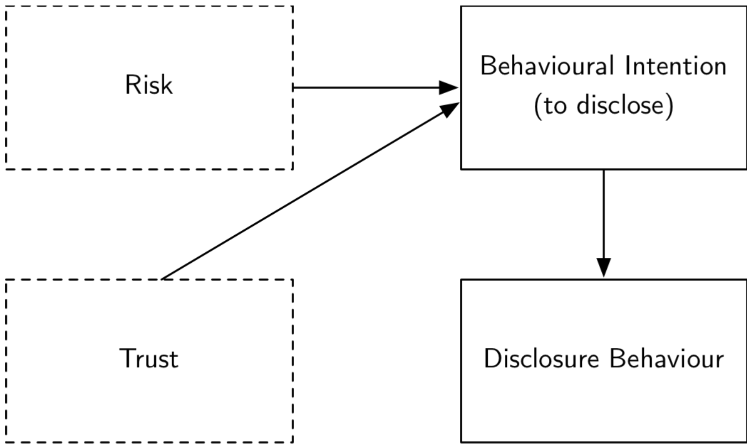
- Norberg et al. [14] were the first to introduce the concept of ‘Privacy Paradox’. They wanted to ascertain whether risk considerations also induce a user’s actual disclosure behaviour

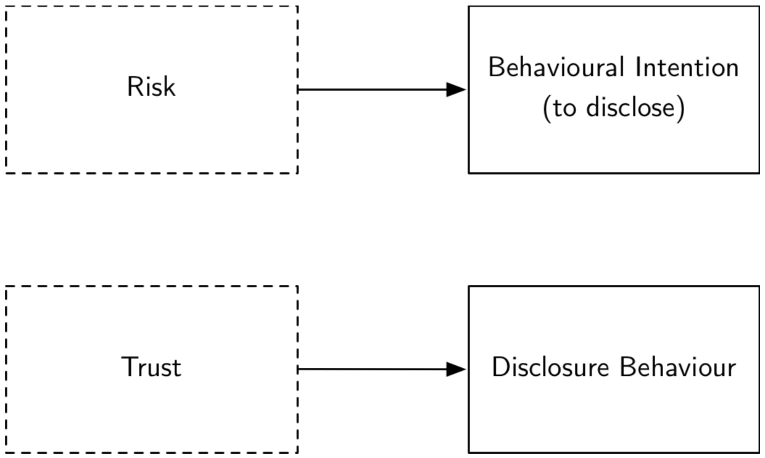
- On the basis of Privacy Attitudes, Westin et al. [22] classified people into three categories:
  - Privacy Fundamentalists
  - Pragmatists
  - Unconcerned

- Privacy attitudes and behaviours are described and explained using two concepts: **Risk** and **Trust**

Attitudes and Behaviours: Risk and Trust #1

- Historically, it was believed that risk and trust together influence behavioural intention, which in turn influences behaviour
- Privacy Paradox states that risk and trust work independently
- Risk influences intention, whereas trust influences disclosure behaviour





Attitudes and Behaviours: Risk and Trust #2

Risk is determined by:

↓

- Perceived privacy threats
- Perceived protection
- Previous experience in dealing with the company / service

Trust is determined by:

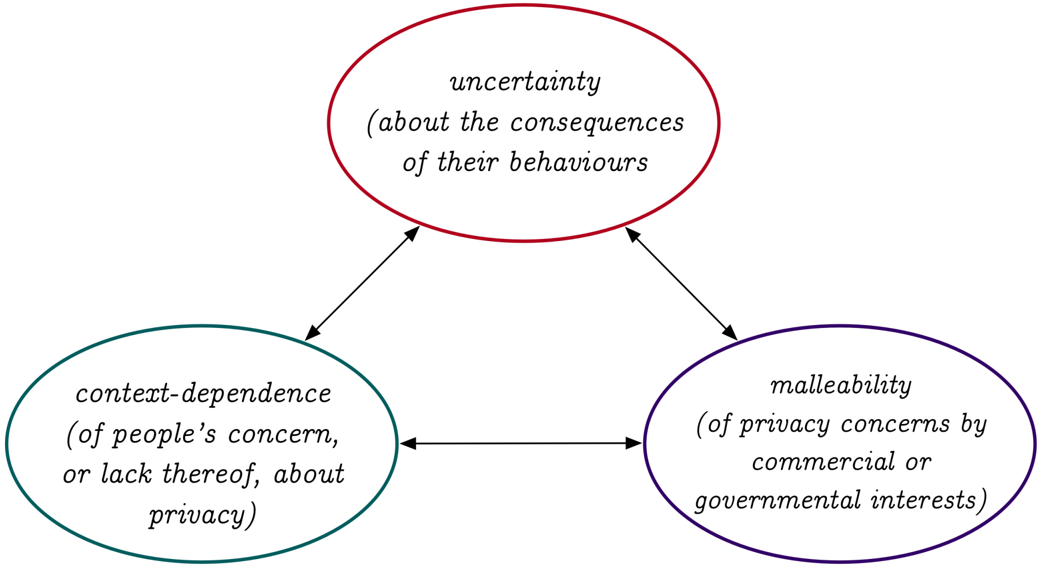
↓

- Company's reputation, brand name, status
- Privacy calculus: a trade-off between risks and benefits
- Context and relevance

Factors Affecting Privacy Perception

According to Acquisti et al. [30], the factors affecting privacy perception are:

- **Uncertainty** about consequences of privacy-related behaviours and their own preferences over them
- **Context-Dependence** of people's concern (or lack thereof) about privacy
- **Malleability** of privacy concerns by commercial or governmental interests



Factors: Uncertainty

Uncertainty: Incomplete or asymmetric information

- **Lack of clarity** about what information is being collected, how it is being used, and how it could be used in the future (e.g. chameleonic EULAs) [30]
  - Explicitly collected information: e.g. Address, e-mail ID, city
  - Implicitly collected information: e.g. Purchase history, IP address

- Uncertainty stemming from **privacy paradox** [31]

- Humans' **desire to be public**, share and disclose (Social Penetration Theory): Humans are social animals, information sharing is a central feature of human connection [32]

Factors: Context-Dependence

Context-Dependence: When people are unsure about their decisions, they take cues from the environment

- **Amount of closeness** to the person with whom information is shared (in case of recommender systems, brand confidence/familiarity) [33]

