On having Robots/AI report their decisions

This document describes the issues perceived as requiring attention from law & policy makers, both for the UK and internationally, in the matter of Artificial Intelligence programs and robots having legally to explain their decisions in the light of a dispute. These include the opaqueness of rule-based programming, the scope of AMA\(^1\) responsibility, “Big Data”\(^2\) driven decisions, and the lingua franca of explanation.

1) Opaqueness of rule-based and other AI program decision making.

A major feature of rule-based expert systems is that they are not programmed using procedural scripts, stepping from one stage to another in the way of a BASIC or COBOL program, but Expert Systems\(^3\) using logic languages such as PROLOG describe logical constraints and seek goals. The precise manner in which the program performs this goal-seeking behaviour is not always very visible, can be huge and very complex, particularly if using heterogeneous data as a basis. It is possible to output to a log file the status of the system at any given stage in the process, but these log files are likely to be huge and incomprehensible except to experts. (see point 4). The maintenance of huge log files will be a large cost to the user organization.

2) On the scope of AMAs responsibilities in moral decision-making

We are currently working on the development of an Artificial Moral Agent program, called Project ERGON, and have developed a first draft prototype called ArbiTer2, a system which makes judgments in Human Relations Disciplinary cases. So far the reasonableness of the decisions is encouraging. However, in the course of our R&D we have identified a phenomenon in ethical decision-making which is a matrix of time and hierarchy in making moral decisions. For example, one relatively rarely makes life-and-death decisions in comparison to daily routines of balancing a visit to the dentist over picking the children up from school? These moral decisions have both a weight and frequency component. The issues for AMA design are that to program a response to a once-in-a-lifetime decision to cover liability may be onerously costly, and in many cases a redundant one?

---

1 Artificial Moral Agent – a computer system which attempts moral judgment of right from wrong for example.
2 “Big Data” are systems which process really huge amounts of data in data-driven programs.
3 Expert System – a branch of Artificial Intelligence which develops adviser systems.
With regard to product liability, is there scope for a cap or ceiling on such liability to ensure insurance for SMEs is not prohibitive to innovation and growth?

3) “Big Data” driven decisions

With machine learning systems the process is data-driven, not program driven. Thus it may well be almost impossible to predict and therefore test for all the scenarios the system may face? In the case of really Big Data the log files for processing would be arcane to decipher even for an expert. This may also be true of Expert Systems and AMAs in 1) and 2) above?

4) The Lingua Franca of explanations.

The legal requirement in the GDPR legislation and other requests from the EU Parliament for systems to explain their decisions in cases where humans are affected and involved means that there needs to be a standard lingua franca of outputs agreed internationally to make sure that non-experts, courts and media can understand what went on. The other side of this is that every system would have to have as a specific component a huge translation module, absorbing processing time, just to record history? It would not be impossible to develop such a lingua franca, or translation table from computer codes to English, for example, but would require a joint design committee formed from members of the judiciary, insurance, developers including micro-SMEs, legislators and representatives of the public, “the person on the Clapham Omnibus”.

I heartily recommend that such a committee be formed to agree the parameters, design and use of such a translation function for decision explanation, a ‘lingua franca’.

April 2017

---

4 EU General Data Protection Regulation
5 Micro-SME - Very small and medium sized enterprises including individuals; “one-person bands”.