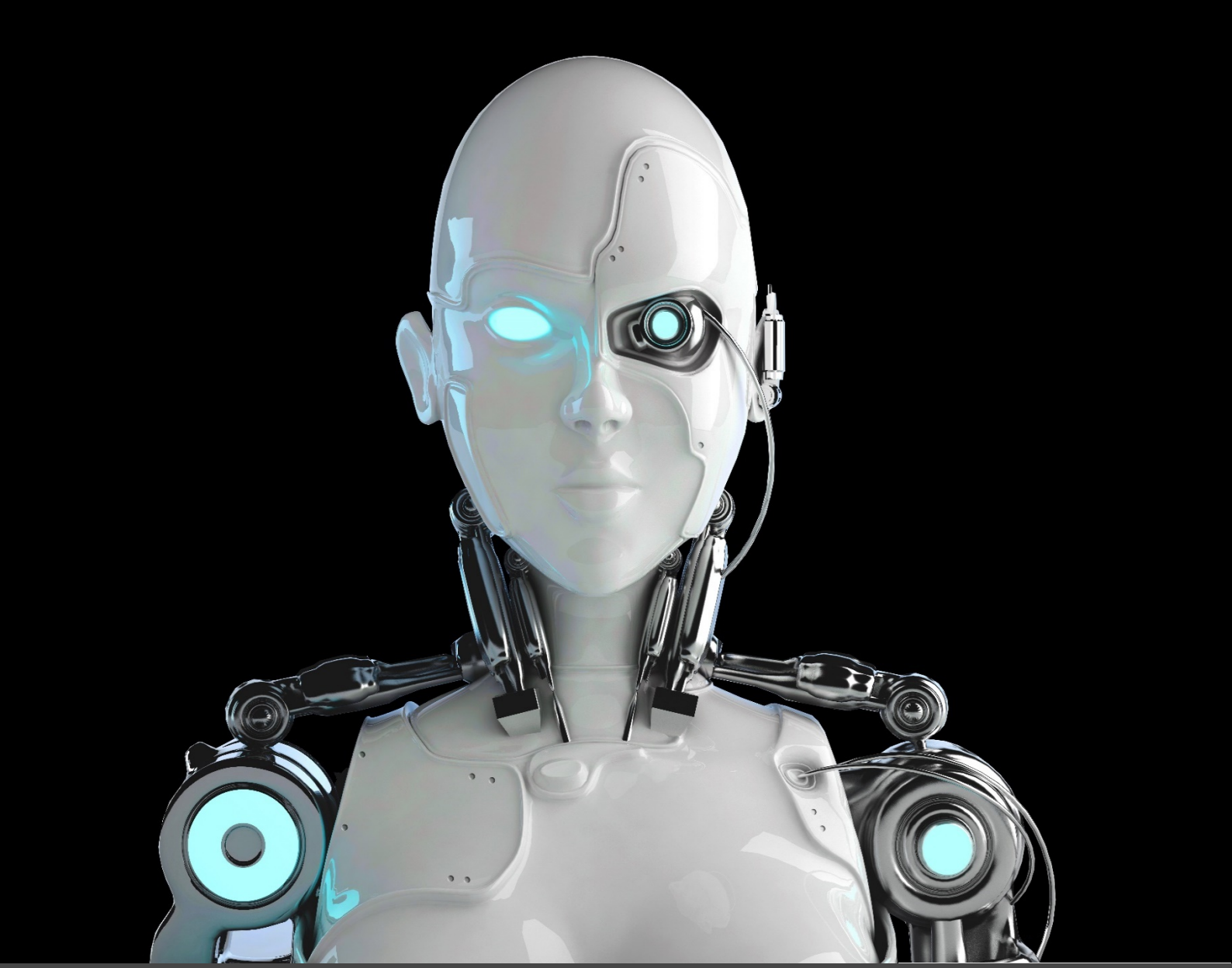


# Comprehensive AI governance needed

Yann Bonduelle, Batsuren Enkhtur, Maria Axente, Ilana Golbin, Anand Rao, Kevin Allison, Paul Triolo

pwc

eurasia group politics first

Hiroyuki Yoshida[1] spent his life toiling in the orchards around his family home in Japan's Yamanashi prefecture. Although he raised a family on his modest income, it was backbreaking work. Today, at the age of 92, Mr. Yoshida is bedridden and dependent on his wife, Michiko, for his daily care. An exuberant man who used to enjoy playing with his great-grandchildren, Mr. Yoshida has become withdrawn, as his physical health has worsened. Michiko worries that her husband may no longer be able to stay in their family home.

Mr. Yoshida's eyes lit up recently when he heard about an engineer who had developed a prototype for an ambitious new exoskeleton that could help people with mobility problems learn to walk again. The exoskeleton used artificial intelligence (AI) to anticipate changes in terrain, to adapt to a user's unique gait and to respond to subtle, intentional movements to change direction. Intrigued, the family contacted Mr. Suzuki, the engineer. Mr. Suzuki and Mr. Yoshida's family agreed to try the prototype.

After a few months of tuning the exoskeleton to Mr. Yoshida's size, the first test went well: Mr. Yoshida was able to walk a few steps. However, the third test ended with a bad fall that took Mr. Yoshida back to bed, this time for good as he had seriously damaged his knees and broke a few ribs. At the news of the accident, an enquiry took place.

The enquiry found three main causes of the accident: First, the exoskeleton's AI balancing algorithm had been trained indoors over thousands of hours by more than 100 young volunteers with sport or traffic accident injuries who wanted to learn to walk, run and play sports again. Mr. Yoshida was the first older person to try the exoskeleton. The enquiry found that the exoskeleton responded too quickly to Mr. Yoshida's movements, which led him to overreact and lose balance. In addition, on the day of the third test, the ground where the test took place was slightly wet, a condition that had never been experienced in the training data. Finally, it appeared that Mr. Yoshida had been emboldened by the successes of the first two tests. He was less cautious than he had been and should have been, given the wet conditions.

In its conclusions, the enquiry was tough on Mr. Suzuki. It recognized that he had followed accepted robotics development processes, but noted he had used no AI governance for the AI used in the exoskeleton. The enquiry was particularly concerned about the disparity between the conditions under which the training data were captured (young people learning to walk and run indoors after accidents), and the specific conditions of the work done with Mr. Yoshida (helping a senior citizen do limited walking outdoors), concluding that this amounted to using biased data to feed the AI. Without a formal approach to check the AI's ability to adapt to a different context and to different conditions, Mr. Suzuki's investors felt it was too risky to continue development without significant changes, adding costs and delays, before an eventual commercial launch of the exoskeleton. Given this new context, investors pulled out. While Mr. Yoshida's family were hoping to get him to become progressively self-sufficient at home, he ended up diminished and injured, and eventually moved into a nursing home.

This fictitious story poses important real-world questions directly relevant to helping makers of AI-powered applications – which we call Smart Tech applications[2] – maximize the benefits of their technologies while minimizing associated risks.

Many Smart Tech applications will be deployed over the coming decade, designed to help billions of people live longer, healthier and happier lives, including people like Mr. Yoshida. Like any innovative technology, Smart Tech applications may cause harm. Good governance that maintains public trust in Smart Tech in general and its AI component in particular is critical, not only because people's lives and wellbeing are at stake, but also because a safe, vibrant and innovative Smart Tech industry is one of the best hopes for tackling some of the world's biggest challenges.[3]

In this paper, we offer a new, comprehensive way to think about AI governance, designed to build and strengthen public trust in Smart Tech applications. We believe adopting a holistic, open, proportional and end-to-end AI governance will help minimize risks of failures or malfunctions of AI in Smart Tech applications.

This paper is written for executives working to use AI in their operations, or in new products or services.

We have two objectives:

1. Help readers understand the importance of AI governance and how they can contribute to building and strengthening public trust in AI and Smart Tech applications.
2. Offer a practical AI governance framework to ensure that Smart Tech applications are developed and deployed in a manner that is safe and ethical.

