Lonely road: speculative challenges for a social media robot aimed to reduce driver loneliness

Felipe Valle, Alexander Galozy, Awais Ashfaq, Kobra Etminani, Alexey Vinel, Martin Cooney

Center for Applied Intelligent Systems Research (CAISR)
School of Information Technology, Halmstad University
301 18 Halmstad, Sweden
martin.daniel.cooney@gmail.com

Abstract

Driver monitoring is expected to contribute greatly to safety in nascent smart cities, also in complex, mixed-traffic scenarios with autonomous vehicles (AVs), vulnerable road users (VRUs), and manually driven vehicles. Until now, one focus has been on detecting bio signals during the relatively short time when a person is inside a vehicle; but, life outside of the vehicle can also affect driving. For example, loneliness, depression, and sleep-deprivation, which might be difficult to detect in time, can increase the risk of accidents–raising possibilities for new and alternative intervention strategies. Thus, the current conceptual paper explores one idea for how continuous care could be provided to improve drivers' mental states; in particular, the idea of a "robot" that could positively affect a driver's health through interactions supported by social media mining on Facebook. A speculative design approach is used to present some potential challenges and solutions in regard to a robot's interaction strategy, user modeling, and ethics. For example, to address how to generate appropriate robot activities and mitigate the risk of damage to the driver, a hybrid neuro-symbolic recognition strategy leveraging stereotypical and self-disclosed information is described. Thereby, the aim of this conceptual paper is to navigate through some "memories" of one possible future, toward stimulating ideation and discussion within the increasingly vital area of safety in smart cities.

Introduction

In nascent smart cities of the near future, we envision complex and potentially chaotic mixed-traffic interactions between autonomous vehicles (AVs), vulnerable road users (VRUs), and manually driven vehicles. To help enable safety in such general urban scenarios, the four year SafeSmart research project seeks to develop approaches to automate communication, localization, control, and testing, that also consider the human factor toward reducing the risk of accidents. For example, drivers who are depressed or sleep-deprived have accidents more often (Garbarino et al. 2017; Rosso, Montomoli, and Candura 2016; Meuleners et al. 2015).

Some methods for inferring health state have been proposed, which mostly focus on times when a driver or passenger is in a vehicle, sometimes also in connection with surveys and health records. For example, eye openness, facial expressions, and micro-nods, as well as data from wearables such as smart clothes, can be analyzed in the cloud, to send alarms to medical staff and family members, commands to a vehicle to change to autonomous driving, and prompts or music (Chen et al. 2018). However, this data might not be enough to accurately detect depression or sleepiness. For example, depressed persons do not always appear sad, but can experience manic phases or hide their emotions, and use of wearables tends to drop over time. Overriding a driver's control of a vehicle raises ethical concerns about risks of false alarms and responsibility in the case of accidents, and such detection might occur too late for successful intervention. Also, manual health checks and surveys are informative but tend to be conducted rarely due to time constraints and the paucity of human healthcare providers (a doctor's appointment might be only once a year or less). Therefore, *continuous monitoring and care* could be effective.

One potential target could be *loneliness*, "the distressing feeling that one's desired social needs are not being met", which is a key factor influencing both depression and sleep-deprivation (Benson et al. 2021). Various new technologies could be used to combat loneliness in drivers, who might be young or old, new or experienced professionals. One technology that has been proposed to reduce loneliness through social interactions is the "companion robot"; however, physical robots are not yet popular in human environments, due to current limitations in capabilities and cost. By contrast, approximately 4 billion people use social media, many leaving a daily digital footprint, suggesting its usefulness as a way to get data and interact.[1] Such social media mining has been successfully used, e.g., to gain insight on traffic safety upon detecting anomalous traffic (Pan et al. 2013).

But, how can a person reduce their loneliness via social media? Just being on social media is not enough to reduce loneliness: similar to the phenomenon of urban loneliness, people can find themselves alone in a crowd of others, and comparison with others can result in disappointment. However, whereas passive use can conversely increase loneliness, interactive use of social networking services (SNS) has been connected to lower loneliness (Yang 2016). This benefit can come from various forms of social activities, like vi-

[1] https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/ – Accessed: 2021-05-04.

