Women Leading in AI – manifesto launch

Ivana Bartoletti – opening speech

Dear Friends and Colleagues,

We are really delighted to welcome you this evening and, first, let me introduce the Women Leading in AI team (Allison, Rebecca, Samara, Emma, Liz, Sanya) who have worked relentlessly over the last few months preparing the report, and for the event tonight. And thank you Jo Stevens MP for hosting us. I will speak for about 15 minutes introducing the recommendations we have produced – and to do that, I want to navigate you through the rationale underpinning them.

Why are we here and what is the problem? Let's nail it down.

AI has been around for a long time now but what we have seen over the last few years has been an explosion of AI impacting on daily life. Sometimes cutting through the hype has proven tricky. Not a day passes by without reading of a whizzy new product coming to life. Without doubt AI is powerful and genuinely promising – and across different fields. We can set it to detect cancer earlier, fight fake news that undermines our democracies and support our cyber-defence. We can live better, support kids with disabilities and personalise medical treatments.

So far – so good.

But over the last few months there have been enough examples of negative, downright disturbing outcomes to turn tech ethics into a growth industry. At conferences and international events, corporates have shown us their ethical toolkits, academics have discussed how to mitigate bias and businesses have been racing to explain the ethics underpinning their cutting edge products.

So far – so good. Is the ethics bubble about to burst though? I think so.

The debate is showing its limitations. Not just around the questions of how we turn ethics into practice, obviously not an easy task. But in the framing of the discussion around what ethics IS in the first place.

For the benefit of the one or two here who might not know it, the Joanna Bryson is a global name in AI and she says: "AI changes nothing, and everything." And I agree with her.

It changes nothing in the sense it is about computation, tasks we do being delegated to machines. No more than we are doing already.

But it is also true that AI changes **everything** –because AI is much more than technology, much more than algorithmic fixes – it shapes our lives, and that is why we are here today and why we started Women Leading in AI Network.

AI is more than tech. It is power.

It is power, reshaping global relationships as countries are engaged in a race to equip themselves with the latest technology.

It is power because some are gaining a lot, and some aren't – and that divide is getting bigger.

It is power because we know algorithms can lock people out of the housing ladder or prevent them from accessing credit. And there are two problems with that:

First, if algorithms are building assumptions about us based on variables we are not even aware of, "it becomes a human rights problem" as Sandra Watchers says.

Second, if algorithms make mistakes in their decision making and lock the most vulnerable out – who then do not challenge decisions, it becomes a problem of automating inequality. Should we really be surprised if AI mysteriously seems to take the side of the powerful? Hadn't we better wake up and push in the opposite direction?

And it is power because if unchecked, AI can turn stereotypes into prejudices into active discrimination as we are seeing already – and across sectors. Recently, Amazon recruiting tool had to be dismissed as it was only recruiting men.

So what is the answer? The answer is that ethics, to be meaningful and powerful, requires regulation to support it.

There is nothing inevitable about how AI is to be deployed. People make these decisions. And now is the time to ensure the decisions are made for the common good, and channelled by our human values.

Over the years Parliament has regulated many sectors and many areas from broadcasting to healthcare – and AI must be no different. Any industry worth its salt has nothing to fear from regulation.

Corporations defining their own ethics practices is neither common sense nor for the common good. In my day job, I spend all day long demonstrating to organisations that in the long term, their profits are not going to be eroded by going beyond mere compliance.

Business cannot shy away from agile, business friendly and innovation-friendly regulation.

And it is reassuring that many are asking for regulation now, especially in the aftermath of the big data revolution which has fuelled online manipulation and helped build the persuasion architecture that is affecting our democracies and polluting public debate.

Regulation must happen now – at sector level, as regulating AI in healthcare is different from regulating it in financial services. This is why we are proposing these things are done:

> ❖ Establish a regulatory function, sitting alongside the Information Commissioner and the Centre for Data Ethics with the power to audit algorithms, set out clear codes and parameters, and be there to support citizens querying the impact of automated decisions affecting them – from pricing to housing. A regulatory function with teeth able to challenge the effect of proprietary algorithms, able to scrutinise them.

- ❖ Introduce mandatory Algorithm Impact Assessments, AIAs, for those algorithms having a strong impact on people's lives, including for example housing, identification of vulnerable people or allocation of benefits.
- ❖ Introduce a Certificate of Fairness to give a big, public tick to algorithms that are audited, and allow companies that deploy them to benefit from a liability safe zone. That means that if companies can demonstrate due diligence in every aspect and yet still some inadvertent error happens, they get some credit and some protection from part of their liability risks. This is to encourage innovation and good practice and reassure business they won't suffer a commercial disadvantage for doing the right thing.

Assessing algorithms means talking about bias and fairness.

I do not have to tell you as you are wise and savvy in here. AI is about classification – but **we must now deal with the politics and the consequences of these classification practices**.

Classification is **a product of history**, **culture and power** in which each category of data tends to promotes one point of view and silences another.

So, what happens if we keep churning out algorithms without looking at the politics of classification? If we do not put enough attention into actively looking at the data coders use? At the diversity of those embracing coding as a profession?

This must end now – and we need a bold public policy respond to tackle this.

But most important, ***what happens if we get rid of the judgment call***? A dispute can be hard to fight – especially for those who cannot afford it. And the risk is that we automate inequality: one of the biggest risks of our time. As public bodies scramble for resources, automated decisions with no human judgement call can place the life of many at risk – especially the most vulnerable. This is why we need a regulator being able to scrutinise algorithms making decisions affecting people's lives. It is not just about challenging outcomes – it is also about ensuring citizens know if they are being subjected to automated decision.

Neither the UK Data Protection act nor GDPR go far enough on this. We need to know the systems deployed have gone through rigorous assessments. IT is like going on a plane – you may not know how the plane works but you are reassured by the fact that the aviation industry has serious regulations governing how planes are built. A plane crash is a horrific but fairly rare event.

We have said before that AI is more than an algorithm fix. Ultimately, we can have the most perfect algorithm and yet be using it for something unethical. The fact that something is technically possible does not mean it has to be deployed. But who is making these decisions? This is why we need citizens engagement in all this – and the workers.

Over the last few years we have seen a surge in workers taking responsibility to oppose some of the uses of technology being pursued by their employers. Or challenging internal practices of often too unaccountable big tech companies – like we have seen with Google.

We need a way for employees to have a say – and to do so without fear. Workers representation on ethics committees could be a way forward. But when assessing the impact of AI deployment, companies need to bring their employees with them. This is why the AIA should also include the impact on the workforce and society – AI is too big to be left to commercial interests alone.

And finally – this is a debate for all.

Although we can understand why some people take that approach, we do not want to smash AI in the mistaken believe it would save workers. AI is a revolution we want to embrace but there is nothing inevitable about it. Algorithms do not have to become policy makers, dictating our cities, our work, our home, our tastes, reading habits or voting intentions. That is for humans to do.

This is why – and no apologies for hijacking the language – **we need politicians and society to take back control of technology**. Not to curb it but to enable it so that it works for everyone, makes us better and creates benefits for all.

Progress, especially that made by women, is always at risk of being unpicked - But in the case of AI, it can also send us whizzing backwards.

We need strong public policy solutions now, so we make ethics a reality and bring the decisions about our tech future back to the public realm, where they belong.