[1] In this paper, when the text is in a light gray box, it means that we are writing about a fictitious story.
[2] Throughout this paper, "Smart Tech applications" refers to combinations of AI with emerging technologies such as Robotics, Drones, Blockchain, 3D printing, IoT, AR and VR. These technologies have been identified by PwC as the Essential Eight. In addition, when we mention AI within a Smart Tech application, we mean the autonomous and intelligent system within the Smart Tech application.
[3] The Global Risks Report 2020, WEF (accessed 21 Jan, 2020) https://www.weforum.org/reports/the-global-risks-report-2020

# Why AI governance matters

## The Macro view

From raging wildfires and typhoons to global epidemics and riots, every day brings more evidence of the disruptions being fueled by climate change, water crises, resource scarcity and demographic and social changes.[4] These Global Megatrends threaten human well-being, economic growth and the environment. Fortunately, it looks like technology can help.

Few technologies hold more promise to help people, businesses and governments address these challenges than AI. For example, AI-powered products and services have the potential to lead to new medicines,[5] speed the transition to a low-carbon economy,[6] and help people enjoy dignity in retirement and old age.[7] The economic gains alone could be enormous. PwC estimates that by 2030, AI could contribute an additional US$ 15.7 trillion to the global economy.[8]

Yet for many people, AI also sparks fear and unease. Concerns such as speculation about "killer robots," job losses due to automation, deep fakes or privacy fears related to facial recognition have already become part of the public discussion about the benefits and risks of AI.

The concerns are not just theoretical. AI-related failures have been implicated, among several other factors, in the death of a pedestrian who was struck by a self-driving car undergoing testing in Arizona in 2018.[9] In October 2019, an article in the journal Science alleged that a software program used to help determine medical care for more than 200 million Americans exhibited systematic racial bias.[10]

AI will increasingly automate tasks that today typically require human intervention and judgement. Errors or biases stemming from the AI included in Smart Tech applications may go unnoticed until they have already caused harm. Beyond the direct cost caused by an AI malfunction, there is also the indirect impact that the failure of a major technical system leveraging AI could cause. This could substantially undermine public trust in AI technologies, thereby slowing down adoption of technology solutions to society's problems.

From nuclear power to genetic engineering, history is full of examples of new technologies where safety risks or concerns about their impact on society have led to a political backlash that affected how those technologies were developed and deployed. These risks are particularly acute for AI, which is coming of age in a much more volatile political environment than previous generations of technology innovations and has the potential to impact more industries and people, than technologies before. Recently, many facets of AI have come under increasing public and political scrutiny. These include:

- **Public concerns about data privacy and jobs.** The way businesses and governments collect and use data, an essential component of AI, is under intense public scrutiny following a series of high-profile privacy scandals and growing concerns about potential abuse of personal information. Meanwhile, at a time of rising concern over income inequality and economic stagnation, AI's potential effects on employment have become an important political issue.
- **Regulatory concerns about safety and the market power of big tech companies.** Governments, industry and civil society around the world are accelerating their efforts to develop ethical guidelines and technical standards for safe and responsible AI. At the same time, large technology companies that are working at the forefront of many new Smart Tech applications are under intense scrutiny from regulators concerned about their growing market power and control over large amounts of personal data.
- **Security concerns about national and economic security and geopolitical competition.** AI and the ICT[11] systems that make it work, including cloud computing, semiconductors, and underlying data, are at the center of a growing geopolitical competition between the US and China. They are also increasingly seen as matters of economic and national security by other governments around the world, with policymakers in major democracies concerned that AI-enabled applications could help strengthen authoritarian systems.

---

[4] The Global Risks Report 2020, WEF (accessed 21 Jan 2020) https://www.weforum.org/reports/the-global-risks-report-2020
[5] AI and robotics are transforming healthcare, PwC 2017 (accessed 22 Jan 2020) https://www.pwc.com/gx/en/industries/healthcare/publications/ai-robotics-new-health/transforming-healthcare.html
[6] Harnessing artificial intelligence for the Earth, PwC 2019 (accessed 22 Jan 2020) https://www.pwc.com/gx/en/services/sustainability/publications/ai-for-the-earth.html
[7] Realizing Society 5.0, The Government of Japan 2017 (accessed 22 Jan 2020) https://www.japan.go.jp/abenomics/_userdata/abenomics/pdf/society_5.0.pdf
[8] Sizing the Prize: What's the real value of AI for your business and how can you capitalise?, PwC 2017 (accessed 24 Jan 2020) https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf
[9] It's 2020. Where are our self-driving cars?, Vox 2020 (accessed 15 Feb 2020) https://www.vox.com/future-perfect/2020/2/14/21063487/self-driving-cars-autonomous-vehicles-waymo-cruise-uber
[10] Dissecting racial bias in an algorithm used to manage the health of populations, Science 2019 (accessed 31 Jan 2020) https://science.sciencemag.org/content/366/6464/447
[11] ICT stands for Information and Communications Technology

This is a high-stakes environment in which to innovate. Yet as Smart Tech applications are applied across a growing cross-section of the global economy, the risk of poorly implemented AI causing economic, psychological, or physical harm, or even loss of life, will increase. With a major incident, the public and regulatory backlash that follows could damage the trust earned by many successful applications of AI. Like with Mr. Yoshida's and Mr. Suzuki's story, this backlash could slow investments in the sector and curtail innovation in AI and related Smart Tech applications that might otherwise help address pressing global challenges such as an ageing society, food crises and extreme weather disasters.

Good AI governance – the best practices, rules and guidelines for how businesses and others working on Smart Tech applications should use data, check for flaws, and ensure accountability in the deployment of AI – can help reduce the risk of bad outcomes for people and society, while also building trust in AI and Smart Tech applications. That is why AI governance matters.

## Smart Tech applications

With relentless advances in many emerging technology areas, the potential of Smart Tech applications to address pressing global challenges is vast. The opportunities to work in groundbreaking fields and help society at the same time make it an exciting time for emerging technology communities. Care robots and voice assistants for seniors are just a few examples of Smart Tech applications that are increasingly being adopted, improving the quality of life for a growing segment of developed countries: the elderly population.[12]

But there is also apprehension. There have been well-publicized cases where Smart Tech applications have caused trouble.[13] For instance:

- Issues with unforeseen malfunctions, such as an accident with a security robot[14]
- Misuse of data, such as harvesting private conversation records from voice-activated digital assistants[15]
- Issues with structural changes, such as mass displacement of workers by smart automation[16]

Adding to this apprehension are issues with the "hype" associated with the most advanced AI technology innovations, such as concerns that emotion-recognition technologies designed to assess a user's state of mind may lack a sound scientific basis.[17]

Fortunately, by emphasizing a comprehensive approach to AI governance, companies using AI in Smart Tech applications can mitigate risks of malfunction and misusage while users and broader society benefit from progressive adoption of effective Smart Tech applications. This is why AI governance matters.

## AI governance so far

Over the last few years, governments, universities, NGOs and large companies have begun developing high-level frameworks and guidelines for AI governance to ensure that risks associated with Smart Tech applications can be mitigated without unnecessarily impeding innovation.

These efforts are arguably most advanced within the private sector, both within individual companies and through multi-stakeholder organizations. For example, the Partnership on AI[18] was founded in late 2016 by Amazon, Facebook, Google, DeepMind, Microsoft and IBM. It now includes more than 100 partners from academia, civil society, industry and nonprofits. Company-specific guidelines and principles have also been published by major tech companies and are already applied to guide the development of AI research and applications.

