Arm Response to European Commission Consultation on AI

Introduction

Arm is a global high tech company with Headquarters in Cambridge in the UK. Our core business is the design of microprocessors, including ones for use in 'AI' systems.

Last year Arm published a manifesto on AI and Ethics. https://www.arm.com/blogs/blueprint/wp-content/uploads/2019/11/Arm-AI-Trust-Manifesto-2019.pdf

We are now working on how best to put that into operational practice in a way which we hope could provide a model for other companies too.

Arm's comments are focussed on the sections of the Paper which relate to building trust.

Broad Comments

Arm agrees with the Commission that building public trust in AI is a crucially important part of securing the benefits AI can bring in various sectors of society. We agree further that is something which the entire AI sector has to aspire to.

AI supply chains are long and complicated. The challenge of building trust has to be shared along the supply chain: it cannot be reasonably left only to the company who finally puts an AI product on the market.

In that sense there is merit in the Commission's idea of an 'ecosystem of trust'. But this needs to be a global effort. The market will need to be global if the full benefits of AI are to be realised.

The Commission paper itself talks of the role of AI in helping attain the Sustainable Development Goals and the 2030 Agenda. So we need to think about how we can stimulate the *global* sector to develop trustworthy AI. The EU may be able to point the way; but it needs to make sure it can bring the global sector with it.

Problem definition

The paper's description of the problem seems quite negative about the benefits of AI and lacks detail on the extent of the wide ranging problems it alludes to, particularly in the section dealing with risks to fundamental human rights.

It might be better too if it included a definition of AI. This is not easy. But it would be helpful to know what the Commission believes is encompassed by 'AI'.

Ideally it would help provide balance if the paper could include material on how many AI systems are already in wide use, what sectors they are in, and how many have been genuinely problematic at the level suggested in the paper.

It would also be helpful if the paper - both in the problem definition section and more widely - was more careful to distinguish AI technology per se from the *applications or uses* of the technology. For example, the paper comments that 'AI may also be able to trace and deanonymize data about persons'. This is not a problem with AI per se, but with the use of AI. The answer to tackling that problem is not to stigmatize AI but to regulate against unacceptable re-identification of personal data.

The section on safety and liability rightly draws attention to the importance of both issues. But it might be helpful to distinguish what is unique about the AI cases, and how we might handle them. For example, is it correct that the current framework of product liability law can broadly apply to AI? If so we may not need new product liability laws, but we need to make sure we have identified any distinctive issues linked to AI which make the application of existing liability law more complex in these cases. For example are there issues around the need to have transparency of the proprietary algorithms used to power AI, and how the courts will get access to this information and make expert judgments about it.

High Risks

The Commission proposes identifying high risk AI applications by using two filters (a) is the AI being used in a particularly sensitive sector like health , transport, energy and ' parts of the public sector' and (b) in addition , is the application being used in a manner where significant risks are likely.

But the clarity of this approach is muddled by the paper also proposing two further specific areas which could be high risk , processing recruitment applications and remote biometric identification. The impression given is that despite the two filters principle, there is likely to be a long list of other specific applications which are judged 'high risk'.

'Risk' is often contextually dependent. Some judgments about the severity of risk will not be straightforward. How, for example, would we judge an AI medical device which is say 90% accurate, but human beings doing the same job are only 75% accurate?

Two points may need to be considered more fully here :

(i) It may be better to envisage an approach which is essentially context based. If we go down this route, the prime driver of regulating for health AI devices would not be a new piece of AI focussed regulation , but the existing health authorities who are better able to weigh up the advantages and risks of new health devices, whether they use AI or not. In many sectors it will be the sector experts who are best able to judge the risks attendant on introducing specific AI technology.

(ii) Does existing regulation already provide a framework for tackling the real problem anyway? For example any concerns around 'remote biometric identification' might be better approached through the existing (or modified) GDPR framework rather than through a specific AI piece of legislation.

Types of Requirements

The section of the paper on 'Types of Requirements' could also be clarified by identifying more precisely where existing legal frameworks suffice. For example when the paper talks about avoiding prohibited discrimination in training data, it may be that existing legal frameworks are adequate since the discrimination is already prohibited.

Sometimes this section introduces wide ranging concepts which will be hard to translate into law: for example the proposal that AI systems are trained on data sets that cover ' all relevant scenarios needed to avoid dangerous situations' could be difficult to put into effective law. (Would the risk of suicide in a loan applicant whose application had been turned down by an AI system be included?)

Addressees, Compliance and Enforcement, Labelling

The paper is correct to assert that there are many actors in providing an AI system. The problems with a regulatory framework designed to ensure they all conform to their individual responsibilities is (i) the difficulty of identifying exactly who is responsible for specific issues, (ii) being confident there are no gaps, and (iii) the need for any framework to keep pace with new developments.

Effective legislation in this area will therefore take time, and might prove elusive.

It might be better to think in the first instance of a different option, which is similar to the Commission's proposal, but perhaps easier to manage.

This would be to consider – perhaps as a trial – a system of voluntary 'assurances' provided by companies in the AI supply chain.

Each company would issue an 'assurance' of the steps they had taken to deal with ethical issues arising in their part of the chain. They would take these issues from the assessment checklist prepared by the Commission's High Level Expert Group on AI Ethics. This is a comprehensive and adaptable list, capable of being quickly revised by expert opinion as new issues – whether linked to new technology or to new societal concerns – emerge.

If each company in the supply chain gave an 'assurance' appropriate to their expertise, then the company putting the product on the market would be able to point to these 'assurances' (and their own) as an indication that specific ethical issues had been taken into account along the supply chain in designing and building the AI system. (This would be without prejudice to any legal determination of where liability lay in the event of a problem.)

This is a common process in industry. It has been pointed out in this context that 'Many industries use transparent, standardized, but often not legally required documents called supplier's declarations of conformity (SDoCs) to describe the lineage of a product along with the safety and performance testing it has undergone.' (IBM Research Paper Feb 2019, FactSheets: Increasing Trust in AI Services through Supplier's Declarations of Conformity.)

Such a voluntary system would help drive up trustworthiness across the whole AI landscape.

Additional mandatory conformity assessments and third party testing could be explored for those AI systems being used in particularly sensitive sectors (eg health) and, as suggested above, driven mainly by the sector specific authorities best placed to judge eg health devices.

There has been a debate about the usefulness of labelling schemes in the context of cyber security, particularly in respect of IoT Security. A label is often held up as providing clarity to the eventual customer. But there are potential drawbacks.

These include: - who affixes the label? Not all the components of an AI system are suitable for ‘labelling’ (take software, training, the deep tech behind the machine).

- will the company putting the product on the market be prepared to label a scheme, unless backed up by the sort of ‘assurances’ described above?

- and for how long can we expect a label to remain valid after the system has ‘left the factory’? Over the air upgrades are an important part of security, and may play a role in updating AI systems. How do these fit into a labelling scheme?

Stephen Pattison

VP Public Affairs

Arm

Stephen.Pattison@arm.com

May 2020