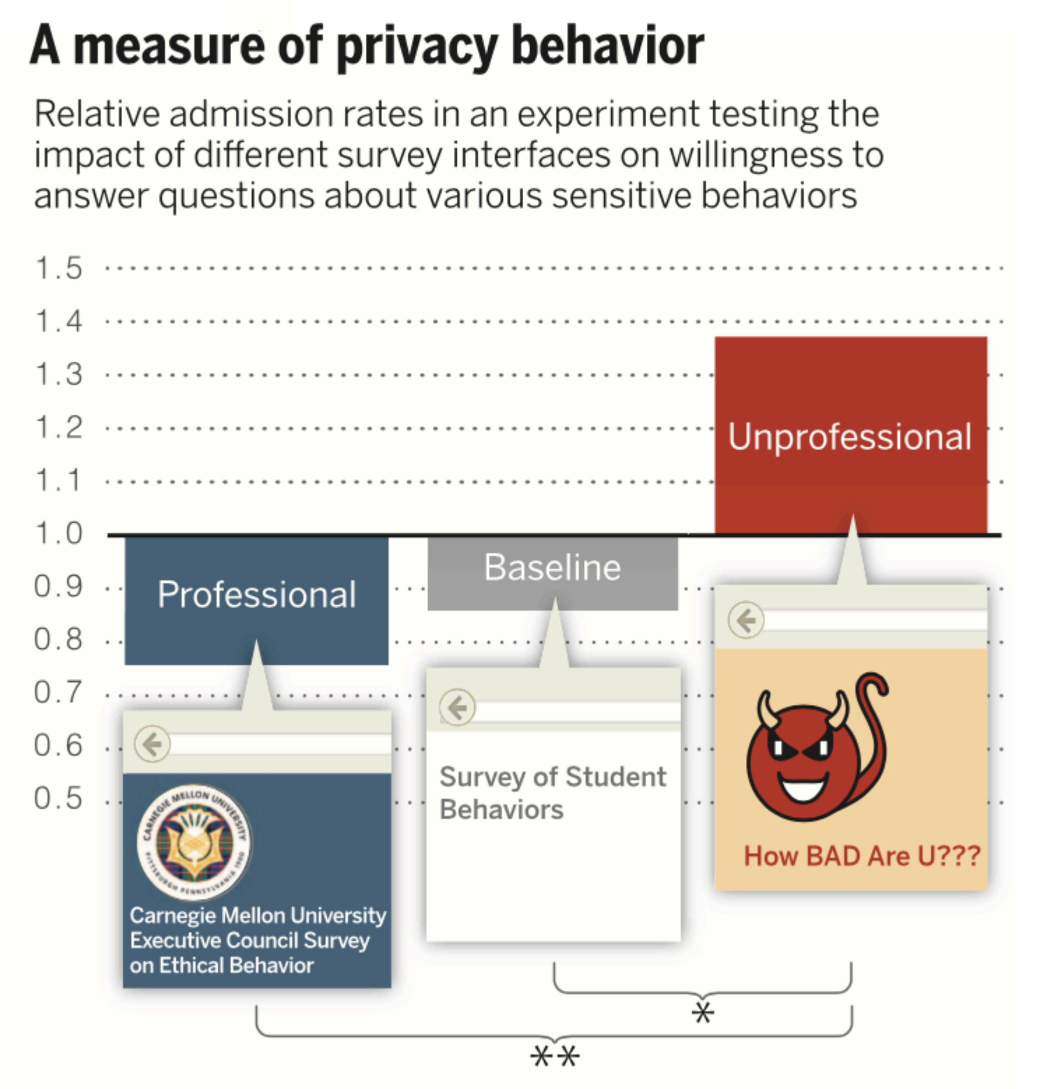
- **Government regulations** (or lack thereof) [34]

- Complexity of **user-interface design** [35]

- External **environmental factors** and ambience [36]

