Personalized and autonomous are already everywhere; let's focus on awareness rather than trust.

Henriette Cramer
Mobile Life @ SICS, Stockholm, Sweden
henriette@mobilelifecentre.org
henriettecramer.com / mobilelifecentre.org

ABSTRACT
Personalization and autonomous adaptation to users has become a mainstream feature. The diversity is immense, and spans domains ranging from personalized online search and recommenders, to social robots. Based on the findings of a set of studies of people's responses to autonomous and adaptive systems, and current commercial developments, I highlight a number of challenges related to user trust in such systems, focusing on transparency, social strategies and social mediation effects. Rather than focusing on convincing users to use systems, a more pressing challenge is how we can increase awareness and understanding of the autonomous and user-adaptive systems already there.

Author Keywords
Personalization, autonomous, adaptive, trust, control, transparency

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Human Factors; Design; Measurement.

INTRODUCTION

Personalization is everywhere. Rather than rapidly *becoming* mainstream, systems that autonomously adapt to users *are* mainstream. Adaptation is widely ingrained in regular users' day-to-day online interactions. Every day millions of people interact with autonomous systems that decide which information they see, which ads are presented to them, the books or movies they should buy, and what application interfaces look like. Google, Facebook, Yahoo News, Amazon and a myriad of web advertisement networks personalize the information they present to site visitors – as Google's Eric Schmidt has been quoted "…it will be hard for people to watch or consume something that has not been tailored for them."

Beyond online personalization, the range of adaptive and autonomous systems is existence is incredibly diverse, ranging from spam filters and adaptive interfaces, to social robots and smart homes.

Based on the findings of a set of studies of people's responses to autonomous and adaptive systems, ranging from filters to social robots, and current developments in the commercial sphere, I will highlight a number of challenges in assessing 'trust' that deserve more attention in the community. Focus lies in transparency and awareness, and social behaviors. Of special interest are the social mediation effects these systems can have when introduced in human-to-human social situations, including online social sharing. I argue that rather than focusing on increasing trust and convincing users to use a system, a more pressing challenge is how we can increase awareness and understanding in interactions with the (networks of) autonomous and user-adaptive systems already here.

TRUST IN AUTONOMOUS SYSTEMS

Broadly speaking, trust in technology can be defined as "users' willingness to believe information from a system or make use of its capabilities" (Pasuraman & Miller, 2004), or the willingness to rely on a system despite of its risks (e.g. Muir & Moray, 1996). Trust is a multi-faceted concept. Perceptions of dependability, competence and intention play a role. Trust diminishes more easily than it (re)builds (Lee & See, 2004) and single critical errors can play a decisive role in whether or not to rely on a system (Cramer et al., 2009). We have to distinguish between the effects of system properties on trust in a system overall, effects for specific system decisions or recommendations, and delegation. The plausibility of specific system decisions, answers or recommendations - and the manner in which they are presented - will again affect whether they are followed-up upon (Cramer, 2010). The nuances above are often not captured in research, but even if they are, we are now confronted with situations that introduce much more complexities.

For online recommenders, web stores and search engines, rather than trust in 'a system', we are now dealing with situations in which users are not interacting with one website or one system. They are interacting with a complex myriad of networks, tracking systems over various apps, sites and various devices. Information is shared between a multitude of systems from different organizations, whose variety of algorithms personalize what is shown to the user – and users might not even realize they are doing so. For

Submitted to the CHI 2012 workshop on 'End-user interactions with intelligent and autonomous systems'

CHI'12, May 5–10, 2012, Austin, Texas, USA.

embodied agents, robots are being designed that aim to build companionship bonds with their users, and a 'break up' with a system possibly now has the potential to be quite an emotional ordeal. A wide range of autonomous systems now mediate in human social situations, making trust more than an user-to-system affair.

TRANSPARENCY AND CONTROL

Understanding of adaptive and autonomous systems' behavior is notoriously difficult. While explanations can help users' sense of control and trust, they can also lead to unpredictable behavior and not all explanations are created equal (Cramer et al., 2008; McNee et al., 2003). As the adaptive systems adapt themselves to the user, users in turn also adapt their behavior to their perceptions of the system's inner workings. Users will for example tailor their feedback in a manner that they think will yield them the best results, but in actuality only decreases performance. Allowing users control will also not per se increase efficacy of a system. Allowing users access to their profiles for example, does not necessarily mean they are able to understand or improve them (Dzindolet et al., 2003).