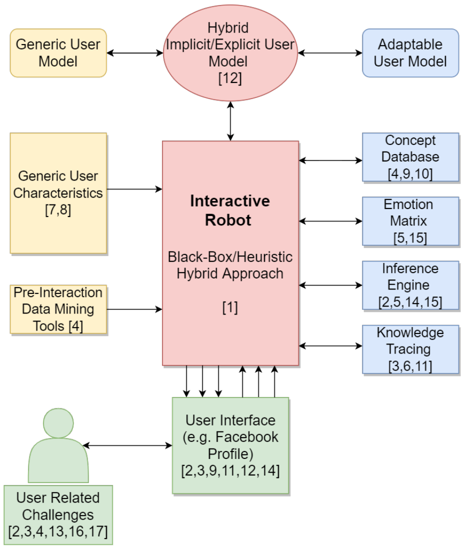

Figure 1: Overview of Interactive Robot Components and their Associated Challenges (Numbered in square brackets)

sual art discussions or videoconferencing (Cohen-Mansfield and Perach 2015), which suggests the usefulness of promoting social interactions.

Thus, in the current paper, we follow a speculative design approach based on *scenario-building* to ask the research question: What if there were a robot that could use social media mining to provide companionship to drivers?(–What might such a robot be like, and what might be some challenges and possible countermeasures?)

Methodology

To address the research question, a *speculative* approach was adopted, that involves identifying potential alternative realities and their challenges. Such an approach is required since we don't yet know how to create the desired kind of robust, effective system: e.g., in 2016, Microsoft's chatbot Tay was shut down within hours of deployment after starting to behave in an undesirable fashion (Chai et al. 2020). The closest study to the current one might be one by Easton and colleagues, which explored the idea of acceptance of a virtual conversation agent by patients with a respiratory disease, via holding some co-design workshops and developing a simplified prototype (Easton et al. 2019).

Here, however, the goal is not to build a prototype, but rather to gain insight, provoke thought, and stimulate discussion, before such technologies can feasibly be implemented and used. Thus, exploration involved first building a basic scenario, then brainstorming challenges and potential solutions. In the scenario we imagined, a robot posts reactions, text, and images aimed toward reducing a driver's loneliness, based on their social media activities.[2] Thus, there are two main tasks, monitoring and care.

[2]The robot does not post during driving, which can be detected by monitoring changes in a driver's location.

Furthermore, we focus here on one example of a social media platform, Facebook, which is the most widely used SNS among adults (over 2B active users per month). Analyzable Facebook signals include reactions (Like, Love, Care, "Haha", Wow, Sad, and Angry), posts (text, photos, videos, groups, stories), tags, check-ins, statuses, and marketplace ads. Actionable driving data could include tracking where a driver has been based on published route data or image locations: e.g., the system could detect overwork if a driver is moving much over long periods of time, erratic behavior if the driver is following new routes that don't seem to be more efficient, or lack of concentration if a driver accesses social media while driving. Health-related information can include "direct" information about illnesses, medicine, medical staff or venues, as well as general mood; for example, posting a message at 3am might indicate sleep problems for a driver who is usually asleep at this time. As well, the kinds of messages posted or not posted could provide insight into a person's social identity, and thereby their ideals, attitudes, and preferences (*preference elicitation*); such signals can be processed through natural language processing, affective image analysis, and sentiment or emotion analysis. The system should then use this information to encourage interaction on social media to reduce loneliness.

Through discussion, we identified five key aspects to consider when designing this type of system: the Interactive Strategy, User Modeling (and Sources of Information), Technical Issues, Ethics, and User Preferences. Each aspect presents its own unique challenges to which potential solutions are explored. Fig. 1 shows an overview of how the proposed system and its individual components could look like; challenges associated with each functional block are linked via numbers to descriptions in the following sections.

Interactive Strategy

The current work represents a simplified foray into a highly complex area, where numerous future improvements will be possible via more detailed analysis and replacing simplified assumptions with richer formulations.

We start with the belief that such a robot should be personalizable with a theory of mind (e.g., use knowledge tracing to infer what a driver knows), able to explain its actions, and have some restricted learning ability. A direct interface to interact outside of Facebook could also be possible. Some related challenges and proposals are listed below:

Challenge No. 1: How a robot should respond was unclear. **Proposal:** Basic content can include information and advice about health and driving when appropriate but might soon become boring if restricted to only this. Furthermore, the robot can be used not only before an accident, but also possibly afterwards to stimulate reflection and healing. However, explicitly envisioning and determining robot responses for all contexts seems difficult.

A hybrid neuro-symbolic system could be used: we can seek to leverage data to build a blackbox approach, and identify some desired characteristics to build a heuristic approach. A blackbox approach can be trained on the data of human experts, possibly people who get many likes and shares, in terms of how they respond to posts from coworkers, friends and family. This approach can be used for "uncritical" factors such as message length and frequency of activity, where we expect there to be some "sweet spot" or range, with a balance between conveying information and being interesting, versus being brief for readability and not obnoxious/spamming.

