مبادرات محمد بن راشد آل مكتوم العالمية
Mohammed Bin Rashid
Al Maktoum Global Initiatives

كلية محمد بن راشد للإدارة الحكومية
MOHAMMED BIN RASHID
SCHOOL OF GOVERNMENT

session

Building High Governance AI Systems

Prof. Melodena Stephens

2023

Contents

Why is AI referred to as artificial intelligence systems? 03

AI from a systems perspective 05

AI versus human intelligence 10

AI system governance 13

Challenges facing intelligence systems 18

Opportunities and trade-offs 21

How can we secure the AI systems of the future 22

"I love to learn about and read as many book summaries as I can. Learning and leadership are prelusion and result, and a prelusion cannot be separated from a result."

His Highness Sheikh

Mohamed bin Rashid Al Maktoum

Reading = Knowledge

Knowledge = Power

Power = Respect

Respect = Success

Success = Happiness

Why is Artificial Intelligence (AI) referred to as Artificial Intelligence Systems?

The reason is that AI is not operated in a vacuum, nor is it just an algorithm, an app, or data, but AI is a system that combines data, software, and hardware.

AI: Take a Systems Perspective

We produce approximately 120 zettabytes of data per year More astounding is that over the last two years, we have produced 90% of the world's data (so talk about biases – how long have humans lived in this world?). Data needs to be managed – you need good quality data for AI – suggesting it needs to be collected, cleaned, relevant, refreshed, and secured to use it for the software part. This process of data management is complex, and there are many choices. **For example**, did you know there are 100 file formats in which you can store data? Sometimes data can get corrupted too!

Next, think of the AI software that is created to process data via algorithms for decision making. There are over 500 programming languages! Just think of all the complications when we translate real languages! Now you are talking to a machine! Before the dot.com crash, languages like COBOL (a first-generation language) and FORTRAN (a second-generation language) were popular. FORTRAN, a leading scientific language of preference, was used to get a man on the moon. COBOL, in 1997, accounted for 80% of computer programming languages. COBOL was used for banking (43% of banking and 95% of ATM machines), medical care, and government services (and still is). Today there is a desperate need for COBOL and FORTRAM programmers.

Newer languages computer are often connected to older ones using APIs (which connect bits of code). A survey of 37,000 developers found 51% of their time was spent on developing APIs, and they had one security incident a month!

When Elon Musk took over Twitter, he commented that they had 1200 RCPs (a type of API – the other is called REST) that was slowing the services down! So yes, software needs language choices, data training, and ongoing HUMAN management, or it will become obsolete, become a security liability or result in challenges like biases, malfunctions, etc.

Finally, you cannot have AI without hardware. Some of this hardware is in the form of public infrastructure. Consider the 5G and 6G wars or the recent cable wars between countries! There are 400 cables on the seafloor responsible for 95% of data transfers. As cloud computing becomes popular (they also can hide carbon emissions in the supply chain, please read my article here), there are massive requirements for running, maintaining, and cooling these cloud server warehouses. Hardware needs development: silicon chips may be vulnerable to hacking (think of the backdoor scandals or the Intel issue where hardware security was being bypassed, giving access to software); hardware is still not sustainable and e-waste is the fastest growing waste (few countries have robust sourcing and recycling plans); and, more importantly, hardware is also a complicated system for data transfer with multiple jurisdictions (cell phone towers, cables, other computers). The gap in time to connect one device to another may be minuscule for you (it is in milliseconds), but that is enough time for hackers to hit vulnerable systems, especially when software is not updated.

Hence, AI needs a systems perspective – where you consider data, software, and hardware. Even if you outsource some of these, do the due diligence and ask the right questions. In government digitalization, purchasing has been problematic because historically, they have bought on price and do not have the basic knowledge of AI systems. Outsourcing is not the answer; it is a poor shortcut.

We need basic awareness and education of AIS, without which we cannot govern them. This education is NOT just digital skills but a fundamental understanding of how these systems work.

AI versus Human Intelligence

The recent letter to have a pause on AI, signed by Silicon Valley luminaries, assumes AI is close to being as intelligent as a human. When the term Artificial Intelligence came out, it referred to machines that mimic human intelligence.

Let me deconstruct this for you by looking at five points:

Very frankly, across these parameters: the human being WINS.

