



UNIVERSITY OF  
TORONTO



**SCHWARTZ  
REISMAN  
INSTITUTE FOR  
TECHNOLOGY  
AND SOCIETY**

---

**STRATEGIC PLAN  
2021–2024**

June 15, 2021

# Table of Contents

Executive Summary .....	3
Background .....	6
Governance & History of the Council of Deans .....	7
Vision.....	8
Mission .....	10
Strategic Objectives.....	11
1 / Interdisciplinarity.....	11
2 / Field-Building.....	13
3 / Regulatory Innovation.....	14
4 / AI for Social Good.....	16
Conclusion .....	18
Appendix A: Landscape of Comparable AI and Society Institutes.....	19
Appendix B: Council of Deans & Terms of Reference .....	20
Terms of Reference.....	20
Members.....	20
Appendix C: Gift Agreement Metrics.....	21

## Executive Summary

**Vision:** *The vision for the Schwartz Reisman Institute for Technology and Society is to be the world's leading institute for innovative research and practical solutions that help ensure that artificial intelligence (AI) and other advanced technologies benefit **all** of humanity.* This vision is the “raison d'être” for the Institute, motivating potential investors, faculty, staff, and partners at the University of Toronto, as well as around the globe, to contribute to the Institute's initiatives.

**Mission:** *The Institute's mission is to deepen our knowledge of technologies, societies, and what it means to be human by actively integrating research across traditional boundaries and building human-centered solutions that really make a difference.*

We will achieve our mission by:

1. *Conducting and integrating world-renowned **research** and academic training – and being truly interdisciplinary.* The Schwartz Reisman Institute for Technology and Society will foster interdisciplinary ideas, insights, and understanding of how technology affects society and individuals.
2. *Serving as a convenor and a hub for paradigm-shifting public **policy** conversations.* We will convene national and international public policy discussions, engaging directly with government officials, civil society organizations, and other stakeholders to accelerate innovation and inform public policy within the context of an AI-enabled future, ensuring that the policy impacts of AI serve the public good while protecting citizens and societies from AI's misuse.
3. *Leveraging our expertise to develop human-centred **solutions** to meet the tremendous demand for innovative, actionable, high value, and globally appealing solutions for building safe, responsible, and inclusive AI and other advanced technologies.*

In our first five years, we are specifically focused on AI and powerful data-driven technologies. In *Appendix A: Landscape of Comparable AI and Society Institutes*, we list major institutes around the world whose work complements ours. We have also formed, or are in the process of forming, partnerships with several of these organizations. We anticipate that our work will expand to encompass powerful biotechnologies in our second five years. Throughout this document, we will refer to AI but intend this designation to also include other powerful data-driven technologies.

**Governance:** The Schwartz Reisman Institute for Technology and Society is governed by a newly constituted Council of Deans appointed by the president of the University of Toronto, and includes representatives from the Division of the Vice President, Research & Innovation (VPRI)

and representative deans and principals. The Council will meet twice a year and is responsible for approving the Institute's strategy and budget; the Council will also advise on key strategic matters such as major partnerships and fundraising. See *Appendix B: Council of Deans & Terms of Reference*.

Additionally, the Institute is in the process of appointing an Advisory Committee, consisting of key industry leaders and renowned subject matter experts in AI, governance, ethics, and other relevant fields who will help guide the future of the Institute. The Advisory Committee will meet at least twice a year beginning in the fall of 2021, and on an ad hoc basis as required.

**Strategic Focus Areas:** Achieving the promise of AI – and avoiding its negative consequences – requires answering two challenging questions, which constitute our *strategic focus*:

1. How will we build the technological, legal, and regulatory infrastructure needed to ensure that AI advancements are responsible and aligned with human goals, ethics, and values?
2. Conversely, how will we make sure that our existing legal and regulatory systems do not either unintentionally or unnecessarily block AI innovation?

**Strategic Objectives:** We have identified four strategic goals to achieve our vision and mission:

1. **Interdisciplinarity: Increasing the range and depth of interdisciplinary research in the AI and society ecosystem at U of T and beyond.**
  - Cross-disciplinary seminars, discussion groups, conferences, workshops.
  - Connecting scholars across disciplines to foster research programs and connecting cross-disciplinary research teams to grants and philanthropy.
2. **Field-building: Fostering and advancing new fields of research in the sciences, social sciences, and humanities to deepen our understanding of powerful technologies and human relationships to technology.**
  - Identifying novel, promising new lines of research.
  - Promoting research collaborations, funding, publications, seminars, conferences, and workshops that profile and advance these lines of research.
  - Growing the community of researchers globally engaged in these lines of research.
3. **Regulatory Innovation: Transforming the world's approach to regulating powerful technologies.**
  - Cross-disciplinary research contributing new forms of regulation with a focus on democratic, agile, and efficient governance that works at global scale, addresses growing public/private power gaps, and aligns technology with human flourishing, agency, and values.

