Conversational AI by **Professor Julie Wall**

Professor Wall, of the University of East London, has been working on Artificial Intelligence (AI) programs to detect fraud, revealed by the way people use language in statements they make to organisations who routinely record phone calls, for instance the police or insurance companies. As a result there is much information available to 'train' AI software.

Speech and natural language technology have advanced at a rapid pace in recent years. Speech recognition was an early target for Artificial Intelligence, which has now become so good that it can be taken for granted.

This advance has been driven in part by the development of Graphics Processing Units and the deep learning frameworks that use them. These developments have markedly affected the way in which humans communicate with computers and are currently driving numerous commercial products, loosely termed 'Conversational AIs', that rely on speech, natural language processing, and natural language understanding.

Fraud may be first-party, where a person makes a fraudulent claim; or third-party, where a person has been hoodwinked by a dishonest vendor who has (for example) taken over their bank account.

The usual responder to a call is in a call centre, its staff seldom or insufficiently trained in how to spot deception, which is becoming more sophisticated and widespread, and who may become desensitised with the volume of calls even with training. An operator will notice stress and tone in a call but this is unreliable. If a person calls twice a voice comparison can be made automatically (but is also not reliable); or other (meta)data known about them can be checked (have they made a previous fraudulent claim?). Such checking is not available in real time but is usually done overnight.

Conversational AI is based on machine learning and deep learning approaches. It takes the next step and aims to check the way, and how, people speak, and how they respond to questions; it can be trained to sense when a person is being deceptive. (Do callers have to pause to think?) It has to avoid giving incriminating evidence and to be convincing. Forensic Linguistic analysis involves grammar, and the way things are said, including pauses, etc.

Professor Wall played a recording by an actor of an anxious person in trouble, being asked an open question. This was transcribed into text and various words highlighted with **Markers** for:

(Missing Pronouns**)** and Negation - Attempting to distance yourself from a statement is suspicious.
Interestingly, negation alone has very little impact, but another negation, at very close proximity drastically increases the deceptiveness of the conversation.

Repetition - people are buying themselves time while thinking.

Explainers - Giving unnecessary information, over-explaining: "I broke three bones in my body." the words 'in my body' are unnecessary.
Disfluencies (ums, ahs & over long pauses) together with Explainers, before Negation is not sensitive, but before a positive statement is very sensitive.

Hedging - Shows uncertainty, being unclear about what happened; it either did or didn't happen.

Lack of Memory - is suspicious.

None of the markers is in itself incriminating, but more than one can raise concerns. The Proximity of markers in an analysis can be significant.

Q: Tell me what happened?

A: Um, well, (X) wasn't, wasn't in the house because (X) been out and when (X) come home there was a flood in my bedroom because my ceiling is leaking. I don't remember seeing anything before. I think it happened when I was out.

This example does not convince – although to a naïve audience (us) it was credible.

Professor Wall then played a recording of a real conversation between someone reporting a minor crime and a duty policeman in the Lancaster police department. This was taken as legitimate, and

two policemen were dispatched to the scene. It was a trap and a policeman was killed. A voice-to-text transcription was made and later analysed – and found to be suspicious.

It can take an analyst a couple of hours to mark up a two page transcript, or six hours for an emergency call.

LexiQal is a programme developed by Professor Wall and her colleagues in a two-year project, which automates this process to give real-time assistance to call centre staff. Briefly it:

- derives Acoustic and Linguistic markers from the speech input;
- makes a Transcript of the call and adds speech marker information.
- This is fed to a marker Proximity detector; and
- fanned out to a number of individual Marker analysers (*as described in the example above*);
- which each produce a Deception Indicator (in the range 0 – 1)
- these are combined to give an overall Deception Score. A score over 0.8 requires investigation.

Care has to be taken to ensure data privacy (GDBR), requiring that how a decision was made can be explained.

BERT, an open source program developed by Google, gave the key to designing Marker analysers. It uses machine learning and deep learning approaches which are behind Conversational AI. A large supercomputer was used to create a vast multilingual database from which to derive word connections.

A relatively simple programme based on it can be set to look for specific Markers in a given database. Such a database is bespoke for each usage – e.g. Insurance or Policing. For development trials of LexiQal there was no pre-existing database so they made their own – much taken from films (script writers try to make their dialogue realistic, and actors may ad-lib).

Insurance and Policing were the businesses they were first involved with, where there were already large databases, though hardly used in this application as existing manual processes were so slow with results coming long after the event.

LexiQal programmes do not need supercomputers, and can be run on laptops. Once trained they can provide real-time assistance to staff - who may not be working in call centres. It has been a success and is being taken up in a range of businesses, with several companies offering to provide it.

There have been surprises – not everything was as expected by experts.