While transparency and understanding sometimes are assumed to be the basis for increased trust, increasing user understanding is not always in the best interests of someone who wants to promote usage of a system. Even when trust is warranted, and bad intentions aside, acceptance of a system is not always improved by making the system more transparent. If a system's criteria for its decisions are transparent, it is also easier for users to detect incongruencies with the way they think a decision should be made (Cramer, 2010). However, making a system more transparent can at least increase the chances that user perceptions of competence and actual competence (as far as it can be assessed) of a system match.

Before people can actually be facilitated in controlling the systems they are interacting with, the first challenge is actually being aware that you are interacting with an adaptive system. Users may not even realize that the information they are receiving is actually filtered - let alone how this may be achieved. Especially when user behavior is tracked across different systems and different websites, it becomes very difficult to understand what information is used to adapt, which behaviors or preferences are stored, and what the consequences of behavior at one site will be on another.

System designers need to pay attention to ensuring awareness about systems' activity and adaptivity. A full understanding of a system's inner workings might not be feasible, or even necessary. However, often how the most basic user understanding and basic control is not apparent: am I using a personalization system, and is it even possible to 'turn it off'?

SOCIAL STRATEGIES & PERSUASION

Considerable research efforts focus on making autonomous systems such as robots 'team players' (e.g. Klein et al., 2004) or social companions, highlighting expected social behaviors and bonding.

Social behaviors can affect the perceptions of users and their attitudes towards systems. We have seen that personality traits such as locus of control, empathic tendencies and attitudes toward a particular type of technology, and affect responses to a system. Reacting in line with the user's situational expectations and affective experience is important. Dependability, credibility and closeness decrease when a system shows behavior incongruent with the situation. However, we have seen that users sometimes still attribute an agent greater empathic ability when it, even when unwarranted, remains optimistic (Cramer, 2010). Whether intentionality is attributed to a system and whether perceived goals behind system's requests coincide with the user's current goals are additional factors that could affect whether expressiveness is experienced as a `sincere' positive feature of user-system dialogues.

Reactions to social system behaviors also depend, at least partially, on whether a system's interaction style matches the user's personality. A system's persuasive qualities can for example be enhanced by personalizing persuasion techniques and which social behaviors are helpful and 'acceptable' is up for debate (Eckles & Kaptein, 2010).

Social mediation & Responsibility

Beyond social strategies in interacting with users, research on social implications of adaptive and autonomous agents is yet scarce. The way in which autonomous and user-adaptive systems currently mediate in social situations remains under-discussed.

As Periser (2011) and Zuckerman (2011) point out, filters now decide which stories are important, and different users can have widely different experiences while using the same system. Rather than human gatekeepers as in the past, autonomous algorithms now decide which news stories we see; we could be living in a filter bubble (Periser, 2011), in which we will not see perspectives that do not fit our profile. This introduces questions of civic responsibility in what information should be shown to users: what they are likely to consume, or what 'is important' (but who then gets to decide what is important?). Other unforeseen consequences of user adaptation are often left undiscussed, what for example happens when security is breached and user profiles, activity and preferences are available to others?

Adaptive and semi-autonomous agents will also affect social relations in a more direct way. People recommenders already arrange social contact on a large scale (e.g. Facebook.com's friend and activity suggestions). While Foucault et al. (2007) showed that gossiping agents can

have a positive effect in interpersonal relationships, less positive effects are certainly not unimaginable. Especially when designing long-term 'companions' such as social robots that interact with multiple people, sometimes information should be 'forgotten' (Vargas et al., 2009) and difficult questions arise on what can be told to other users, and what cannot.

From a more traditional point of view, system mistakes can be seen as user error, or as a design flaw putting the blame on the system developer. For autonomous and adaptive systems this is much less clear. Where responsibility lies when a system makes a mistake when users have 'trained' a system is not a simple question, especially when they might have understood the system differently than designed; or are unaware they are actually 'training' it. In Cramer et al., (2009), for example not all users fully assigned the responsibility for spam filter errors on their filters alone. For instance, a participant who decided to fully rely on his filter mentioned false positives resulting from email subjects being 'vague' or 'too spam-like'. This may imply that some people feel that while filter performance is important, senders should also take into account systems possibly operating on the receiving end of their emails. While this may appear a mundane issue, this has actually been quite a shift; instead of systems adapting to people's emails, a new responsibility was put on email users to figure out, and adapt to, system criteria when communicating with others.