- Integration of technical solutions with legal and regulatory frameworks.
  - Fostering entrepreneurship and investment in innovative regulatory technologies.
  - Working with governments, industry, and civil society to build global legal frameworks for accountable regulatory markets.
- 4. AI for Social Good: Working with industry, governments, and civil society organizations to maximize the benefits that new powerful technologies can bring to people across income, class, gender, ethnicity, and global divides.**
- Identifying dynamics and barriers that block development and deployment of AI and other technologies that benefit groups underserved by market incentives.
  - Cultivating research initiatives that generate technology and governance solutions that benefit groups underserved by market incentives.
  - Building AI and governance capacity in low- and middle-income countries as well as historically marginalized groups such as Indigenous and migrant populations.

**Key Performance Indicators (KPIs):** Finally, we have outlined KPIs for each of our strategic objectives:

- 1. Interdisciplinarity KPI:** # of research products by U of T faculty/postdocs/graduate students in Schwartz Reisman-related topics that are interdisciplinary (as determined by research methodology or co-authorship/team composition); e.g., publications, grants awarded.
- 2. Field-building KPI:** # of novel research domains contributed by U of T faculty/postdocs/graduate students in Schwartz Reisman-related topics (as determined by peer survey).
- 3. Regulatory Innovation KPI:** # of Schwartz Reisman-inspired governance mechanisms/designs implemented or piloted by governments, industry, or civil society organizations.
- 4. AI for Social Good KPI:** # of Schwartz Reisman-inspired governance mechanisms/designs implemented or piloted by governments, industry, or civil society organizations that specifically support AI for social good. These mechanisms/designs could be described in various communications/media in top-10 outlets (national news media, global news media, technology media).

Our KPIs also include specific measures in our gift agreement with our founding donors, outlined in *Appendix C: Gift Agreement Metrics*.

## Background

Created in 2019 as a result of a historic gift from Canadian entrepreneurs Gerry Schwartz and Heather Reisman, the Schwartz Reisman Institute for Technology and Society explores and addresses the ethical and societal implications of technology. In its first five years, the Institute is focused on AI. AI is rapidly transforming our economic and social environments, and, in an increasing number of instances, threatening or causing harm to human flourishing, social stability, and the well-being of vulnerable populations. At the same time, the potential public benefits of AI, especially in the public sector, are frequently blocked by legal and regulatory obstacles that are poorly adapted to the modern digital environment. The Schwartz Reisman Institute for Technology and Society was created for the purpose of developing the research and new approaches needed to maximize the benefits and minimize the harms flowing from powerful technologies. Metrics for the Institute established at the Institute's founding are listed in *Appendix C: Gift Agreement Metrics*.

The Institute will eventually be housed in the **Schwartz Reisman Innovation Campus**, a 750,000-square-foot complex designed to anchor U of T's unique cluster of world-leading AI scientists and biomedical experts, its world-class entrepreneurship network, and the country's largest concentration of student- and faculty-led start-ups, creating the country's biggest university-based innovation node.

The new Campus will turbocharge the next wave of Canadian innovation, advancing how AI, biomedicine, and other disruptive technologies can enrich lives.