Non-governmental bodies, such as the International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC) and the IEEE,[19] are working on their own concepts for standards around AI, incorporating both ethics and governance criteria. The World Economic Forum Centre for the Fourth Industrial Revolution engages with fellows from different organizations globally to raise important issues and propose practical guidance for emerging technologies, like the recently released Artificial Intelligence Toolkit, designed for corporate boards.[20]

---

[12] Japan's robot revolution in senior care, The Japan Times 2019 (accessed 2 Feb 2020) https://www.japantimes.co.jp/opinion/2018/06/09/commentary/japan-commentary/japans-robot-revolution-senior-care/

[13] Thinking About Risks From AI: Accidents, Misuse and Structure, Remco Zwetsloot, Allan Dafoe, Lawfare Blog 2019 (accessed 2 Feb 2020) https://www.lawfareblog.com/thinking-about-risks-ai-accidents-misuse-and-structure

[14] Robot runs over toddler in shopping centre, BBC 2016 (accessed 2 Feb 2020) https://www.bbc.com/news/technology-36793790

[15] More privacy missteps cast cloud over voice-activated digital assistants, The Japan Times 2019 (accessed 15 Feb 2020) https://www.japantimes.co.jp/news/2019/08/03/business/tech/privacy-missteps-cast-cloud-voice-activated-digital-assistants/

[16] Will robots really steal our jobs?, John Hawksworth, Richard Berriman, Euan Cameron, PwC 2018 (accessed 31 Jan 2020) https://www.pwc.com/hu/hu/kiadvanyok/assets/pdf/impact_of_automation_on_jobs.pdf

[17] AI Now 2019 Report, AI Now Institute (accessed 20 Jan 2020) https://ainowinstitute.org/AI_Now_2019_Report.pdf

[18] Partnership on AI (accessed 28 Jan 2020) https://www.partnershiponai.org/about/

[19] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, IEEE (accessed 27 Jan 2020) https://standards.ieee.org/industry-connections/ec/autonomous-systems.html

[20] Artificial Intelligence Toolkit Helps Companies Protect Society and Their Business, WEF 2020 (accessed 5 Feb 2020) https://www.weforum.org/press/2020/01/artificial-intelligence-toolkit-helps-companies-protect-society-and-their-business/

Government efforts to put guardrails around AI are picking up, although not evenly.[21] Some of the most notable efforts include:

- G20 members signing a set of AI principles at the G20 2019 summit hosted by Japan, with the goal of promoting the development of international rules for AI developers[22]
- 42 countries adopting OECD Principles on Artificial Intelligence in 2019[23]
- The European Commission preparing to introduce AI legislation[24] and publishing a whitepaper[25] that builds on ongoing work by the High-Level Expert Group (HLEG) on AI. The European Commission will want its new legislation to have long-reaching global effect, akin to GDPR[26]
- Japan's Ministry of Economy, Trade and Industry (METI) publishing contract guidelines on utilization of AI and data released for corporations[27]
- The UK government releasing draft guidelines for AI procurement to help government agencies purchase AI better.[28]

Many organizations, including PwC, have been developing AI governance methods and tools to facilitate practical implementation of AI governance. These include checking automatically AI models and data for potential biases, assessing the robustness of AI models and novel approaches to improve the interpretability of outcomes of AI algorithms.

We have selected in this paper some of the most important initiatives looking at aspects of AI ethics and governance. While there has been a lot of work in these areas, significant contributions are accelerating, so it is difficult to keep up. One of the most recent visual displays of contributions on AI ethics and governance was shown a few months ago by Fluxus Landscape, a Center for Advanced Study in the Behavioral Sciences (CASBS) at Stanford University partnership with Şerife Wong, under the form of a collaboration platform,[29] illustrating how fast innovation in the AI governance field is moving.

We have attempted to show a simple table (see below with the list of references overleaf) detailing the range of methods and tools available, from understanding ethical principles to applying specific AI governance technical tools, whether they come from pan-national organizations, such as the UN or the EU, or from a company's own internal audit team.

**AI governance methods and tools**

| AI governance bodies | | Ethical principles | Guidelines, recommendations | Laws, standards, contracts, certifications, applied ethics | AI governance technical tools |
|---|---|---|---|---|---|
| Legislators and regulators | Pan-national organizations | OECD (8) | EU (4) | N/A | N/A |
| | National governments | Singapore (10) | Japan (7) | Canada (2) | N/A |
| | Local governments | Dubai (3) | N/A | California (1) | N/A |
| Other external governance bodies | Industry organizations | Radiology associations (9) | IEEE (5) | N/A | N/A |
| | AI application standard and quality certification services | N/A | ISO (6) | N/A | N/A |
| Internal governance bodies | Internal governance teams | PwC (11) | PwC (11) | PwC (11) | PwC (11) |

[21] Government AI Readiness Index, Oxford Insights and the International Development Research Centre (accessed 23 Jan 2020) https://ai4d.ai/index2019/
[22] G20 ministers agree on guiding principles for using artificial intelligence, Japan Times 2019 (accessed 23 Jan 2020) https://www.japantimes.co.jp/news/2019/06/08/business/g20-ministers-kick-talks-trade-digital-economy-ibaraki-prefecture/#.XiqqOcgzZPY
[23] Forty-two countries adopt new OECD Principles on Artificial Intelligence, OECD 2019 (accessed 23 Jan 2020) https://www.oecd.org/science/forty-two-countries-adopt-new-oecd-principles-on-artificial-intelligence.htm
[24] Decoded: Europe's AI cacophony — Human rights enter AI debate — Who's liable for AI harms?, Politico 2019 (accessed 21 Jan 2020) https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-europes-ai-cacophony-human-rights-enter-ai-debate-whos-liable-for-ai-harms/
[25] White Paper on Artificial Intelligence: a European approach to excellence and trust, The European Commission 2020 (accessed 22 Feb 2020) https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en
[26] General Data Protection Regulation (GDPR), PwC (accessed 24 Jan 2020) https://www.pwc.com/gx/en/issues/regulation/general-data-protection-regulation.html
[27] Contract Guidelines on Utilization of AI and Data, Ministry of Economy, Trade and Industry of Japan 2018 (accessed 23 Jan 2020) https://www.meti.go.jp/english/press/2019/0404_001.html
[28] Draft Guidelines for AI procurement, UK Government 2019 (accessed 23 Jan 2020) https://www.gov.uk/government/publications/draft-guidelines-for-ai-procurement
[29] AI ethics and governance map platform, Şerife Wong, CASBS (accessed 27 Jan 2020) https://icarus.kumu.io/fluxus-landscape

(1) **California** AB-2269 Personal rights: automated decision systems, https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB2269

(2) **Canada** Canada's CIO Strategy Council publishes national AI standards https://www.itworldcanada.com/article/canadas-cio-strategy-council-publishes-national-ai-standards/422722

(3) **Dubai** Artificial Intelligence Principles & Ethics https://www.smartdubai.ae/initiatives/ai-principles-ethics

(4) **EU** White Paper on Artificial Intelligence: a European approach to excellence and trust, The European Commission 2020 https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

(5) **IEEE** The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems https://standards.ieee.org/industry-connections/ec/autonomous-systems.html

(6) **ISO** ISO/IEC JTC 1/SC 42 https://www.iso.org/committee/6794475.html

(7) **Japan** Contract Guidelines on Utilization of AI and Data, Ministry of Economy, Trade and Industry of Japan 2018 https://www.meti.go.jp/english/press/2019/0404_001.html

(8) **OECD** Forty-two countries adopt new OECD Principles on Artificial Intelligence, OECD 2019 https://www.oecd.org/science/forty-two-countries-adopt-new-oecd-principles-on-artificial-intelligence.html

(9) **Radiology associations** Ethics of Artificial Intelligence in Radiology: Summary of the Joint European and North American Multisociety Statement https://pubs.rsna.org/doi/10.1148/radiol.2019191586

(10) **Singapore** Singapore Model AI Governance Framework Second Edition https://www.sgpc.gov.sg/sgpcmedia/media_releases/imda/press_release/P-20200122-2/attachment/Singapore%20Model%20AI%20Governance%20Framework%20Second%20Edition%20-%20Framework.pdf

(11) **PwC** Responsible AI Toolkit https://www.pwc.com/rai

So far, most AI governance methods and tools have been designed for creators of AI applications (principles, guidelines, technical tools) and providers of data (mainly GDPR or similar legal frameworks as well as some technical tools). Considering the increasing complexity of the AI powering Smart Tech applications and the accelerating speed of their adoption, now is the time for a more comprehensive and pragmatic AI governance. Involving all stakeholders will yield better results than relying solely on AI architects and data purveyors. This is why AI governance matters.