Following a heuristic approach, a robot could reply to the most frequent, and most infrequent, patterns in a human's behavior (Glas et al. 2017); additionally, a robot could also focus more on messages containing symbols that have an extreme emotional meaning, like skulls and crashes, where the cost for inaction might be high. Thus, such detected driver behaviors and robot responses could act as "triggers" from the perspective of the Fogg Behavior Model (Fogg 2009). To act effectively, we believe such a system's "personality" should demonstrate sincere caring via a balance between *exogeneity* and *endogeneity*: exogeneity here means that there should be some part that is connected to the driver's activity, matching, empathic, user-centric, easy to understand, and "stable"; the latter means that some part can be positive, complex, creative, interesting, and stimulating, showing "variety". Being positive can provide distraction to avoid negative rumination and venting; this can include sometimes liking a driver's own posts or others' posts that the driver liked, even if the driver is not being positive to the robot.

To implement matching, one natural analogy might be to recommender systems, since the robot seeks to find something which would be good to show a human; as such, content-based filtering, collaborative filtering or sequence aware recommender systems could be applied. For example, for content-based filtering, if touring bikes elicit negative emotions in a person, then sport bikes might as well. For collaborative filtering, if one person thinks driving near lakes and forests is relaxing, another person with similar preferences who likes driving near lakes might also like forests. For sequence aware recommender systems, a person might prefer sets of recommendations that follow some logical continuation; e.g., music or videos with similar, increasing or decreasing tempo. Another person with similar listening habits might respond positively to recommendations exhibiting the same temporal structure. There are also direct analogies to exogeneity and endogeneity in desired properties for recommender systems such as "diversity" and "serendipity", or how surprising recommendations are, and "persistence", which relates to contingency/matching.

For example, in line with the responsive art paradigm in art therapy, if a person uploads a sad image of someone dying in a car accident, the robot could also express sadness by showing an image of grieving family members. To interact at this level, more advanced robots should be able to perceive not just basic sentiment, but the semantics of messages, and be able to come up with meliorative responses. Another interesting direction could be to consider allostasis and how internal state regulation could promote a positive and varied action space (Lowe 2020).

Challenge No. 2: Communication signals are deceptively complex. For example, borrowing some nomenclature from human science, the meaning of communications is not always clear to an observer (human or robot) because symbols tend to be *polysemous*, *polyvalent*, *chiasmic*, and *commingling*. Polysemous signals arise from an overdetermined condensation of multiple causes (e.g., posting a black image both due to personal color preferences and to convey a bleak feeling). Polyvalency entails that observers will add their own interpretations, seeing spurious connections where there are none, or nothing where there is a connection (i.e., *pareidolia* or *randomania*). Chiasmic here means intertwined (Shotter 2002), like how a cycle of closed-loop *entrainment* can ensue in which a robot reacts to a driver's post, and the driver reacts to the robot's post, also over possibly many sessions. Signals commingle in composition of a message or image, such that multiple symbols, possibly nested, can exert an influence over one another's meanings. **Proposal:** Agreement between black box recognition and heuristic recognition could be checked, and unclear exchanges avoided or clarified with questioning.

To address such complexities, in general, a rich user model (discussed more below), improved modeling of emotions, awareness of context, advanced recognition and inference capabilities, as well as interaction experience, might be helpful for accurately interpreting a human's communications. Regarding how a robot could seek to deal with multiple symbols, various options exist. If symbols fit together, like cars, sand and water, the robot could interpret this as driving along a beach; else, important symbols within the mixture can be singled out. Additionally, a robot can also make a judgement of how critical it is to convey meaning clearly. In caring for a first time user, monosemous symbols could be preferred for clarity, whereas some errors could be allowed in an informal session with an experienced user where the intention might be more to have fun rather than to communicate a clear message. Some drivers could also benefit from both attention from humans and robots, where robots could also elicit responses which a human therapist might not evoke, thus helping the driver in complementary ways. One caveat is that matching a robot's response to a driver's message using only semantic similarity carries a large risk of failing. For example, if a person posts a positive image of a nurse, the robot could post a negative image of a dying crash victim; a red sunset could be matched with a pool of red blood. For this, emotion can be used as a tool to check that there is a similar underlying meaning in visual communications intended to match.