- Turbulence of **trust boundaries**

Factors: Context-Dependence



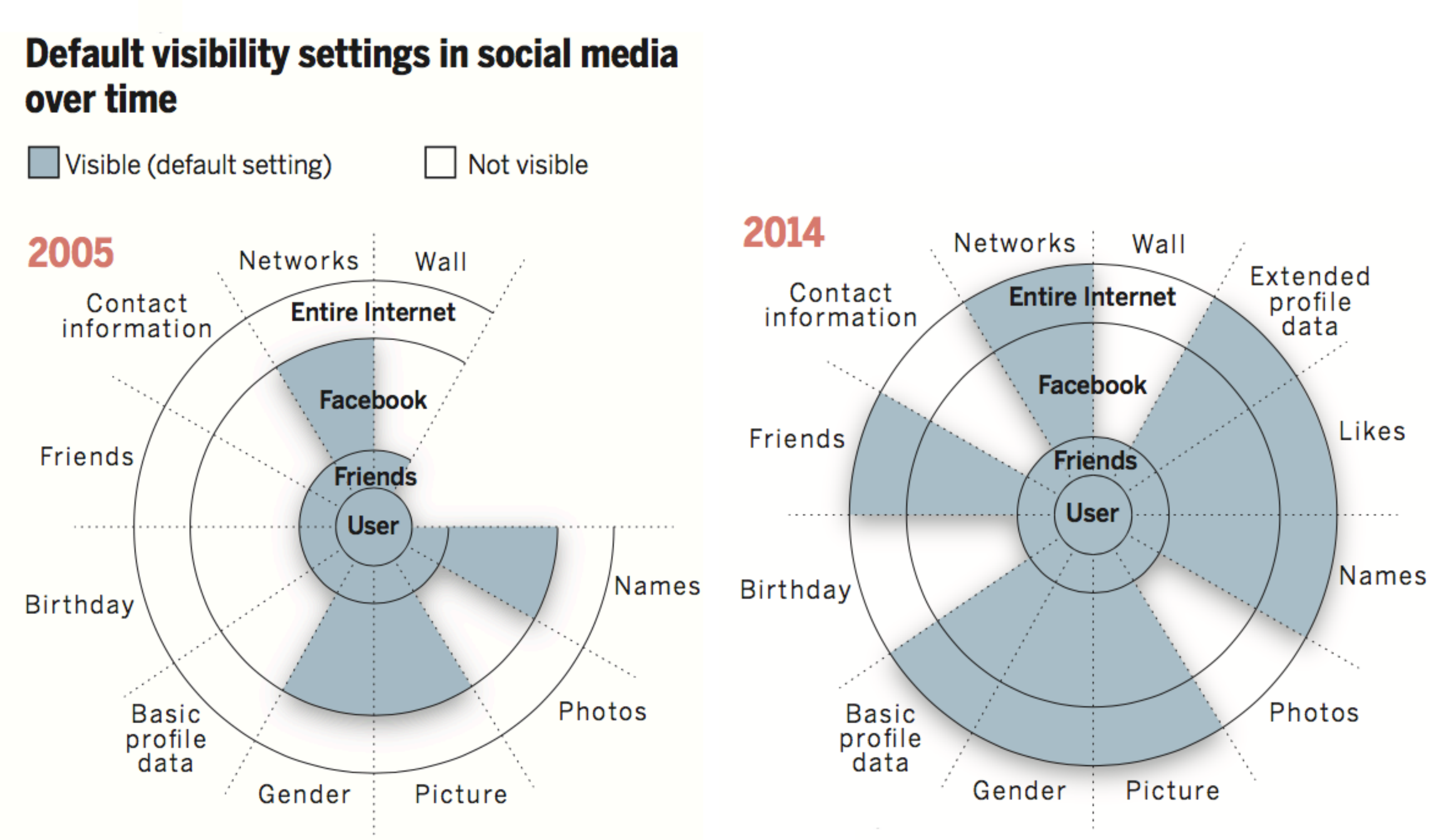
Determining the relationship between user-interface design and disclosure of sensitive information, John et al. [35]. Figure taken from Acquisti et al. [30]

Factors: Malleability

Malleability: Access to personal information is in the economic/ business interests of entities, i.e. they have a vested interest

- **Improper Defaults**: Defaults are considered by many users as implicit recommendations [7], and it is often convenient for them to let the default option be their choice
- **Malicious UI Design**: Frustrates the user to provide more information
- Merely showing a lengthy (often esoteric) **privacy policy**: gives the user a misplaced feeling of being protected, even if they don't read it
- **Not alert/warn** the user when a potentially harmful privacy setting is used

Factors: Malleability



Default visibility settings in social media over time

Default visibility settings on Facebook over the years. Figure taken from Acquisti et al. [30]

CHAPTER 3

- Transparency and Control
- Privacy Nudges
- Tailor-made Privacy Decision Support

Transparency and Control: Introduction

- Both transparency and control are fundamental to gaining a user's trust and helping them engage in a privacy calculus

- Users often want to claim full control over their privacy, but eschew the trouble of exercising this control

- Fully-transparent systems have been shown to hinder system usage rather than encourage it [39]. Only 0.2% people actually read EULAs [40]. EULAs with trust seals are counterproductive

- Keeping human psychology and behavioural factors in mind, we suggest some practices to strike a balance between transparency and control

Transparency and Control: Suggestions

- **Real Effort** to explain the most-relevant contents of the EULA, in a manner which is easy to understand and absorb, e.g. video, animation or dialogue

- Description of the privacy aspects of the system should be made available for **future reference**

- Future EULA modifications should **avoid relaxing current norms**. If unavoidable, provide an alternative to the user to opt-out

- All privacy controls **condensed into one page**, e.g. Dashboard

- Make it possible for advanced users to know about **fine-grained privacy aspects** of the system and how the collected information is used