Comparison between AI and Humans

Multiple types of intelligence: spatial, linguistic, musical, logical/mathematical, existential, interpersonal, kinesthetics, and intra-personal [Gardner & Hatch, 1989).

Constantly learning and relearning.

Has the capacity to create ne-knowledge.

The brain runs on 23 Watts of energy.

Human Brain – 86-100 billion neurons, 1000 Trillion??? Con-nections.

Our DNA can store all the knowl-edge in the world accumulated by 2025 in 4 Evian bottles of 500 ml each!

AI is still relatively narrow – it depends on what it can be trained on.

Needs human supervision.

Has the capacity to create new data by combining old data sets and following instructions or looking at patterns.

The average computer consumes 600 kWh per 8 hours a day.
Transformer systems like Chat GTP run on training. The model used in January 2023 was be-tween 1.1 to 23 million KWh!!!

540 billion parameters (connections).

The number of Data servers would have to increase, and they would create 3.2% of the world's greenhouse gases!

AI is not equal to human intelligence, but the challenge is that we are defaulting decisions to AI like it is. This practice is dangerous as we are giving the AIS agency on the assumption that it is intelligent for various reasons – misinformation, hype, or poor awareness of AIS. This needs to change.

AIS Governance (Artificial Intelligence Systems)

There have been many AI governance initiatives. However, there is no agreement worldwide on what it means and the relevance of the recommendations and their implementation. Partly because this is a very competitive space. According to OECD, there are 800 AI policy initiatives from 69 countries, territories, and the EU. Another challenge is that you need an "agile government" approach towards policymaking as what you may have recommended or deemed as necessary may be obsolete. Take, for example, the EU classification of the risk of AIS (based on the OECD classification). They had assumed chatbots are low risk, but with the introduction of ChatGPT, they will need to rethink this.

Another challenge seen with evolving regulations and research spending is that the North and West have more power over the global South and parts of the East. Here it is important to acknowledge that a significant amount of funding into AIS has been defense funding. Hence, surveillance and security issues are sometimes more prevalent than human and planet flourishing.

We also see that some countries have better connectivity from an infrastructure point of view. How would this inequality be addressed? An additional but significant issue is that AIS has often been developed in countries with aging populations, like the EU, North America, China, and Japan, as they DO NOT have a replenishing workforce. So job obsolescence and the fallacy that we are bettering people's lives by taking away routine jobs is not always true (especially if they have no future job replacement, and lose income or if the workloads have not decreased). Human beings have different skills: some paint, some love numbers, and some are good with their hands. Why are we taking away their jobs? It takes 10-15 years to reskill, and the older they are, we have the additional burden of old age and pension, which we must fix first, along with education making adopting of AIS a wicked problem. Think of the recent tech layoffs coming out a pandemic and into a recession.

Further, much of the AIS is developed using human knowledge (how we do accounting, putting in a screw, painting, writing, etc.). Using human talent as part of the AIS supply chain and making humans redundant becomes an ethical issue. One of the Universal Human Rights is the Right to Work. Also, some countries with young populations – Sub-Saharan Africa, Middle East North Africa, and parts of Asia - need more job creation, some of which will be in low-skill areas. Governments need to consider this carefully.

So what does high governance in AIS mean: it means the ability to steer through the future and the opacity of AIS and its impact (positive and negative OR intentional or unintended).

We also need to understand the vulnerabilities of these systems and the maliciousness of intent – it is estimated daily we have 30,000 zero-day moments or vulnerabilities which open our software to hacking. These attacks will not decrease but only escalate with greater AI adoption.

AIS needs anticipatory governance – how can we predict AIS consequences and be future-ready to optimize the positive impact, mitigate the negative, and have adequate guardrails.

We don't have time to waste! The scale and speed of adoption of new tech are accelerating. ChatGPT took only five days to reach 1 million customers. When the EU was crafting its massive GDPR regulation (it is 261 pages, and lawyers need lawyers to understand it), it took years. GDPR was proposed in 2012, adopted in 2016, and implemented in 2018. **WE DO NOT HAVE THIS LUXURY OF TIME.**

I do not think the responsibility of AIS governance is that of governments alone – it requires massive public education to ensure we all understand and contribute to good governance. Furthermore, it is not just tech companies that are responsible – it is any company that adopts AI, recommends AI, or researches AI. There is no excuse!