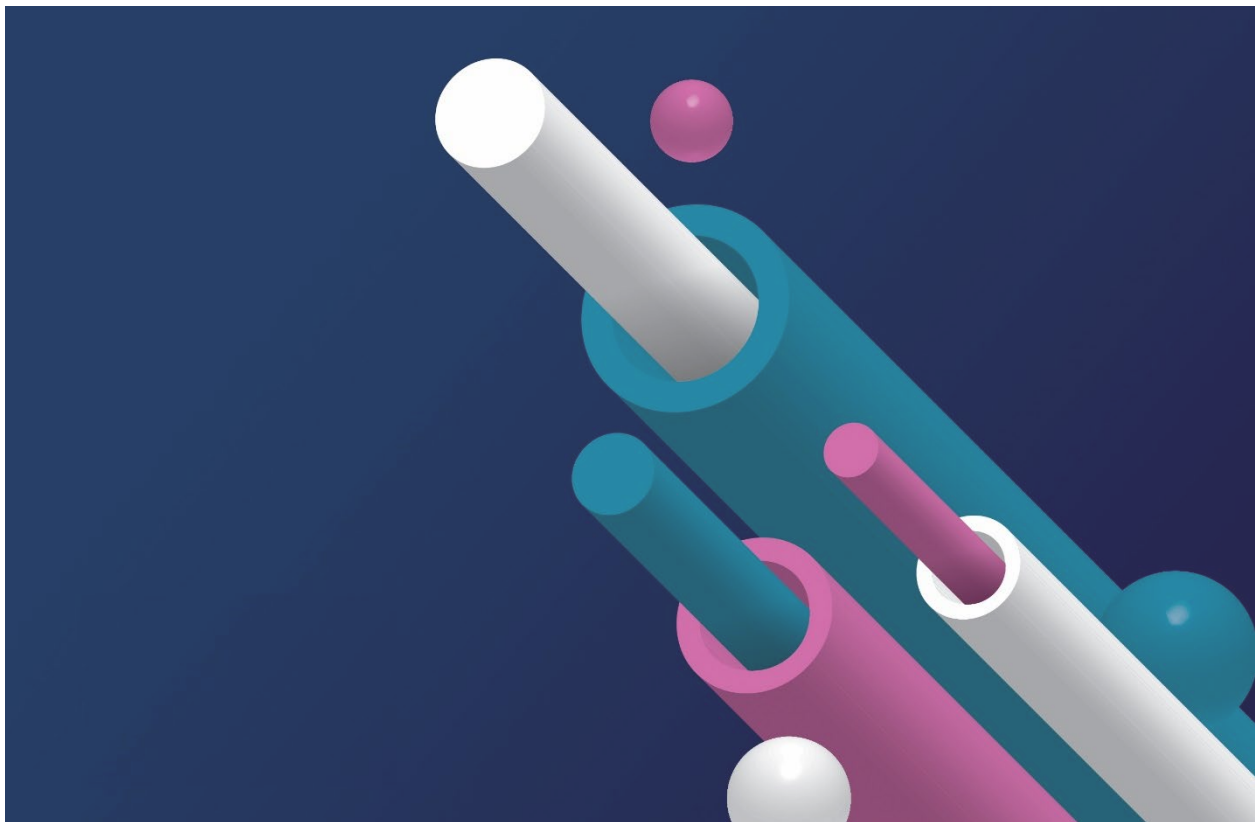


## Governance & History of the Council of Deans

The Schwartz Reisman Institute for Technology and Society is governed by a Council of Deans appointed by the president of the University and includes representatives from the Division of the Vice-President & Provost, the Division of the Vice President, Research & Innovation, and representative deans. The Council of Deans meets twice a year and is responsible for approving the Institute's strategy and budget. See *Appendix B: Council of Deans & Terms of Reference*.

The Council (then constituted as the Steering Committee) first met in 2019 to approve the appointment of the inaugural director and chair of the Institute, [Gillian K. Hadfield](#).

The Institute is in the process of appointing an Advisory Committee, consisting of key industry leaders and renowned subject matter experts in AI, governance, ethics, and other relevant fields who will help guide the future of the Institute. The Advisory Committee will meet at least twice a year beginning in the fall of 2021, and on an ad hoc basis as required.



## Vision

On a scale comparable to the agricultural and industrial revolutions, artificial intelligence (AI) is promising to transform every sector of human life: medicine and public health, financial systems, manufacturing, communications, government, policing, the legal system, and more. But achieving the promise of AI – and avoiding its negative consequences – requires answering two challenging questions:

1. First, how will we make sure that AI advancements do not cause harm, that they are responsible and aligned with human goals, ethics, and values? Why are fairness-informed algorithms, for example, not a manufacturing requirement just as seatbelts are in automobiles? Why don't machine learning engineers need to be licensed to build AI infrastructure in the way civil engineers or lawyers do?
2. Second, and conversely, how will we make sure that our existing legal and regulatory systems do not either unintentionally or unnecessarily block beneficial AI innovation? In healthcare, for example, AI systems can diagnose some cancers better than human physicians.<sup>1</sup> But data governance rules developed before the advent of massive amounts of digitized medical innovation are hampering the aggregation of data to support the use of powerful digital technologies. Why don't we have trusted data infrastructures in place to simultaneously protect privacy and enable the use of powerful technologies to improve health and save lives?