# Comprehensive AI governance

In this section, we introduce a new way to look at AI governance: holistic, proportional, open and end-to-end.

## Holistic AI governance

A holistic, or systemic, AI governance process includes all stakeholders involved in the design, development, usage and maintenance of a specific AI-powered Smart Tech application.

The process of designing, developing, using and maintaining a Smart Tech application requires contributions from many stakeholders:

- Investors provide funding
- Smart Tech application makers and sellers, including AI platform providers, provide workforce and investments
- Data aggregators and providers provide data, workforce and investments
- Smart Tech application users provide data and feedback and often pay for usage
- Legislators, regulators and governance bodies provide guidance and controls
- The public at large provide feedback.

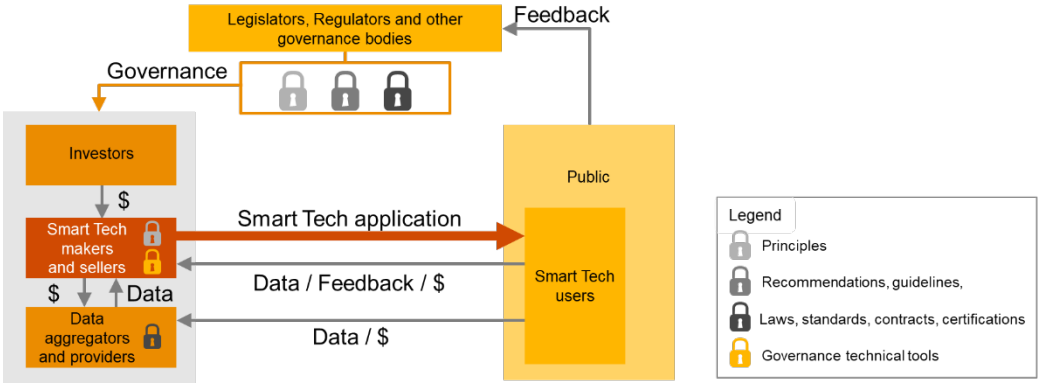
All stakeholders directly or indirectly either influence, or are affected by, Smart Tech applications launched in the market.

Much of the debate around AI governance has so far centered on data usage and AI algorithm development, focusing on the responsibility of data aggregators and AI designers and developers.

It is well understood that developing and deploying a successful Smart Tech application requires first the use of the right data from accurate, fit for purpose, complete or as unbiased as possible data sets. Then it requires developing and training an algorithm whose actions can ideally be sufficiently explained to individuals and organizations who need to trust its conclusions.

There has not been much discussion of the contributions of other stakeholders to the success of a Smart Tech application. We have defined success as (1) avoiding failures of Smart Tech applications and (2) contributing to building people's trust in these applications.

The diagram below shows the interactions among stakeholders and what they exchange: data, feedback, application, governance and money. The lock icons in the figure indicate the type of AI governance methods and tools usually applied in the design, development, usage and maintenance of Smart Tech applications.




When all stakeholders play their part and share the common goal of designing, developing, using and maintaining Smart Tech applications to avoid failures of AI and build society's trust in Smart Tech applications, the odds of AI making a positive impact increase.

We suggest redefining the scope of AI governance so that all stakeholders involved in the design, development, usage and maintenance of a Smart Tech application are fully aware of their own responsibilities, what AI governance they should conduct, when and how.

Concerted efforts among stakeholders are vital to ensure mitigating risks and building long-term trust in a Smart Tech application. We might see in the future convergence of specific standards and regulations regarding responsibilities of stakeholders. This would be to ensure there are as few gaps in responsibilities and conflicting roles as possible. In the meantime, stakeholders will be able to rely on proliferating assessment tools, such as the ones PwC has developed in our Responsible AI Toolkit.

A holistic view of AI governance makes it easier to review the overall governance conducted by all stakeholders involved in a Smart Tech application.

With a holistic AI governance lens, it becomes clear that active feedback from users and the public is needed to improve Smart Tech applications and to develop appropriate rules, guidance and laws. Emphasis on data privacy and individual choice, for example, took a long time to come. Most recently, if the public had not voiced concern about data privacy issues caused by some well-known voice assistants, there would not be any option available for opting out of data harvesting. It could be argued that earlier measures and legislation could have minimized or delayed the current tech backlash.

Investors, especially those in early VC funding rounds, have the power to influence funded companies to various degrees. In the last few years, for example, some investors have taken steps to prioritize ethical behavior along with financial returns, helping to mitigate some decisions made by startup management.[30]

## Open AI governance

An open AI governance process provides visibility into AI governance to stakeholders. This does not mean that the AI itself is to be made visible to stakeholders.

We have shown that a holistic view of stakeholders helps identify stakeholder's responsibilities in ensuring the successful deployment of an AI-powered Smart Tech application. A holistic view also makes it easier to check that all responsibilities are met and whether there are gaps of governance that need to be filled. For example, is there a user organization in place to collect user feedback? Is the user organization providing fair and balanced feedback? Is the Smart Tech application seller addressing the feedback? Who checks that these governance gaps are closed quickly? Who is responsible if they are not closed? What processes are in place to incentivize compliance and shared responsibilities? Are remedies needed to ensure non-compliers comply?

[30] Bankrolling Ethics: Do Tech Investors Have a Responsibility to Protect Democracy?, Harvard University 2019 (accessed 3 Feb 2020) https://epicenter.wcfia.harvard.edu/blog/bankrolling-ethics-do-tech-investors-have-responsibility-protect-democracy

We anticipate that open AI governance will require lead developers to get used to increased scrutiny on AI governance while performing their day jobs: development of new Smart Tech applications with many other contributors, each with their own objectives. Mediating possible conflicts among contributors in an open governance context will require new processes and new skills, especially in communication and mediation. There could be a role for independent certifiers to allow lead developers to focus on development while professional certifiers deploy their expertise into this new context.

We see two models of AI governance collaboration developing. One model is decentralized: most of the stakeholders involved in the design, development, usage and maintenance of an AI product decide voluntarily to share among themselves the governance steps taken by each contributor. In that model, stakeholders share the sum of the AI governance risks coming from each contributor.

In practice, data providers would work with AI developers to check how possible data bias or asymmetry might lead to mistaken or biased AI output. The decentralized model requires deeper sharing collaboration than currently practiced in commercial ventures and probably more complex legal contracts. Such closer collaboration could potentially help data providers and AI developers in quickly finding a solution to joint risks, such as adding a complementary data set to overcome initial data limitations.

This model may be more attractive when contributors have similar sizes and risk profiles. The strength of the overall AI governance attached to such a decentralized model is only as good as its weakest component. We would expect contributors to pay attention not only to their own AI governance, but also to everybody's share of the overall governance.

The second model is centralized. This is the case when contributors have very different sizes and capacity to take on risks, most likely the situation when a large lead developer works with a network of smaller firms or sub-contractors. In this case, the lead assumes the responsibility of the overall risk. Each of the small firms then focus only on their share of the AI governance. The lead developer has the incentive to perform due diligence with respect to the governance of the smaller firms or sub-contractors, and thereby requires transparency from each firm.

We can also see the possibility that governments or regulators become increasingly worried that repetitive failures of AI governance are affecting society's trust in Smart Tech applications. And that this lower trust is impacting adoption of new technological solutions to society problems - such as aging population or productivity. There is a risk that regulators may want to impose a Smart Tech application's stakeholders to share publicly their AI governance program and progress. It is difficult to see conditions under which this model would be welcomed by technology innovators across the world. However, if AI governance blunders continue, it may be an option that governments reluctantly impose in specific areas.

A truly open AI governance will most likely require users and regulators to be actively involved, most probably via an open platform. It does not feel that we are quite there yet, although there are a few initiatives trying to implement sharing platforms to store, exchange and sell trusted data and accompanying documentation. These are referred to as Data Trusts.[31]

## Proportional AI governance

A proportional, or fit-for-purpose, AI governance process recommends governance measures that vary in thoroughness as a function of the technical complexity and the magnitude of potential impact of malfunction or failure of a specific AI-powered Smart Tech application.