Challenge No. 3: How to measure system performance was unclear. **Proposal:** Monitoring can be evaluated in comparison to manual analysis by a human expert. For care, the simple increase in interaction on social media could be objectively evaluated in numbers of posts and lengths of messages (or the degree to which driving habits seem "healthy"). Subjective surveys could be used during initial development and periodically throughout. More complex metrics could involve sentiment analysis in general or in responses to the robot, also over different periods of usage, although analysis of the behavior of irregular users might be inconclusive. Poor performance could then be improved by requesting feedback from the driver.

User Modeling and Sources of Information

To implement such a robot's behavior strategy, an adequate user model is required.

Challenge No. 4: Some drivers might not be active on social media, hide their account, or have multiple accounts. **Proposal:** In such cases, a direct medium could be recommended, such as receiving emails about driving and health, or the robot could be set to interact with more than one known account. Such cases could also be escalated to human assistants, in line with the idea of "human-in-the-loop AI".

Challenge No. 5: There can be ambiguity and noise in people's messages. For example, if a driver expresses anger about an article on irresponsible driving, are they angry about such drivers, or about the article's conclusions? Furthermore, some noise can result from clicking the wrong button, or typos. **Proposal:** Robust mining methods are desired, which can interpret underlying meaning, also by inferring the *referent*, and deal with noise.

Challenge No. 6: Changes occur to social media and people over time. **Proposal:** The robot can be adapted through continuous development, and model tracing used to track changes in a person over time (e.g., vehicles they drive or preferences). How to retain engagement in long-term interactions is an open challenge, which could be addressed via endogeneity and computational creativity (e.g., self-disclosures seeded by current events).

Challenge No. 7: It's unclear what kind of personalization model to use for this context. A simplified personalization approach relying on users to answer a small number of open-ended questions might not be able to deal with profile sparseness and cold-starts, problems often faced in the recommendation literature (Bobadilla et al. 2012)–especially if a user is unwilling or incapable of training the system through self-disclosure. Conversely, a generic model might cause problems when using tricky or "triggering" concepts for which a person's perception differs most from others; e.g., if a driver has had a bad experience with a jeep. **Proposal:** A hybrid model with both implicit stereotypes and explicit user self-disclosure could be used. This model could contain both a list of stereotypical properties that affect perception, as well as maps between specific symbols and perceptions, possibly as a modular layered architecture. When personal information is initially sparse, basic stereotypical information like location, gender and age, could be used if a driver wishes[3], until eventually only user disclosed information is used. Thus, the model is dynamically updated and can deal with missing data (implicit or explicit), while explicit knowledge has precedence over implicit and generic values; in other words, the model is a hybrid of a stereotype-based user model and a fully personalized adaptive model, depending on the information available. The trade-off between these models may be adaptive managed by a sequential decision making agent that uses upper confidence bound (Lai and Robbins 1985; Li et al. 2010; Galozy, Nowaczyk, and Ohlsson 2020) or probability matching techniques (Thompson 1933; Agrawal and Goyal 2013).

[3]although there is a risk of mistakes, as noted above, the benefit is having some way to interact, and a design to avoid damaging mistakes could be applied.

Challenge No. 8: If a driver accepts the use of stereotypes to enable initial interactions, which properties might be useful to consider are unclear. **Proposal:** Emotion offers a simplified way to qualify how a person can be expected to react to stimuli. Some variables that could affect emotional perceptions include culture, preferences, social connections, context, age, gender, and personality. For example, Gothic subculture attributes positive characteristics to expressions that are otherwise often considered to be negative, such as skulls or crashes. One person might like the thrill of racing, whereas another might detest dangerous driving. A driver might enjoy a stimulus more when with a friend, or not want to show certain emotional reactions to a stranger or superior. And, if a driver sees a message when angry, cold, wet from rain, tired, or sick, they might perceive it more negatively.

Technical Challenges

Subsequently, trying to implement solutions to the previous challenges brings forth some technical problems.

Challenge No. 9: It is unclear how to intervene acceptably for groups of users with different, conflicting profiles (e.g., new or professional driver, elderly or teenager, female or male). Also, group dynamics can be complex; the effects of mixing robots and humans in a large social network requires further study. **Proposal:** We suggest as a simple solution, rather than blending weights, finding an intersection of acceptable concepts, or concepts that minimize some joint loss function; although this runs the risk of censorship that might not be required given the context.