Challenges to AIS

I will introduce you to two decision-making scales from our latest book which is coming soon. The first is the ability for rational decision-making and a comparison between humans and AI. A human by nature, and when compared to an AIS, thinks more slowly and intuitively (contrary to what we think!). We need to think about what unique skills humans and AI bring toward so-called "rational" decision-making (first of all, humans will never be entirely rational, so the data we have may always be biased) Just because we cannot explain intuition (we still know very little about the brain) does not mean it is irrelevant?

A human can never compete with a machine for task that require speed of computations – but why should we? We must acknowledge the unique skills humans can bring to a team of AI and humans (see another article here). When you combine teams of humans and machines, you want to ensure they complement each other.

Why is it we will accept a "black box" decision of a computer or Machine Learning, because of the number of parameters but not a human's intuition?

When AIS depends on other AIS systems, the matter tends to get exacerbated, leading to AIS group think that quickly spirals downwards (have you tried getting two chatbots to talk to each other?). This was seen with the 2010 Flash Crash, where a trader created an algorithm to gamify the system to think share prices were dropping, and this triggered panic selling as it got flashed on other computerized systems. So much so that the S&P 500 plummeted 5% or 600 points in four minutes, temporarily erasing US$1 trillion in value, and the Dow Jones Industrial average, fell 1000- points in 10 minutes. This time around, the stock markets have introduced “circuit breakers” or a cooling down period. Human-in-control period!

We need to acknowledge the role of intuition in decision-making. So-called 'rational thinkers' may not like this word, but we know little about how the human brain "thinks." Let me illustrate this point with another example. In the Cold War days, USSR set up an early defense warning system of 40 satellites called Oko. The period was tense as USSR had just shot a South Korean airliner down, killing 296 people, and NATO had escalated its military exercises. Stanislav Petrov, a lieutenant colonel of the Soviet Air Defence Forces, was in charge when there was a warning on the Oko system, that the USA had launched an intercontinental nuclear missile. The system was new, and he hesitated to inform his superiors, even though that was the protocol, knowing it would result in a nuclear war. A few minutes later, the system blasted again, signaling four more missiles. He never informed his superiors (these missiles would have taken 15 minutes to reach USSR). Afterwards, they found out that the system had malfunctioned, mistaking sunlight reflections from clouds as a missile. Was it intuition? While he can never explain at that moment what was in his head. Lurking inside, was the fact that the USA had more than 35,804 nuclear warheads in 1983, and later he said - why would they just launch one?

When we think of the role of AIS in decision-making, we need to consider whether AIS should assist, augment or replace human decision-making. Gartner recommends that in complicated, complex, or chaotic situations to use human decision-making with the AIS assisting or augmenting, NOT replacing. Yet we have let AIS make decisions for us daily (since they are perceived as low risk), with social feeds, movie recommendations, automated cars, text completion, friend recommendations, or navigation tools. What skills are we making obsolete?

The second scale that will help us govern AI better is one of Accountability In case of failure, who is responsible? Here we refer to the human in in charge (so they have authority) and hence need to be accountability. This accountability is at multiple levels: top management to the technical departments and even the supply chain. The Boeing MCAS system is a great example. This accountability space was complex as you needed accountability by Boeing (senior management and technical staff), customers, FAA, contractors, and even the individual governments involved. Perhaps if accountability were assigned at the design stage, we would not have the tragic loss of people (more on this case later).

Opportunities and Trade-offs

There is tremendous potential in AIS – it is estimated to contribute US$2 trillion (Statistica) to US$15.7 trillion by 2030 (PwC), disproportionately benefiting North America, Europe, Developed Asia, and specifically China.

A trade-off is a compromise, and we make many such compromises when we adopt AIS systems. For example, jobs. McKinsey & Company reckons that, depending upon various adoption scenarios, automation will displace between 400 and 800 million jobs by 2030, requiring as many as 375 million people to switch job categories entirely. A study finds that many new jobs AIS creates are lower paying (which makes no sense in times of recession and where savings are not enough for retirement!), especially when you think of these AIS being developed on human knowledge and talent! Another trade-off could be sustainability. As we adopt more AIS, we create an enormous negative impact on the environment – this includes extraction, manufacturing, transportation, and of course, recycling (which we do not do well). Significant investments need to be made in R&D – for example, hardware efficiency will need to double faster than every 1.1 years to keep emissions under the recommended levels.