Although AI is still in its early phases, we are seeing ample evidence that the challenge of ensuring it does not cause harm is already with us. Powerful facial recognition technology, for example, is already threatening to destroy ordinary concepts of privacy and upend trust, as a single photo or video capture can be used to track just about anyone anywhere on the internet. At the same time, outdated approaches to data governance are hampering efforts to use powerful machine learning systems to build smart cities or innovate improved techniques for the prevention, diagnosis, and treatment of disease.

These challenges are not limited to data governance. They are presenting themselves in every domain of the advanced democratic market economy: determining how to regulate autonomous vehicles, implementing fair automated decision-making in government services and industry, responding to the increasing concentration of wealth and power in private technology companies, ensuring that automation does not cause massive disruption to employment, and so on. The root

---

<sup>1</sup> [Transforming healthcare with AI: The impact on the workforce and organizations](#), McKinsey & Company. Executive Briefing, March 10, 2020.



of these challenges is the same: our existing legal, regulatory, and policy infrastructure – built for the manufacturing nation-state economy of the 20th century – is poorly suited to keep up with the speed, complexity, and global reach of AI and other digital technologies. There are widespread calls for regulation to ensure AI benefits us all, but there are few ideas about how to implement that regulation.

To date, the global community has produced little beyond abstract ethical principles and guidelines, and a call for industry self-governance. Governments feel ill-equipped to assess these complex technologies, and it is increasingly clear that solutions will not lie in crafting ever-more complex pieces of legislation that are ill-informed and outdated practically from the get-go. Securing the benefits of AI while ensuring that the powerful transformations on our doorstep do not destroy the stability and inclusiveness of our societies will require reinventing our approaches to democratic governance and market regulation, and a transformation in the way we build AI itself.

How can we develop the technologies and techniques that ensure AI systems are correctly reading the human environment and generating recommendations or actions that are aligned with what diverse human communities aspire to achieve? How can we ensure that the technologies needed to monitor AI development and ensure its responsible deployment are also built? How can we attract the necessary investment and innovation in regulatory technologies? How can we ensure that the computer scientists and tech companies building AI technologies – who are themselves often highly unrepresentative of the population in terms of gender, race, and income – do so in a way informed by human values, ethics, and democratic ideals? How can we ensure that AI does not widen the wealth and knowledge gaps between us? That the benefits of AI are secured for the benefit of all?

These are the challenges of inventing the technical, legal, regulatory, and democratic infrastructure we need for the age of AI.

*The vision for the Schwartz Reisman Institute for Technology and Society is to be the world's leading institute for innovative research and practical solutions that help ensure that AI and other advanced technologies benefit all of humanity.*

It is imperative that we accelerate research, policy, and solutions to ensure that powerful AI systems are built in ways that promote human well-being, with fundamentally new approaches to governance and regulation needed as AI systems reshape what we know, how we interact, and the distribution of power. AI is driving forward at a furious pace in contexts that should deeply concern us, and paradoxically, idling in contexts that could lift us up. Now is our time to act.

## Mission

The Institute's mission is to deepen our knowledge of technologies, societies, and what it means to be human by actively integrating research across traditional boundaries and building human-centered solutions that really make a difference.

We will achieve our mission by:

- ▶ *Conducting and integrating world-renowned **research** and academic training – and being truly interdisciplinary.* The Schwartz Reisman Institute for Technology and Society will foster interdisciplinary ideas, insights, and understanding of how technology affects society and individuals.
- ▶ *Serving as a convenor and a hub for paradigm-shifting public **policy** conversations.* We will convene national and international public policy discussions, engaging directly with government officials, civil society organizations, and other stakeholders to accelerate innovation and inform public policy within the context of an AI-enabled future, ensuring that the policy impacts of AI serve the public good while protecting citizens and societies from AI's misuse.
- ▶ *Leveraging our expertise to develop human-centred **solutions** to meet the tremendous demand for innovative, actionable, high value, and globally appealing solutions for building safe, responsible, and inclusive AI and other advanced technologies.*



# Strategic Objectives

We have organized our goals for building research, policy, and solutions into four strategic objectives to achieve our vision and mission. The first two are specifically addressed to **research** and the second two to **policy and solutions**, but we expect research to inform policy and solutions and vice versa.