Challenge No. 10: In most cases, we expect there will be too many concepts that can be represented in posts to get direct feedback on all. **Proposal:** Concepts can be related via a taxonomy structure like in WordNet, an online lexical database which groups nouns and verbs into hierarchical trees (Miller 1995); e.g., a "dump truck" is an example of a "truck", which is an example of an "automotive vehicle". Thus a system could adapt automatically based on new information, updating weights encoding a driver's perception (e.g., valence and arousal for emotion)–for example, via a back-propagation-style update by computing the gradient of a loss function with respect to the weights (Rumelhart, Hinton, and Williams 1986). Adaptation could also occur on the fly in interacting; for example, a robot could post a short test message to gauge reaction, and continue if the reaction is positive or switch the topic otherwise.

Challenge No. 11: Cross-application interoperability can be a concern for modeling drivers; e.g. across social media networks, or for autonomous driving support, vehicle updates, vehicle-to-vehicle communication, etc. **Proposal:** Interoperability via a shared/unified model can be difficult in open, dynamic, and exploratory scenarios such as the current one, such that conversion could be more practical (e.g., semantic web approaches could be useful), provided that sufficient care is given to avoiding that restrictions on the use of personal data become blurred.

Ethics

Some typical dangers of AI systems include being opaque, too-easily misapplied at large scales, and damaging. Damage from health data mining, the main concern in this section, could include being penalized in regard to work or insurance/pensions, victimized through identity theft, used for others' financial gain, or subjected to stigma or embarrassment, etc. (Stockdale, Cassell, and Ford 2018).

Challenge No. 12: How to best provide transparency is unclear. **Proposal:** Blackbox methods can be used for high performance, but critical components should also be explainable and transparent–for this, heuristics can be used, or an ensemble. First, the robot should be known to be a robot. Furthermore, we propose accountability via reports, audits, and deterrents. For reports, a transparent (blockchain-based) reporting system can be used. At the start, a supplier's declaration of conformity (SDoC) should be provided to the driver for the robot, possibly upon receiving a license; the system's scope, goals, and limitations can be clarified with a user agreement. Regular updates on any changes to the robot should be sent to drivers. Next, audits by humans or other robots should be conducted periodically on random interactions and robots to check correct operation. Finally, deterrents should be attached to irresponsible behavior: it should be clear who is responsible if the robot does something wrong. Likewise, the driver must still feel responsible for their own actions, and not wonder why the system didn't stop them from causing an accident if it was supposed to be monitoring.

Challenge No. 13: The system should be set up in such a way that it should not be easy to misuse it for different objectives than it was built to address. The challenges and opportunities discussed here are not limited to drivers, as well-being and reduced loneliness are universally desirable qualities. For example, such a system could be used for people with dementia to reduce their loneliness by seeking to infer pleasant memories, disease progression, engagement in physical activities, etc. However, unknown pitfalls could present themselves when seeking to use such a system with a different demographic. **Proposal:** Use on different demographics must be carefully explored in advance. This can also be upheld via agreements.

Challenge No. 14: The robot should respond appropriately to avoid hurting or embarrassing people on social media. We expect that robots in the near future will make mistakes because the technologies involved are in development, and even humans err. **Proposal:** The state of technological readiness should be clearly conveyed to the drivers; e.g., that the robot is a work in progress. The robot should also be designed in such a way that seeks to avoid and limit the possible bad effects of mistakes. For example, the robot's role should be clear, that it is not an expert who will diagnose a person as dangerous or a poor driver, but rather a partner facilitating self-reflection to maintain healthy driving habits. Furthermore, socially acceptable symbols should be used which are appropriate for the context, using a taboo model. For example, affective image datasets can contain sexual images that should not be posted where a child could view them; or, a gaze tracker could identify positive emotions in a driver upon seeing an attractive passenger that could be embarrassing to share with others. Also, negative stereotypes connected to people's insecurities should not be used; e.g., the robot should not post pictures of drive-throughs selling hamburgers and doughnuts in the assumption that an overweight driver must be fond of food. Transparency would also allow the robot to explain its posts, to ensure the lack of derogatory logic.

Another possible source of embarrassment could result if a robot's behavior becomes corrupted by *trolls*.We propose specifically incorporating a model for detecting and reacting to trolls, as well as limiting the robot's ability to learn and adapt. The robot should not just censor certain keywords, which might create unreasonable censorship, but take context into consideration.