The AI that powers Smart Tech applications range from tried and tested to groundbreaking and complex. Depending on an application's context, the impact of a potential malfunction could range from catastrophic to benign. For example, misdiagnoses by medical AI are likely to have a much greater impact than algorithms generating poorly targeted personalized advertisements. We think that the depth and breadth of AI governance should depend on the complexity of a specific Smart Tech application's AI and the potential impact of a malfunction or failure.

Consider the next generation of recommendation engines for shopping websites. They may use facial recognition techniques to authenticate the user, then use that information to facilitate direct access to their online payment system. It may also eventually be possible to identify and predict customers' emotions by analyzing facial expressions coupled with other inputs such as the speed of typing or lingering of a cursor to gauge satisfaction with products displayed and therefore narrow down the set of products to recommend.

If all works well, the user's personal data (facial expressions, speed of typing, cursor lingering) may guide the shopping website to find what the user really wanted but could not describe well, providing clear tangible benefits.

[31] The new ecosystem of trust, Nesta 2019 (accessed 29 Jan 2020) https://www.nesta.org.uk/blog/new-ecosystem-trust/

However, using advanced facial expression recognition AI techniques could also raise ethical concerns, as the application combines the use of private data with complex new technologies not yet tested on the general public outside of China. New AI-powered algorithms which could potentially affect millions of online shoppers should probably be subjected to a rigorous governance approach, possibly involving independent observers to attest to the thoroughness of the AI governance applied. The AI governance in such a case should scrutinize potential risks before the technology is deployed and should provide a way to monitor user and public feedback once launched.

Proportionality of AI governance is in line with suggestions from Germany's data ethics commission, which recently recommended applying different controls to algorithmic decision-making systems based on their impact, using a 5-point potential impact rating scale.[32]

We think that the AI governance to adopt for a given Smart Tech application, and for each stakeholder, depends primarily on the following seven factors.

## 1. Type of AI technologies used

If a Smart Tech application has already been tested in the market and best practices are established, then the governance can be informal rather than formal. If the AI technology is groundbreaking and/or produces results that are impossible to explain, a comprehensive governance approach is needed.

This is likely to be a challenge for complex deep learning solutions as there are very few existing ways to explain results or to check how results have been generated.

## 2. Societal values

Governance is influenced by societal values and the behavioral norms in specific societies. Societal values determine where society stands with respect to ethical trade-offs such as "the individual vs. the state," "regulation vs. innovation," "privacy vs. customization or safety" and "transparency vs. vulnerability".[33]

In China, payment systems using facial recognition are on the rise and could become the norm in several years.[34] This evolution may indicate that people in China are willing to have images of their faces stored in central databases to get the benefits of faster, easier payments (no need for a phone or card).

In EU countries, governments have made protecting personal data privacy paramount. The EU and some member state regulators are considering limiting the use of facial recognition in some applications.[35]

The position a given society takes on ethical trade-offs will determine what kind of AI governance should be deployed for a specific Smart Tech application. And that of course will depend on the location of the developers, sellers and prospective users of the AI, which will increasingly make the problem more complex as Smart Tech applications are increasingly the result of multinational collaboration.

## 3. Privacy level of data used

If a Smart Tech application uses sensitive personal data, such as medical records or images of people's faces captured in public spaces for surveillance purposes, stricter controls around data handling will be required.

For example, it recently came to light that the London police had shared facial images of people captured by a security camera in the King's Cross Train Station for use in a facial recognition system without notifying the public.[36] The incident spurred the police to negotiate a data-sharing agreement with King's Cross authorities to govern the future use of such surveillance technology. The London Metropolitan Police said in late January 2020 that they would begin deploying Live Facial Recognition (LFR) technology in operational settings in selected locations around the city. In its press release, the Met emphasized the conditions in which the technology would be deployed and the effort to protect people's privacy.[37]

Beyond the UK, the EU has launched GDPR laws to protect EU citizen's privacy. California is in the process of finalizing laws providing even stricter privacy controls for California consumers.[38]

---

[32] Germany's data ethics commission releases 75 recommendations with EU-wide application in mind, Algorithm Watch 2019 (accessed 4 Feb 2020) https://algorithmwatch.org/en/germanys-data-ethics-commission-releases-75-recommendations-with-eu-wide-application-in-mind/

[33] Gaining National Competitive Advantage through Artificial Intelligence (AI), PwC 2019 (accessed 26 Jan 2020) https://www.pwc.lu/en/technology/gaining-national-competitive-advantage-through-ai.html

[34] Smile-to-pay: Chinese shoppers turn to facial payment technology, The Guardian 2019 (accessed 4 Feb 2020) https://www.theguardian.com/world/2019/sep/04/smile-to-pay-chinese-shoppers-turn-to-facial-payment-technology

[35] Google DeepMind NHS app test broke UK privacy law, BBC 2017 (accessed 4 Feb 2020) https://www.bbc.co.uk/news/technology-40483202

[36] Facial recognition row: police gave King's Cross owner images of seven people, The Guardian 2019 (accessed 4 Feb 2020) https://www.theguardian.com/technology/2019/oct/04/facial-recognition-row-police-gave-kings-cross-owner-images-seven-people

[37] Met begins operational use of Live Facial Recognition (LFR) technology, Metropolitan Police 2020 (accessed 31 Jan 2020) http://news.met.police.uk/news/met-begins-operational-use-of-live-facial-recognition-lfr-technology-392451

[38] AB-375 Privacy: personal information: businesses., California 2018 (accessed 29 Jan 2020) https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375

## 4. Number of people affected

Some Smart Tech applications are used by millions of people, such as Amazon's shopping recommendation algorithm. Others, such as an object recognition system on a drone used by a construction or mining company to predict the amount of daily work and equipment required are only used by, and likely to impact, a few people. AI algorithms used by the city of Los Angeles police for predictive policing are somewhere in between, with many thousand impacted.[39]

AI governance standards should vary as a function of the number of people that could be impacted by malfunction or failure. Unfortunately, the real world is not that simple. Many Smart Tech applications experience complex lives: they are created with a purpose in mind, sometimes with a small number of targeted users, and evolve over time, gathering in the process many more users. So, the initial number of people that could be affected by malfunction or failure can often grow exponentially during the lifetime of a Smart Tech application, and this should also impact the type of AI governance to be applied.

## 5. Severity of risks

AI is increasingly used in high-stakes applications that can have big financial and reputational implications and affect people's health and well-being.

For example, COMPAS[40] is an AI sentencing decision support tool used by US courts to predict defendants' likelihood of re-offending. Using COMPAS, a judge decides the sentence for re-offending. COMPAS has shown[41] racial bias in its prediction of re-offending and to be no better than using the average of the opinions of 20 randomly selected internet users.

Given the impact of a potential failure in such an algorithm, the governance applied to the development of a new generation of such a tool should be stricter than that applied to the launch of a song recommendation function on an audio streaming service website or app.

In client work, PwC reviews 6 categories of AI application risks: Performance, Security, Control, Economic, Societal, Ethical.[42]

Evaluating the performance risk of a specific AI application in several dimensions (for example cost, health, reputation) can be achieved with a variety of qualitative and quantitative techniques, such as Risk Analysis and Multi Attribute Utility Theory.[43]

## 6. Probability of risks

Several qualitative and quantitative techniques can be used to assess the probability of specific risks. For measurable risks, statistics is an obvious method. For non-observable risks, subjective probability assessments are a well-known method.[44] Methodological difficulties occur when comparing the risk of several AI technologies measured with different methods.

## 7. Range of AI governance tools

We reviewed earlier the current state of AI governance. There are already plenty of AI governance methods and tools available, and many more are created every year. Choosing the right set of available AI governance methods and tools will also depend on internal skills and external pressure, current or expected.

As public pressure for stronger controls continues to grow, we expect regulators will start to collaborate more and catch-up by issuing tighter recommendations and laws. These will become difficult to ignore. So far, there have been little real long-term financial or reputational impact when data or AI malfunctions are the result of accidental or deliberate behaviors. We expect this will change. As it does, it will force Smart Tech application developers to move rapidly first to user-centered AI governance then to society-centered AI governance.

