Goodbot, Badbot? Engineering trust into conversational interfaces

Dave Horsfall

As more people interact with bots driven by Artificial Intelligence, it is important to understand how to create relationships based on trust. The FinTrust Project at Newcastle University is looking at the role of machine learning in banking, particularly in the context of automated chatbots and how the presence of socio-emotional features impacts the intention to use. Researchers from social sciences and computing are examining how trust can be measured, how trust is gained and lost, and what implications this might have on design considerations.

The team has used theories on trust in automation, emotion, and theories of computer-human interaction to categorise trustworthy traits. Engineering these social cues as features into conversational interfaces has offered an interesting start to my RSE career, and spans a variety of areas from Natural Language Processing to careful UI Design. In this talk I will firstly discuss my step from industry into the RSE profession at the end of March, just a few days after severe lockdown restrictions where imposed by the UK government. The challenges of starting this new role and collaborating with colleagues I have been unable to meet in person have been significant, but ultimately very rewarding. I will explain my approach within this multi-disciplinary project to use Dialogflow to quickly leverage Google's powerful machine learning. This allowed us to build an operational chatbot prototype that was capable of exhibiting varying degrees of social presence, such as politeness, empathy and active listening. The talk will review the infrastructure design for AI powered chatbots, and illustrate how machine learning can interpret the intent of what users are saying, to gather information and respond appropriately.