A robot could also seek to detect humor and diffuse conflicts without seeming sarcastic, deceptive, or derogatory. Additionally, a robot should not appear like a "Big Brother" that is looking over a driver's shoulder, as it might feel disagreeable to be tracked and evaluated, like an invasion of personal space, unless it is clear that the driver is okay with this. The robot should also not get in the way, or be perceived as nagging, which might cause a driver to start to ignore it. Rather, it should be clear that the robot wishes to interact positively with, and only report to, the driver.

Challenge No. 15: Matching could lead to negative effects if a robot always seeks to match the emotional quality of a person's posts; e.g., there could be a danger of undue sheltering, related to filter bubbles and echo chambers. Shelter means to "prevent (someone) from having to do or face something difficult or unpleasant"[4].A filter bubble is a state in which a user is isolated through technology from information that disagrees with their own viewpoints, such that their view of the world can become skewed (Pariser 2011). Echo chambers are clusters of users (here a robot and user) characterized by homogeneous aligned thought that are believed to facilitate the propagation of misinformation and radicalization of thought through repeated biased reinforcement (Cota et al. 2019). In other words, if the robot is overly protective and reinforces unhealthy behavior, drivers could be denied stimuli that could help them to grow and otherwise negatively affected. For example, a perception that it's okay to drive dangerously, or that other drivers' opinions are generally wrong, might not be good to nourish. Conversely, matching can also be used to corrupt robots, like with Tay.

Proposal: Excessive matching should be avoided. There should be an exploration-exploitation trade-off to allow diversity, achievable through, for example, adaptive models with behavioural constraints (Balakrishnan et al. 2018).

User Preferences

To facilitate acceptance and avoid early abandonment of the robot solution, a user's needs in regards to personal integrity and privacy should be respected and addressed.

Challenge No. 16: Drivers will probably not want this service if it could mean being fired or suffering financial

[4]https://languages.oup.com/google-dictionary-en/ – Accessed: 2020-10-28.

penalties. **Proposal:** The boundary between home and work should generally be respected, but use of such a robot could be obligatory, given the seriousness of the threat of accidents. For example, drivers already require licenses, since the consequences of accidents or assaults can involve multiple fatalities (especially for drivers of heavy vehicles). Initialization of the service could take place at the same time as a license is obtained. To reduce the possibility of damage, drivers should have control over the robot. Any mined medical information should be reported only to the driver, and not posted publicly, to avoid the risk of falsely reporting a driver as dangerous, and the driver should be allowed to delete posts. Moreover, private data should be securely dealt with to prevent facilitating identity theft or discrimination/stigma based on the emotion profile. Models can be trained in a decentralized manner (Kairouz et al. 2021) or through the use of differential privacy (Ji, Lipton, and Elkan 2014), respecting the user's security concerns.

Challenge No. 17: If a driver is willing to self-disclose information but forgets or does not use Facebook or other social media, what other possibilities could exist? **Proposal:** One alternative way to detect tricky concepts could be to directly ask their caregivers or acquaintances, or mine their data. If this is not possible, how could the robot fit a model for people's emotional reactions, in some way that doesn't require much effort?

A physical or virtual companion robot that spends much time with a person (e.g., following or walking beside them, or running on their mobile device) could observe reactions to some stimulus, e.g., using a camera to detect facial expressions and gaze, open-ended questions in everyday conversation, or even a brain-machine interface, or perspiration/galvanic skin response. Another possibility is for a robot to detect preferences based on where a person drives or what a person chooses to surround themselves with, either in their car, at home, or at their workplace (and what is absent); e.g., object detection could be used by a robot in an environment or from photos. One difficulty is that a car could contain a mixture of items required for work, items received from others, and personal items, which might be difficult to tell apart. Another source of information could be a person's interactional context, e.g., purchase history (Hariri, Mobasher, and Burke 2012). A robot could also seek to detect a person's personality by interacting (Abe et al. 2020). If age and gender are not explicitly disclosed to the system, they could be inferred with computer vision or sound.

To gain information, a humanoid form might be more effective than an abstract form like a phone (Jinnai et al. 2020). And, not all tricky concepts might be easy to detect, like Google's example of a person in a wheelchair chasing a turkey; e.g., if a person hates red cars since their acquaintance had an accident in one. Thus, while perfect personalization is unlikely, implicit knowledge can be stored in a list.