Regulators will also have to evolve and adopt a more practical approach. If a regulation is difficult to enforce, regulators may want to make it more easily applicable or offer alternatives.[45] For instance, assurance of the quality of Smart Tech applications will likely remain necessary for potential high-risk cases. However, developing "meta" AI

[39] How the LAPD Uses Data to Predict Crime, Wired 2019 (accessed 24 Jan 2020) https://www.wired.com/story/los-angeles-police-department-predictive-policing/
[40] A Popular Algorithm Is No Better at Predicting Crimes Than Random People, The Atlantic 2018 (accessed 24 Jan 2020) https://www.theatlantic.com/technology/archive/2018/01/equivant-compas-algorithm/550646/
[41] The accuracy, fairness, and limits of predicting recidivism by Julia Dressel and Hany Farid, Science Advances 17 Jan 2018, Vol. 4, no. 1 (accessed 28 Jan 2020) https://advances.sciencemag.org/content/4/1/eaao5580
[42] Responsible AI Toolkit, PwC 2019 (accessed 23 Jan 2020) https://www.pwc.com/rai
[43] Decisions with Multiple Objectives by R. L. Keeney and H. Raiffa, Wiley, 1976.
[44] Foundations of Decision Analysis, Global Edition, Pearson, by Ronald A. Howard, Ali E. Abbas
[45] It's Hard to Ban Facial Recognition Tech in the iPhone Era, Wired (accessed 29 Jan 2020) https://www.wired.com/story/hard-ban-facial-recognition-tech-iphone/

engines to identify AI design or data weaknesses is not yet feasible. Instead, making sure that responsible AI governance measures are applied at every step of the development process for Smart Tech application — or, for example, testing Smart Tech applications for boundary conditions with different sets of data — should be considered.

To sum up, the conceptual formula below lists the seven factors described so far. The formula is a simple way to help stakeholders in the design, development, usage and maintenance of a Smart Tech application define the AI governance required to minimize risks of malfunction and failure and contribute to build trust, not only in their specific Smart Tech application, but also in all future Smart Tech applications, one layer at a time.


The formula is also helpful to check how AI governance should change over time, as technology, societal values and the range of available governance methods and tools will themselves change over time.

## End-to-end AI governance

Most of the efforts in AI governance are targeted at the pre-launch of a Smart Tech application. To provide complete AI governance, the process needs to start with the design stage of the application and end with its retirement. This may not be necessary when a Smart Tech application's purpose, focus and user base does not change much over time.

However, an increasing number of Smart Tech applications start their lives being designed for a specific function and are tested on a small number of users. Risk of malfunction is contained and expected impact of a potential malfunction is benign. With the advent of agile development, reality is quite often different, as successive versions are launched with increased functionality to an ever-larger number of users. For example, a recent mobile phone application was launched to show a person's image as their "future face" or how they will look when they are older. It was meant to be fun with a relatively small user base, but quickly became very popular, calling into question the purpose of collecting these images.

The speed of the user base growth of some Smart Tech applications can be astonishing, with significant increases in users sometimes happening in a matter of a few days or weeks. That requires quick reaction from the AI governance framework and processes, which has typically worked in monthly or annual cycles. Adapting AI governance to the pace of evolution of Smart Tech applications, especially in the consumer space, is a real challenge. End-to-end AI governance poses logistical challenges and requires developing automated tools.

Despite these challenges, end-to-end AI governance is likely to be the only effective method to ensure that appropriate controls can be exercised when the number of people, that could be affected by the malfunction or failure of a Smart Tech application, starts increasing.

To conclude this section, we are convinced that AI governance can only succeed durably if it is holistic, open, proportional and end-to-end. We are aware that this will require more formal processes to be put in place and more collaboration among stakeholders.

# Mr. Yoshida's story

## The old man and the exoskeleton

With our comprehensive approach to AI governance, the story of Mr. Yoshida, the bedridden orchard worker, might have ended differently.

The following scenario describes how Mr. Suzuki, the developer working on an AI-augmented exoskeleton, could apply our governance framework to develop a new product to help Mr. Yoshida walk again.

The main contributors to the creation of the autonomous skeleton are 1) Mr. Suzuki, an engineer who uses AI and robotics to create this Smart Tech solution in the role of a Smart Tech application maker and seller; 2) Mr. Yoshida, the user of the exoskeleton, in the role of the AI user; and 3) Mrs. Yamada, the mayor of the town where the exoskeleton is used, in the role of local government. Other contributors we have described earlier in the paper such as investors, the public and legislators and regulators are not covered in this scenario.

We follow how the main characters progress through the successful development of the exoskeleton while determining the governance necessary to earn the trust of the local community and achieve success for the user. As this is an illustration, the situation and names are fictitious.

Hiroyuki Yoshida is a 92 year-old retired orchard farmer who lives near Iino in Yamanashi prefecture in Japan. He and his wife have two children who live in Osaka and Tokyo. The children are both in their 60s and married, with children and grandchildren of their own. Mr. Yoshida's wife, Michiko, is slightly younger than he is and in good health.

Before retiring, Mr. Yoshida spent his entire working life in the orchards within 10 km of his current home. Years of physically challenging work in the orchards and advanced age eventually took a toll on his back to the point that he became bedridden a few years ago. His wife provides the care and companionship he requires, and his friends, mostly retired farmers, visit him frequently. Even so, Mr. Yoshida has become increasingly withdrawn, far from the exuberant character he had always been before he lost mobility.

Recently, his children saw a TV program about a small firm led by a visionary engineer, Masataka Suzuki, who licensed technologies to build the next generation of exoskeletons, working with major robotics companies. Mr. Suzuki was convinced that exoskeletons had thus far been designed far too conservatively, focused solely on helping workers lift heavy loads, patients in hospitals, for example, or heavy, unevenly shaped objects in warehouses and factories. He felt that with a little investment, experiment and luck, he could work with his local university's AI lab to develop an AI-powered exoskeleton designed to help people like Mr. Yoshida walk again. Mr. Suzuki envisioned using AI in the exoskeleton to bring stability and smoothness of movement. Equipped with sensors, the exoskeleton could anticipate changes in terrain and respond accurately to slight, intentional movements of the wearer, while taking into account the wearer's specific way of balancing.

Mr. Yoshida's children contacted Mr. Suzuki and after a few meetings with the family to explain that his current prototype needed to be tested progressively, they all agreed to try it. Since that time, Mr. Suzuki has worked closely with Mr. Yoshida and his doctor to adapt the exoskeleton to Mr. Yoshida's body and train the AI to give it some autonomy. Mr. Yoshida's goals were ambitious. He first needed to be able to put the exoskeleton on and take it off by himself. Next, he wanted to be able to meet his friends for a drink at the local Orchard Farmer Union building, a 15-minute walk from his house along a quiet, narrow road. Ideally, he also wanted to be able to visit his doctor or nurse at their offices in the next town, which meant walking to a bus stop, getting on and off the bus, crossing the road, climbing stairs to a second-floor office and returning home.

Mr. Yoshida felt that progress was slow, with great improvements one week and disappointing results the next. But he enjoyed the regular exercise routine he had picked up to regain some strength. Mr. Suzuki however was delighted. After a year of development and refinement work, time had finally come to start the real test: using the exoskeleton every day.

The progress made during testing and refining his custom exoskeleton has Mr. Yoshida beaming. He has not yet shared much of the progress with his grandchildren and great-grandchildren, and he is looking forward to surprising them with a live demonstration. Mr. Yoshida is now familiar with the odd-looking exoskeleton.

He knows how to shift his weight to indicate the direction in which he wants to move, and he is even able to start a slow, gentle jog for a few meters on the flat road in front of his house.

Lately, Emi Yamada, the town's mayor, has been interested to learn more. This Smart Tech application could help many of her older constituents stay active into their later years, improving their physical health and emotional well-being and giving them the ability to contribute more to their families and the community. She is a bit apprehensive however, and feels some responsibility to keep her townspeople safe. What if Mr. Yoshida's exoskeleton stops when he crosses the road? What if he falls down on the bus, distracting the driver and causing an accident?