Gmail's Google Chat also for example automatically – without any user involvement or request - decides who to show as online contacts based on email contact, ignoring perceived differences between emailing, online chats and seen as available and online. In an earlier case involving Google Buzz such automatic social sharing led to a widely circulated report of a woman whose reportedly abusive ex-husband got access to information about her online activity after Buzz algorithms identified him as someone she often received emails from, raising considerable safety and privacy concerns (TechCrunch, 2010). Too often, users are left out of the loop and the decision on what should be socially shared has been made for them before they even notice. Now the responsibility is put on them to find out that these things are happening, how they work, and to actively opt-out, if at all possible.

CONCLUDING THOUGHTS

In devising evaluation methods for autonomous and adaptive systems we should not only make sure that 'end-users positively evaluate the systems', but that we also consider their awareness of system activity and potential social implications, including responsibility in offering users control, and not relying on automatic gatekeepers alone.

Making a system trustworthy to end-users should go beyond their perception that the system is trustworthy. Trust should be based on awareness of activities of systems, and the actors behind them, and what their information is used for. A basic understanding of what an adaptive system does, and leaves out, is arguably more important than user satisfaction per se. But how do we achieve such transparency if user understanding is notoriously difficult to achieve; when current personalization systems are not just 'one system', they are complex networks of systems that collect information from a wide range of sources; and when a full understanding of all intricacies of a machine learning algorithm is impossible?

Arguably, the biggest challenges in trust lie in the systems that already present millions of users with personalized of available information and current events. Users often do not have the option to 'turn it off'. Rather than focusing on increasing trust, we should focus on giving end-users control over their data, and to increase awareness in both users and developers of systems' potential social implications.

AUTHOR BACKGROUND

Henriette Cramer is a senior researcher at the Swedish Institute of Computer Science (SICS) and the Mobile Life Centre in Stockholm, Sweden. Her research focuses on user studies and concept development in the area of mobile location-based experiences, location-sharing, 'Research in the Large' methodology for using widely distributed apps in research, and users' interaction with autonomous 'things'. She currently also coordinates SICS' activities within a human-robot interaction project.

Henriette's PhD-thesis at the University of Amsterdam (2010) revolved around people's interaction with autonomous and adaptive systems, ranging from spam filters and recommenders to social robots. Focus was on the effects of both transparency features and social strategies on (perceived) user understanding, control and trust.

REFERENCES

1. Cramer, H. People's responses to autonomous and adaptive systems, PhD-thesis, University of Amsterdam, The Netherlands, 2010.
2. H. Cramer, V. Evers, M. van Someren, S. Ramlal, L. Rutledge, N. Stash, L. Aroyo, and B. Wielinga. The effects of transparency on trust and acceptance in interaction with a content-based art recommender. User Modeling and User-Adapted Interaction, 18(5), 2008.
3. H. Cramer, V. Evers, M. van Someren, and B. Wielinga. Awareness, training and trust in interaction with adaptive spam filters. CHI'09, 909–912, 2009.
4. M. Dzindolet, S. Peterson, R. Pomranky, L. Pierce, and H. Beck. The role of trust in automation reliance. Int J. of Human-Computer Studies, 58(6): 697–718, 2003.
5. Eckles, D. & Kaptein, M.C. Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling , Persuasive'10, 2010.

6. Foucault, B. and Mentis, H.M. and Sengers, P. and Welles, D.Provoking sociability, CHI'07, 2007.
7. G. Klein, D. Woods, J. Bradshaw, R. Hoffman, and P. Feltovich. Ten challenges for making automation a 'team player' in joint human-agent activity. IEEE Intelligent Systems, 19(6):91–95, 2004.
8. B. Muir and N. Moray. Trust in automation. part II. Ergonomics, 39:429–460, 1996.
9. J. Lee and K. See. Trust in automation: designing for appropriate reliance. Human Factors, 42(1), 2004.
10. S. McNee, S. Lam, C. Guetzlaff, J. Konstan, and J. Riedl. Confidence metrics and displays in recommender systems. Interact'03, 176–183, 2003.
11. R. Parasuraman and C. Miller. Trust and etiquette in high-criticality automated systems. Communications of the ACM, 47(4):51–55, 2004.
12. Pariser, E. (2011) The Filter Bubble, Penguin Press, NY.
13. TechCrunch, Google Buzz Privacy Issues Have Real Life Implications, 12 Feb 2010. http://techcrunch.com/2010/02/12/google-buzz-privacy/
14. Vargas, P. et al., Advocating an ethical memory model for artificial companions from a human-centred perspective, AI & Society, 2011.
15. Zuckerman, E. (2011) Desperately Seeking Serendipity, CHI'11 keynote.