Conclusion

We present a speculative scenario involving a robot intended to continuously interact with a driver on social media to reduce loneliness, by monitoring activities and posting or responding; sensitive data such as monitoring results should be visible only to the owner, whereas typical, casual interactions can be public. The idea is to leverage easily accessible, plentiful data and interactional affordances that are not currently being used in this context, to "reach" more drivers and in a better way. This can be a complement to other approaches that require drivers to do something extra, like wearables, a physical robot, or mobile surveys. In identifying 17 related challenges, we propose potential solutions regarding interactive strategies, user modeling, and ethics. Limitations include the preliminary, theoretical nature of our study, as well as the focus on driving and Facebook, which might be replaced by another social network in the near future; a robot implementation and other applications and platforms will be explored. Although real-world deployments related to this ambitious proposal might still be far-off, the aim is that such robots could exert a positive influence on interacting humans, thereby contributing in an indirect way to traffic safety and quality of life in our evolving smart cities.

Acknowledgments

We gratefully acknowledge support from the Swedish Knowledge Foundation (KKS) for the "Safety of Connected Intelligent Vehicles in Smart Cities – SafeSmart" project (2019–2023), the Swedish Innovation Agency (VINNOVA) for the "Emergency Vehicle Traffic Light Pre-emption in Cities – EPIC" project (2020–2022), and the ELLIIT Strategic Research Network.

References

Abe, K.; Nagai, T.; Hieida, C.; Omori, T.; and Shiomi, M. 2020. Estimating Children's Personalities Through Their Interaction Activities with a Tele-Operated Robot. *Journal of Robotics and Mechatronics* 32(1): 21–31.

Agrawal, S.; and Goyal, N. 2013. Thompson Sampling for Contextual Bandits with Linear Payoffs. In Dasgupta, S.; and McAllester, D., eds., *International Conference on Machine Learning*, volume 28:3 of *Machine Learning Research*, 127–135. Atlanta, Georgia, USA: PMLR.

Balakrishnan, A.; Bouneffouf, D.; Mattei, N.; and Rossi, F. 2018. Using contextual bandits with behavioral constraints for constrained online movie recommendation. *IJCAI International Joint Conference on Artificial Intelligence* 2018-July: 5802–5804. doi:10.24963/ijcai.2018/843.

Benson, J. A.; McSorley, V. E.; Hawkley, L. C.; and Lauderdale, D. S. 2021. Associations of loneliness and social isolation with actigraph and self-reported sleep quality in a national sample of older adults. *Sleep* 44(1): zsaa140.

Bobadilla, J.; Ortega, F.; Hernando, A.; and Bernal, J. 2012. A collaborative filtering approach to mitigate the new user cold start problem. *Knowledge-Based Systems* 26: 225–238.

Chai, Y.; Liu, G.; Jin, Z.; and Sun, D. 2020. How to Keep an Online Learning Chatbot From Being Corrupted. In *2020 International Joint Conference on Neural Networks (IJCNN)*, 1–8. IEEE.

Chen, M.; Tian, Y.; Fortino, G.; Zhang, J.; and Humar, I. 2018. Cognitive internet of vehicles. *Computer Communications* 120: 58–70.

Cohen-Mansfield, J.; and Perach, R. 2015. Interventions for alleviating loneliness among older persons: a critical review. *American Journal of Health Promotion* 29(3): e109–e125.

Cota, W.; Ferreira, S. C.; Pastor-Satorras, R.; and Starnini, M. 2019. Quantifying echo chamber effects in information spreading over political communication networks. *EPJ Data Science* 8(1): 35.

Easton, K.; Potter, S.; Bec, R.; Bennion, M.; Christensen, H.; Grindell, C.; Mirheidari, B.; Weich, S.; de Witte, L.; Wolstenholme, D.; et al. 2019. A virtual agent to support individuals living with physical and mental comorbidities: co-design and acceptability testing. *Journal of medical Internet research* 21(5): e12996.

Fogg, B. J. 2009. A behavior model for persuasive design. In *Proceedings of the 4th international Conference on Persuasive Technology*, 1–7.

Galozy, A.; Nowaczyk, S.; and Ohlsson, M. 2020. Corrupted Contextual Bandits with Action Order Constraints.

Garbarino, S.; Magnavita, N.; Guglielmi, O.; Maestri, M.; Dini, G.; Bersi, F. M.; Toletone, A.; Chiorri, C.; and Durando, P. 2017. Insomnia is associated with road accidents. Further evidence from a study on truck drivers. *PLoS one* 12(10): e0187256.

Glas, D. F.; Wada, K.; Shiomi, M.; Kanda, T.; Ishiguro, H.; and Hagita, N. 2017. Personal greetings: Personalizing robot utterances based on novelty of observed behavior. *International Journal of Social Robotics* 9(2): 181–198.