She discusses with Mr. Suzuki how to reduce these risks and keep Mr. Yoshida, fellow citizens in their vicinity and future exoskeleton wearers safe. After the first nine months of use, locals have gotten used to seeing Mr. Yoshida in his exoskeleton increasingly farther away from his house, accompanied by Mr. Suzuki and his team. Mr. Yoshida has actually become something of a celebrity and other community members with similar mobility limitations have volunteered to test a new, lighter version of the apparatus.

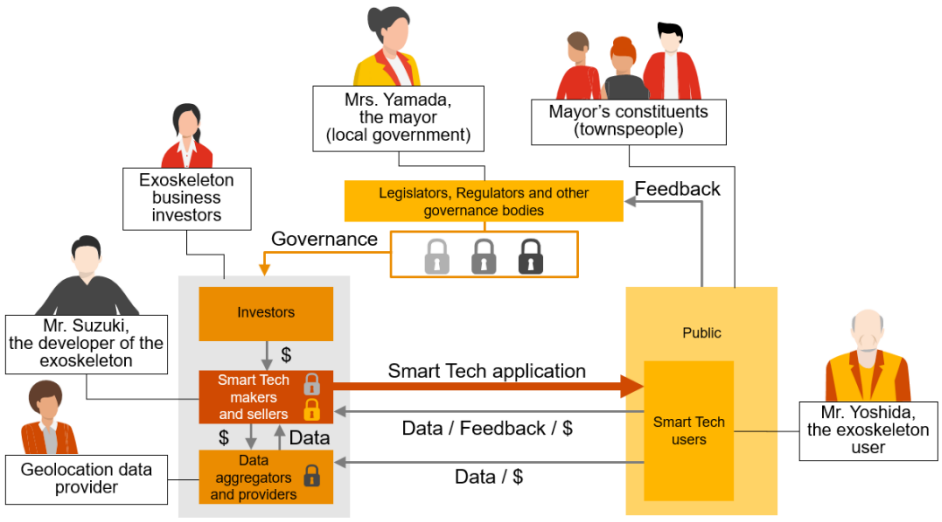
An accident, however, motivated Mr. Suzuki to integrate another form of AI into the device and spurred Mrs. Yamada to institute measures to protect exoskeleton wearers and the general public. One day, Mr. Yoshida stumbled outside of his doctor's office. He bumped into a passerby who fell and cut his forehead. The pedestrian recovered and Mr. Yoshida was unhurt. Mr. Yoshida's doctor ran some tests and discovered his eyesight was slowly deteriorating, which could explain why he had not seen the curb well and failed to adjust his weight to "tell" the exoskeleton to step down.

The accident gave Mrs. Yamada the idea to develop an "exoskeleton wearing license" process to minimize risk to the local community and exoskeleton wearers. This was also the reason why Mr. Suzuki decided to collect health data while Mr. Yoshida used the exoskeleton. This would help his doctor to monitor how Mr. Yoshida was doing and would help identify when he needed treatment or interventions, such as new glasses or additional physical therapy. It also made Mr. Suzuki think that he should go back and label the data collected to recognize the conditions it was collected under - for example tagging the wearer's health parameters to the usual GPS location and movements of the exoskeleton. This would help filter the right data to train the exoskeleton's AI for the next potential wearer. Mr. Suzuki also realized that he should create a more comprehensive, thorough AI governance process in advance of the next application, rather than the one-step-at-a time approach he found himself using with Mr. Yoshida.

Thanks to the exoskeleton, Mr. Yoshida has enjoyed four years of mobility he would not otherwise have achieved — along with a great sense of pride to be the first to walk with such a high-tech "helper." He has made two new friends in Mr. Suzuki and Mrs. Yamada. He has reconnected with his old farmer chums. And he is no longer such a burden on his wife. These are all wonderful developments, but the best part, he says, is the look of awe in his great-grandchildren's eyes as they take long walks at his side admiring their "hybrid" great-grandpa.

## The role of AI governance in the story

The figure below shows how each person maps to a role in the process of designing, developing, using and maintaining a Smart Tech application.




Let's look now at the AI governance steps taken by each of the main contributors.

Mr. Suzuki, the Smart Tech developer of the exoskeleton, set out to create a highly reliable Smart Tech solution by integrating the right level of AI governance into his creation's design, testing and continuous improvement by doing the following:

1. He took time to explain the exoskeleton technology and realistic risks and benefits to Mr. Yoshida and his family: how his exoskeleton worked; how it would be adapted to Mr. Yoshida's needs; the fact that getting a functioning solution would likely take months, requiring patience and determination from Mr. Yoshida; the risks that Mr. Yoshida would take and ultimately the risk that it may not work; the routine he recommended, testing new features progressively and checking with Mr. Yoshida and his family at every step of the development.
2. He explained what data they will be collecting; how it will be used to personalize the experience for Mr. Yoshida; how that data will be used for overall design improvements of Mr. Yoshida's exoskeleton and for the next wearers of exoskeleton. Mr. Suzuki also explained to Mr. Yoshida that there was an option not to share his data with Mr. Suzuki's company and the implications.
3. Mr. Suzuki made sure his AI algorithm used highly accurate geolocation data. He first checked the general accuracy of the geolocation data with the provider, then conducted geolocation accuracy tests of the data in the exact areas where Mr. Yoshida was expected to use the exoskeleton.

4. He also made sure that the exoskeleton's training data used to ultimately command movements given live pictures of the ground in front and around the exoskeleton was sufficiently diverse, being able to recognize the roads and pedestrian areas Mr. Yoshida would traverse. Additional data cleansing measures were conducted to de-noise the geolocation data.
5. Mr. Suzuki made sure that real-time data collection as well as validation of select data samples allowed him to monitor the performance of the technology, including adjusting movements in real time and alerting the user of any potential problems. Doing this also provided more training data that could be used to train similar exoskeletons used by other customers and to improve the accuracy of computer vision environment detection by using data from a greater variety of environmental situations.
6. He ensured the exoskeleton's basic design could safely carry Mr. Yoshida's weight in all likely conditions and areas, and that Mr. Yoshida could "communicate" with it using only slight movements according to his physical stature and ability.
7. He ensured that the exoskeleton's "shoes" were well adapted to the walking surfaces where Mr. Yoshida would be expected to use the exoskeleton: around his home, on the road and pavement in the town, up and down stairs and on the bus.
8. Overall, Mr. Suzuki conducted progressive usage tests, introducing progressively more complex tasks until he, Mr. Yoshida and Mrs. Yamada felt it was safe to test in real conditions.
9. Mr. Suzuki also provided regular training visits and explicit warnings about conditions in which the device had not yet been tested or certified — for example: "use on paved roads only."
10. He also took time to explain risks and benefits to Mrs. Yamada and her constituents.

While all these steps increased development time, they also helped minimize the short-term risks associated with introducing an exoskeleton in the community. This has helped build trust in the exoskeleton's ability to help local seniors be more mobile — and thus happier in their twilight years — without creating a public safety hazard.

For his part, Mr. Yoshida also took his responsibility seriously:

1. He took the time to understand the exoskeleton's benefits and risks, its technology, how to use it safely and progressively, and the behavior required of him, including taking regular exercise to gain strength and have greater ability to handle the exoskeleton.
2. He agreed to undertake a few tests in difficult conditions to experience firsthand the risks of using the exoskeleton outside of its design conditions — for example, in the rain.
3. He provided detailed feedback to Mr. Suzuki so the exoskeleton's AI could be improved for his own benefit, as well as for the benefit of other future users of AI-enabled exoskeletons.