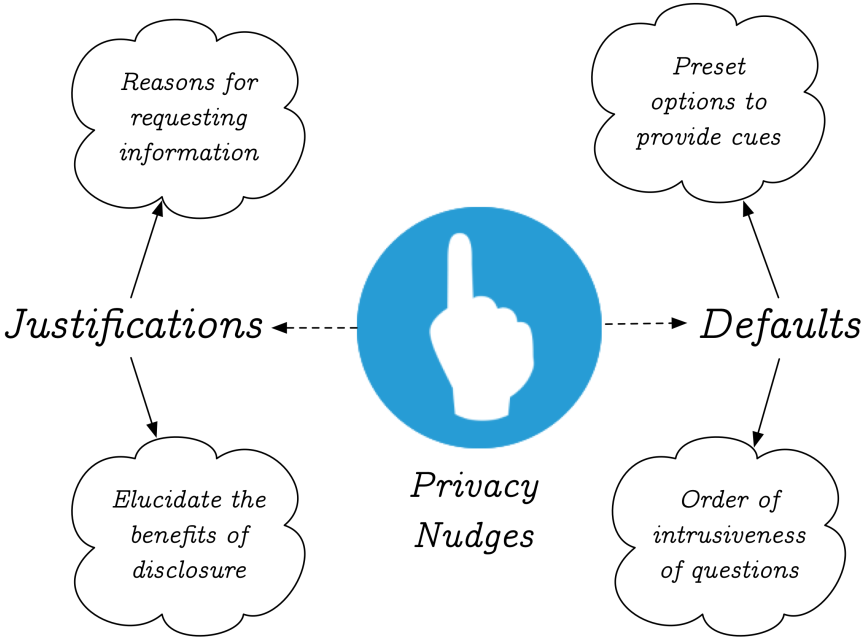
Privacy Nudges

Nudges drive the users towards making the right privacy decisions for themselves

- Reduce the regret associated with decision making [42]

- Help the user make the correct option without limiting their freedom of choice [7]

- Nudges can be of two types:
  - Justifications
  - Defaults


Reasons for requesting information
Preset options to provide cues
Justifications
Defaults
Elucidate the benefits of disclosure
Privacy Nudges
Order of intrusiveness of questions

Tailor-made Privacy Decision Support

Idea: No single privacy setting caters to all users, what is 'right' for one user might be objectionable for another

- Privacy Adaptation: customization of privacy settings, nudges and defaults by modelling behaviour in different disclosure scenarios
- According to me, 'A personalisation system to enable a human-centric personalisation system'
- Middle ground between giving full control to the user, and giving no control at all
- Understanding user's cognitive decision making increases trust