Future-Proofing AIS

Future-proofing AIS for a high level of governance needs the active participation of individuals, entities, and governments.

The first step is to outline your values. This is not a complicated process – there should ideally just be two:

(i). Human-Centric – for the benefit of the human
(ii). Planet-centric – for the benefit of the planet

You can decide the scope of how broad you would like it to be but be ruthlessly honest. Governance cannot be adopted as part of an organizational culture if you are not genuine.

You need to understand where the human-in-the-loop is (for decision-making) and in the supply chain. Is the human a means to profit, to support AI till it can do the task that we think may be better than the human, or is AI and money (as a resource) being used to ultimately help humans and the planet?

It was disappointing to see that Open AI was initially not very transparent about the profit intensions, and there was an assumption they were non-profit, which they have now clarified on the website: "We are governed by a nonprofit and our unique capped-profit model drives our commitment to safety. This means that as AI becomes more powerful, we can redistribute profits from our work to maximize the social and economic benefits of AI technology….[they explain this as follows]… The fundamental idea of OpenAI LP is that investors and employees can get a capped return if we succeed at our mission, which allows us to raise investment capital and attract employees with startup-like equity. But any returns beyond that amount—and if we are successful, we expect to generate orders of magnitude more value than we'd owe to people who invest in or work at OpenAI LP—are owned by the original OpenAI Nonprofit entity. Going forward (in this post and elsewhere), "OpenAI" refers to OpenAI LP (which now employs most of our staff), and the original entity is referred to as "OpenAI Nonprofit."

Second, we need to be Agile, as whatever you do, the future will throw a curveball at you, and all you can do is acknowledge it and get the best minds to work on a solution quickly.

Third, here is a proposed governance framework:

Start with accountability. Make people responsible and have discussions on possible failures so they will go out of their way to ensure better AIS. This accountability is within the organization and across the supply chain.

Next,

decide the level of transparency; too often, engineers do not know what they are working on – just small bits of code and hence cannot design the best systems.

Be honest about why these levels of transparency are there (sometimes it's just plain ignorance or because things are outsourced).

Work on responsible systems, ensuring the data, software, and hardware are robust.

Part of accountability is secure AIS.

The security costs are expensive but critical.

When you have accountability – **think of intent** – why are you building the AIS? Be honest – to make money or as Open AI said to benefit all humanity (not sure how this is being done). When Uber started, they said it was to give a person additional income, earning extra money driving. This assumption is no longer valid, and contexts may change, so honesty is critical here.

For good governance, we need to decide human-in-the-loop, and where possible, this goes back to understanding decision-making and human agency. Next, data management is critical, especially as many of the challenges we face today are data related. Then focus on algorithm governance and finally the system robustness. Remember when employees of Meta were locked out of their offices, when the system failed due to a DNS server issue? Sometimes you need backups that are not high-tech.

There are many AIS frameworks, but before getting lost in them – think about the above critical points. It may help you focus on what really matters – humans and the planet.

كلية محمد بن راشد للإدارة الحكومية
MOHAMMED BIN RASHID SCHOOL OF GOVERNMENT

The reading society is a civilized society, and whoever supports reading supports the industry of civilization, supports a knowledge economy, and supports building generations that build the future of the Emirates

His Highness Sheikh

Mohamed bin Rashid Al Maktoum

Author

Melodena Stephens

Professor of Innovation Management
Mohammed Bin Rashid School of Government

Editor

Aisha Sultan Al Shamsi

Director of Executive Education
Mohammed Bin Rashid School of Government

التعليم التنفيذي

Executive Education

+971-4-3293290
execed@mbrsg.ae.ar
www.mbrsg.ae
/mbrsg
/mbrsgae
/+mbrsgae
/mbrsg
mbrsgae
/company/mbrsg

جميع الحقوق محفوظة 2023 © كلية محمد بن راشد للإدارة الحكومية

يمنع نسخ أو حفظ أو إعادة إنتاج أي جزء من هذه المطبوعة أو إرسالها بأي شكل أو طريقة (إلكترونيا، بالتصوير أو التسجيل أو بأي طريقة أخرى) من دون الموافقة الخطية من الكلية.

يستثنى من ذلك الاقتباسات الموجزة التي يتم تضمينها في البحوث والمراجعات وبعض الاستخدامات غير التجارية الأخرى المسموح بها بموجب قوانين حقوق النشر