Finally, on behalf of the local public, the mayor, Mrs. Yamada, quickly understood the trade-offs at stake: risk to the community if there were accidents and benefits if the exoskeleton was successful. In absence of laws, she used national guidance and her common sense:

1. She took the time to understand the technologies involved, how they could be safely applied in public spaces (roads, sidewalks, buildings, transit), and their risks and benefits for individuals and for her community and society.
2. She consulted with others, Mr. Yoshida's doctor for example, to adapt to an unforeseen change in Mr. Yoshida's eyesight before creating a local pilot test for an "exoskeleton wearing license," with support from interested national authorities. To obtain the license, she required that users pass physical and cognitive tests every year, or more frequently when needed, as well as obtaining recommendations from their doctor and a physical therapist.
3. She communicated with her constituents frequently about the exciting experiment that was taking place in their community with Mr. Yoshida and Mr. Suzuki, making locals aware of benefits and risks, such as when walking close to Mr. Yoshida and his exoskeleton, or when passing them on the road in a vehicle.

Not explicitly described in our illustrative case are industrial robotics and AI organizations that are also learning from this experiment and exploring the development of a certification process to ensure that the AI exoskeleton has safety backups to allow the human wearer to override the AI if necessary. Also, not mentioned are investors in the exoskeleton business, who are very interested in quickly scaling the business up, but are cautious about the risks and are consulting closely with Mr. Suzuki in developing a successful exoskeleton.

Let's look at the factors in our conceptual formula and how they defined the governance used by each stakeholder and why.

Type of AI Technology: Mr. Suzuki used several AI technologies in making the exoskeleton. Here are considerations for two:

- Geolocation – Finding a device's location based on GPS positioning data is a very mature technology. Nevertheless, Mr. Suzuki wanted to ensure quality of reception in all areas where Mr. Yoshida was expected to be with his exoskeleton. This is why he conducted some specific tests.
- Interpretation of pressure on knee pads to decide direction of movement – This is a new technology with little testing done in public space. This is why Mr. Suzuki decided to test slowly.

Societal values: Japan's openness toward robotic technologies led to early adoption of the exoskeleton by Mr. Yoshida. His willingness to share performance and personal data of his tests contributed to increasing Mr. Suzuki's and Mrs. Yamada's knowledge of exoskeleton performance, a great gift to accelerate the launch of other exoskeletons in their community.

Privacy level of data used and number of people affected: The data generated through exoskeleton usage was anonymized and used to train the AI further to improve accuracy of movements for Mr. Yoshida and for future exoskeleton wearers.

Severity and probability of risks: The exoskeleton has been rigorously tested and improved before being commercialized for Mr. Yoshida. In the next stage, Mr. Suzuki will refine the application of the exoskeleton to older people with a larger test - still in the home town of Mr. Yoshida. Robotics companies have offered help to scale up. The Japanese Health Ministry has also reached out to offer hospitals who are candidates to test Mr. Suzuki's exoskeleton. Mr. Suzuki is checking how many of these initiatives he can safely lead in parallel.

Range of AI governance tools: Mrs. Yamada noticed a governance gap and worked with the national government to introduce a local "exoskeleton wearing license." Mr. Suzuki has realized that before his next development step he needs to design a suitable AI governance process.

This example illustrates how collaboration leads to sharing governance responsibilities, mitigating risks of accidents or failure, and establishing the success of the exoskeleton for Mr. Yoshida. Thanks to this close collaboration, further development of the exoskeleton has been made possible, leading to the potential wider use of this innovative solution to diminished mobility.

This story illustrates a practical application of the holistic, open, proportional and end-to-end AI governance process we advocate.

# Conclusion: Let's work together

PwC and Eurasia Group believe the large-scale adoption of Smart Tech applications can only happen if the public trusts that such applications have been tested robustly using a consistent, comprehensive and well-articulated method. So far, most of the governance has been put in place after incidents have identified problems to fix. A better approach is to be proactive, which is why PwC is proposing this more comprehensive AI governance approach.

We hope this paper has provided a valuable, pragmatic contribution to the debate on how AI governance can contribute to making Smart Tech applications successful. We hope our illustrative example shows how a holistic, open, proportional and end-to-end AI governance works best when implemented collaboratively.

We hope organizations that design, develop, use and maintain Smart Tech applications can use our Responsible AI framework and Toolkit to implement the right governance steps to address specific risks of their Smart Tech applications.

Finally, we hope legislators, regulators and other governance bodies will encourage the creation of AI governance communication platforms to help increase and maintain trust among users and the public.

If governments and businesses work together to promote the use of a more open AI governance approach, combined efforts will help reduce the risks of Smart Tech failures and strengthen public trust in the development of Smart Tech applications. This is the best way to build the long-term trust that is essential for the development of Smart Tech applications. This is the best way to start building a more human-centered AI governance.

## Authors

Yann Bonduelle and Batsuren Enkhtur (PwC Japan); Maria Axente (PwC UK); Ilana Golbin and Anand Rao (PwC US); Kevin Allison and Paul Triolo (Eurasia Group).

## PwC's expertise in AI

In recent years, PwC has highlighted the opportunities of AI, including how augmentation and automation promise to transform society and work. We have worked with regulators and governments to help define the right governance and regulation. For instance, PwC AI experts provide strategic advice to public and government bodies, sectoral regulators and expert groups to address risks and governance requirements for AI techniques in line with organizational objectives. PwC has researched and pushed the concept of Responsible AI, developed a Responsible AI toolkit, and is tracking national AI strategies.[46]

By bringing Smart Tech and AI to PwC's own operations, we are on the front line of AI governance with respect to data, algorithms and monitoring Smart Tech applications after launch. Working with organizations to help bring Smart Tech or AI into their operations provides a useful feedback loop, so we can integrate what works best to develop Responsible AI. We advise clients on designing the right strategy and roadmaps for them. We help them with data transformation, supporting execution of Smart Tech and AI use cases, defining operating models for their analytics and AI team(s), operationalizing AI and measuring and accelerating impact. Each situation is different, with different trade-offs among urgency of action, scope, and risk of failure, thus requiring unique approaches. We, of course, provide advice on developing a holistic, open, proportional, and end-to-end AI governance process.

## Eurasia Group's expertise in Geo-technology

Established in 2016, Eurasia Group's Geo-technology Practice advises global corporations and investors on risks emerging at the intersection of technology and global politics, including risks involving AI, cross-border data flows, 5G networks, cybersecurity, autonomous vehicles and semiconductors. The geo-technology practice works with Eurasia Group's other country analysts and sector experts to provide crosscutting analysis that traverses countries, regions and industries affected rising geopolitical competition and changes in government technology policies and regulation.

Founded in 1998 by Dr. Ian Bremmer, Eurasia Group is the first firm devoted exclusively to helping investors and business decision-makers understand the impact of politics on risks and opportunities around the globe. Ian's idea—to bring political science to the investment community and to corporate decision-makers—launched an industry and positioned Eurasia Group as the world leader in political risk analysis and consulting. Eurasia Group's dynamic partnerships with leading firms in the investment, consulting and broader professional services industry complement our politics-first capabilities and expand our suite of client solutions.

## Acknowledgement

This paper builds on the work already published by PwC on Responsible AI (RAI). RAI focuses on creating a framework to minimize and mitigate risks. RAI Toolkit is a suite of customizable frameworks, tools and processes designed to help organizations in harnessing the power of AI in an ethical and responsible manner - from strategy through execution. With our **h**olistic, **o**pen, **p**roportional, **e**nd-to-end AI governance, we broaden the focus of RAI to look at the entire system of AI governance to help the overall AI ecosystem produce successful Smart Tech applications.

[46] National AI Strategy (NAIS) Radar, featured in "Gaining National Advantage through Artificial Intelligence (AI)", PwC 2019 (accessed 5 Feb 2020) https://www.pwc.lu/en/technology/gaining-national-competitive-advantage-through-ai.html

© 2020 PwC. All rights reserved.
PwC refers to the PwC network member firms and/or their specified subsidiaries in Japan, and may sometimes refer to the PwC network. Each of such firms and subsidiaries is a separate legal entity. Please see www.pwc.com/structure for further details.

© 2020 Eurasia Group. All rights reserved. This material was produced by Eurasia Group for use by the recipient. This is intended as general background research and is not intended to constitute advice on any particular commercial investment or trade matter or issue and should not be relied upon for such purposes. It is not to be made available to any person other than the recipient. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic or otherwise, without the prior consent of Eurasia Group.

This content is for general information purposes only, and should not be used as a substitute for consultation with professional advisors.