Privacy Adaptation - Approaches

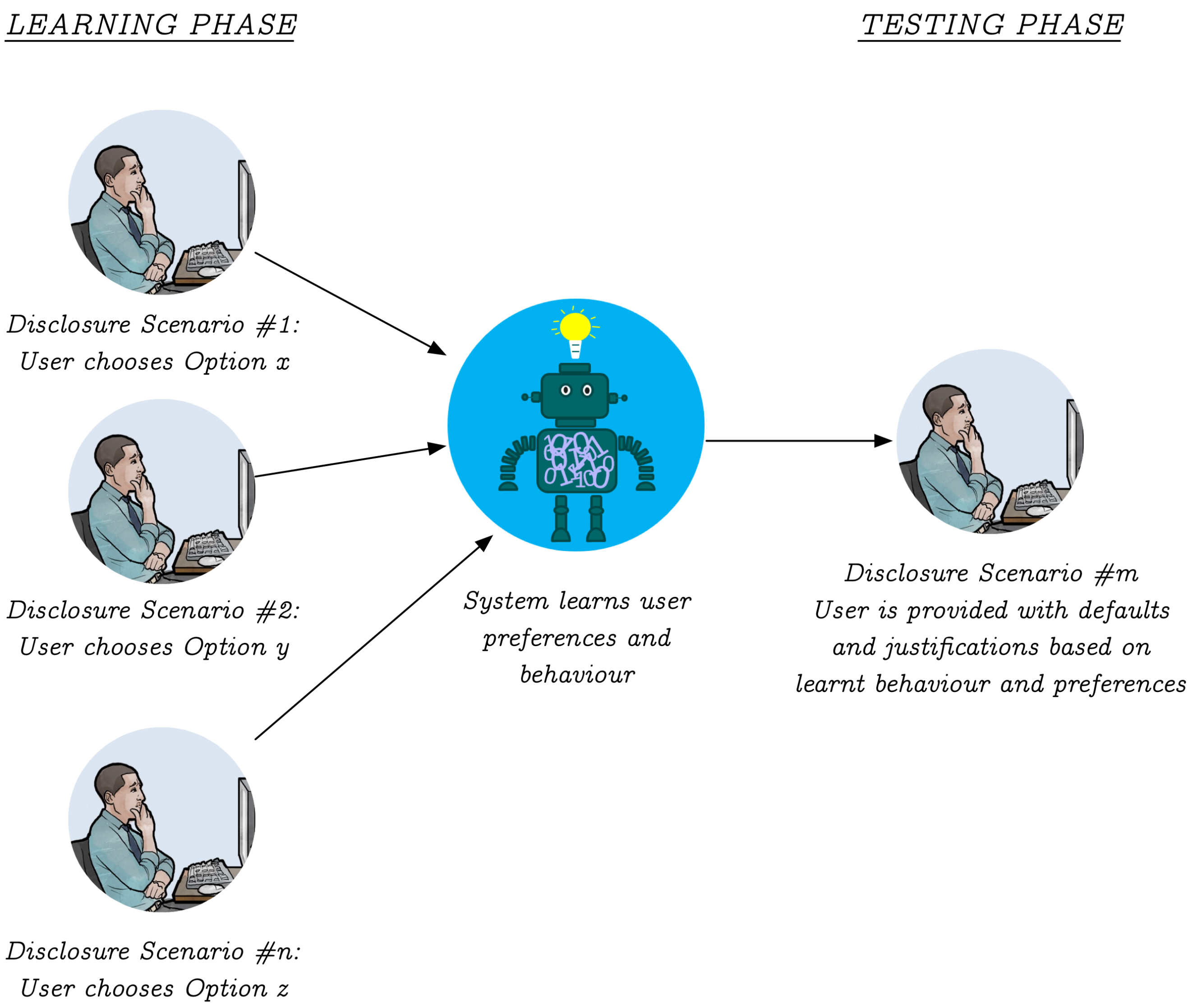
- Knijnenburg et al. [50]
  Measure information disclosure decisions and materialize them into behavioural models

- Knijnenburg et al. [49]
  Group users based on privacy preferences in several domains

- Ravichandran et al. [51]
  Clustering to aggregate users' location sharing preferences, and devise default common policies for similar individuals

- Schaub et al. [53]
  Framework for dynamic privacy adaptation in Ubiquitous Computing

Privacy Adaptation – General Framework



LEARNING PHASE

TESTING PHASE

*Disclosure Scenario #1:
User chooses Option x*

*Disclosure Scenario #2:
User chooses Option y*

*Disclosure Scenario #n:
User chooses Option z*

System learns user preferences and behaviour

*Disclosure Scenario #m
User is provided with defaults and justifications based on learnt behaviour and preferences*

CHAPTER 4

- Discussion and Conclusion
- References

Discussion and Conclusion

- Only finding technical solutions to tackle user privacy is incomplete
- Human information disclosure decisions are influenced by an amalgamation of internal and external factors
- Human-centric privacy systems must take into account the individual privacy preferences of the user
- A suitable balance between transparency and control should be found
- Privacy adaptation is a win-win for both users and system developers as it uses the user’s own disclosure preferences as a yardstick

Discussion

Interesting points raised in Peer Review

- On Personalization Systems vs Recommender Systems
- Real-life practical uses of Privacy Adaptation
- Privacy Adaptation for new users: How would that work?

On Personalization vs Recommender Systems

"Recommendation can be a form of personalization, but not all personalization is a recommendation, and vice versa. E.g., remembering my seating preferences when I book tickets online is personalizing the site to me, but has nothing to do with recommendations. Similarly, it's quite possible to recommend other videos to watch on YouTube without knowing anything about me - you can meaningfully recommend just based on what you know about general user behavior."