Hariri, N.; Mobasher, B.; and Burke, R. 2012. Context-Aware Music Recommendation Based on Latenttopic Sequential Patterns. In *Proceedings of the Sixth ACM Conference on Recommender Systems*, RecSys '12, 131–138. New York, NY, USA: Association for Computing Machinery.

Ji, Z.; Lipton, Z. C.; and Elkan, C. 2014. Differential Privacy and Machine Learning: a Survey and Review.

Jinnai, N.; Sumioka, H.; Minato, T.; and Ishiguro, H. 2020. Multi-Modal Interaction Through Anthropomorphically Designed Communication Medium to Enhance the Self-Disclosures of Personal Information. *Journal of Robotics and Mechatronics* 32(1): 76–85.

Kairouz, P.; McMahan, H. B.; Avent, B.; Bellet, A.; Bennis, M.; Bhagoji, A. N.; Bonawitz, K.; Charles, Z.; Cormode, G.; Cummings, R.; D'Oliveira, R. G. L.; Eichner, H.; Rouayheb, S. E.; Evans, D.; Gardner, J.; Garrett, Z.; Gascón, A.; Ghazi, B.; Gibbons, P. B.; Gruteser, M.; Harchaoui, Z.; He, C.; He, L.; Huo, Z.; Hutchinson, B.; Hsu, J.; Jaggi, M.; Javidi, T.; Joshi, G.; Khodak, M.; Konečný, J.; Korolova, A.; Koushanfar, F.; Koyejo, S.; Lepoint, T.; Liu, Y.; Mittal, P.; Mohri, M.; Nock, R.; Özgür, A.; Pagh, R.; Raykova, M.; Qi, H.; Ramage, D.; Raskar, R.; Song, D.; Song, W.; Stich, S. U.; Sun, Z.; Suresh, A. T.; Tramèr, F.; Vepakomma, P.; Wang, J.; Xiong, L.; Xu, Z.; Yang, Q.; Yu, F. X.; Yu, H.; and Zhao, S. 2021. Advances and Open Problems in Federated Learning.

Lai, T.; and Robbins, H. 1985. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics* 6(1): 4 – 22. ISSN 0196-8858. doi:https://doi.org/10.1016/0196-8858(85)90002-8.

Li, L.; Chu, W.; Langford, J.; and Schapire, R. E. 2010. A contextual-bandit approach to personalized news article recommendation. *International conference on World wide web - WWW '10* doi:10.1145/1772690.1772758. URL http://dx.doi.org/10.1145/1772690.1772758.

Lowe, R. 2020. Maximization of Future Internal States? *Constructivist Foundations* 16(1): 060–062.

Meuleners, L.; Fraser, M. L.; Govorko, M. H.; and Stevenson, M. R. 2015. Obstructive sleep apnea, health-related factors, and long distance heavy vehicle crashes in Western Australia: a case control study. *Journal of Clinical Sleep Medicine* 11(4): 413–418.

Miller, G. A. 1995. WordNet: a lexical database for English. *Communications of the ACM* 38(11): 39–41.

Pan, B.; Zheng, Y.; Wilkie, D.; and Shahabi, C. 2013. Crowd sensing of traffic anomalies based on human mobility and social media. In *Proceedings of the 21st ACM SIGSPATIAL international conference on advances in geographic information systems*, 344–353.

Pariser, E. 2011. *The filter bubble: What the Internet is hiding from you*. Penguin UK.

Rosso, G. L.; Montomoli, C.; and Candura, S. M. 2016. Poor weight control, alcoholic beverage consumption and sudden sleep onset at the wheel among Italian truck drivers: a preliminary pilot study. *Int J Occup Med Environ Health* 29(3): 405–16.

Rumelhart, D. E.; Hinton, G. E.; and Williams, R. J. 1986. Learning Representations by Back-propagating Errors. *Nature* 323(6088): 533–536.

Shotter, J. 2002. Spontaneous responsiveness, chiasmic relations, and consciousness: inside the realm of living expression. *New Hampshire* 34.

Stockdale, J.; Cassell, J.; and Ford, E. 2018. "Giving something back": A systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland. *Wellcome open research* 3.

Thompson, W. R. 1933. On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples. *Biometrika* 25(3/4): 285–294. ISSN 00063444. URL http://www.jstor.org/stable/2332286.

Yang, C.-c. 2016. Instagram use, loneliness, and social comparison orientation: Interact and browse on social media, but don't compare. *Cyberpsychology, Behavior, and Social Networking* 19(12): 703–708.