Sean Owen
Director
Data Science
Cloudera

"We built an engine that personalized results based on a user's individual taste profile. This, to us, was true personalization. But, shortly after we launched, we realized we had to pivot, because users disliked being put into taste profiles. So, Hoppit moved towards a recommendation engine, a search experience that suggested restaurants based on the user's context, not necessarily their personal taste. This was the big difference to me. A personalization engine learns all it can about a user's specific tastes, but it's really hard for personalization engines to take into account the nuances of the human mood and shifting emotions. However, recommendations engines can take into account user intent and user context, and make better recommendations."

Steven Dziedzic
Founder
Hoppit

References

- [7] A. Friedman, B. P. Knijnenburg, K. Vanhecke, L. Martens, and S. Berkovsky, “Privacy aspects of recommender systems,” in Recommender Systems Handbook, pp. 649–688, Springer, 2015

- [14] P. A. Norberg, D. R. Horne, and D. A. Horne, “The privacy paradox: Personal information disclosure intentions versus behaviors,” Journal of Consumer Affairs, vol. 41, no. 1, pp. 100–126, 2007

- [22] A. F. Westin et al., “The dimensions of privacy: A national opinion research survey of attitudes toward privacy.,” 1979.

- [30] A. Acquisti, L. Brandimarte, and G. Loewenstein, “Privacy and human behavior in the age of information,” Science, vol. 347, no. 6221, pp. 509–514, 2015.

- [31] E. Singer, H.-J. Hippler, and N. Schwarz, “Confidentiality assurances in surveys: Reassurance or threat?,” International journal of Public Opinion research, vol. 4, no. 3, pp. 256–268, 1992.

- [32] I. Altman and D. A. Taylor, Social penetration: The development of interpersonal relationships. Holt, Rinehart & Winston, 1973.

References

- [33] H.H.Kelleyand, J.W.Thibaut, Interpersonal relations : A theory of interdependence. John Wiley & Sons, 1978.

- [34] H. Xu, H.-H. Teo, B. C. Tan, and R. Agarwal, “The role of push-pull technology in privacy calculus: the case of location-based services,” Journal of Management Information Systems, vol. 26, no. 3, pp. 135–174, 2009.

- [35] L. K. John, A. Acquisti, and G. Loewenstein, “Strangers on a plane: Context- dependent willingness to divulge sensitive information,” Journal of consumer re- search, vol. 37, no. 5, pp. 858–873, 2010.

- [36] A. L. Chaikin, V. J. Derlega, and S. J. Miller, “Effects of room environment on self-disclosure in a counseling analogue.,” Journal of Counseling Psychology, vol. 23, no. 5, p. 479, 1976.

- [39] L. Bustos, “Best practice gone bad: 4 shocking a/b tests,” 2012.

- [40] Y. Bakos, F. Marotta-Wurgler, and D. R. Trossen, “Does anyone read the fine print,” New York University School of Law Working Paper, 2009.

References

- [42] Y. Wang, P. G. Leon, K. Scott, X. Chen, A. Acquisti, and L. F. Cranor, “Privacy nudges for social media: an exploratory facebook study,” in Proceedings of the 22nd International Conference on World Wide Web, pp. 763–770, ACM, 2013.
- [49] B. P. Knijnenburg, A. Kobsa, and H. Jin, “Dimensionality of information disclosure behavior,” International Journal of Human-Computer Studies, vol. 71, no. 12, pp. 1144–1162, 2013.
- [50] B. P. Knijnenburg, A. Kobsa, and H. Jin, “Preference-based location sharing: are more privacy options really better?,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2667–2676, ACM, 2013.
- [51] N. Sadeh, J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao, “Understanding and capturing people’s privacy policies in a mobile social networking application,” Personal and Ubiquitous Computing, vol. 13, no. 6, pp. 401–412, 2009.
- [53] F. Schaub, B. Könings, S. Dietzel, M. Weber, and F. Kargl, “Privacy context model for dynamic privacy adaptation in ubiquitous computing,” in Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pp. 752–757, ACM, 2012.