1. Significantly increasing the depth and breadth of **interdisciplinary** research and training spanning the sciences, social sciences, and humanities in the domain of AI and powerful technology.
2. Fostering and advancing **new fields** of research in the sciences, social sciences, and humanities to deepen our understanding of powerful technologies and human relationships to technology.
3. Transforming the world's approach to regulating powerful technologies through **regulatory innovation**.
4. Working with industry, governments, and civil society organizations to develop the legal and regulatory frameworks needed to **unlock AI for social good** to maximize the benefits that powerful technologies can bring to people across income, class, gender, ethnic, and global divides.

The purpose of these strategic objectives is to help us to prioritize activities and the allocation of resources. Specific annual plans are constructed with a view to connecting everything the Institute is doing to one of these four objectives.

## 1 / Interdisciplinarity

**We believe that developing the technical, legal, and regulatory infrastructure to mitigate harms and amplify the potential benefits of powerful technologies requires interdisciplinary collaboration that is deeper and broader than what we currently see in the AI and society ecosystem.**

Interdisciplinary research has been an ambition of the research university for several decades. Interdisciplinary hiring has transformed the faculty in many schools and departments, and many research grants today require applicants to include researchers from multiple fields. But the AI and society ecosystem is still not as interdisciplinary as it needs to be to tackle the tremendous challenges we face. AI researchers in computer science are seeking to understand more about

the social and economic impact of their work, but the community of researchers in other scientific fields, the social sciences, and the humanities who are engaging in a deep way with the technology that is rapidly developing is still too small. The Institute believes we need to significantly increase the scale of this community. Doing so requires a deep and intentional focus on how to build interdisciplinarity in this domain, and the development of concrete research infrastructure to support integrated research agendas.

Our approach to fostering interdisciplinary research began by appointing our inaugural Research Leadership team with U of T faculty members from computer science, engineering, law, philosophy, biology, and political science. These research leads were selected for their excellence, but also their enthusiasm for cross-disciplinary knowledge. We then dedicated several months to regular meetings to discuss how each of us sees the issues and challenges of AI. In doing so we have built shared language and vision across disciplinary divides. We have established a culture of translating our disciplinary jargon to colleagues. We then appointed our first complement of faculty affiliates with close attention to broad coverage across disciplines and we fostered a culture of cross-disciplinary engagement with a vibrant weekly seminar, intellectually accessible to and engaging for all. We made interdisciplinary engagement a centrepiece of our rubric for selecting graduate and faculty fellows, as well as the theme for our first academic conference (“Absolutely Interdisciplinary”), which we designed around moderated discussions in which we invite a leading technical researcher in AI to engage with a researcher from another domain in the sciences, social sciences, and humanities in conversation about a common question.

Going forward, this strategic objective will focus our efforts on furthering the breadth of our community, building infrastructure for interdisciplinary work with staff support for identifying potential collaborators in other disciplines, organizing discussion groups and workshops that introduce researchers to the concepts and framing used by different disciplines to analyze problems such as fairness in machine learning and seek to build shared frames and research agendas, and building truly integrated research proposals and teams for grant applications.

Finally, as the Institute grows, we will develop a campus-wide strategy, with the goal of being the hub of collaboration with the divisions represented on our Council of Deans around the theme of AI and society at U of T. We will also prepare a progress report for the Council of Deans for the fall of 2021 and will be looking for support from this group to scale some of our initiatives and accelerate our momentum.

### **How will we know we're successful?**

In addition to tracking our activities and being able to account for how these activities further the goal of increasing the depth and breadth of the interdisciplinary community working on AI and society research, we will also track these key metrics:

- **KPI: # of publications by and grants awarded to U of T faculty/postdocs/graduate students in Schwartz Reisman-related topics that are interdisciplinary (as determined by research methodology or co-authorship/team composition):**

*Methodology: Analysis of publications and grants identified with VPRI for purposes of (e.g.) Times Higher Education rankings.*

*Timeframe: Annual, set baseline 2018-2019.*

- **KPI (GIFT AGREEMENT METRIC): Publications:** Within the first three years of the Institute's operations, scholars affiliated with the Institute will author or co-author 50 papers and one book per year in the Institute's technology and society theme. At least 10 of the publications per year will involve co-authorship of a scholar in a social sciences or humanities discipline working with a scholar in a science discipline. The Institute's goal will be to increase the annual number of publications from this baseline to 35 per cent by the 10th year of the Institute.

## 2 / Field-Building

**We believe that new interdisciplinary fields of research are needed to address the challenges of ensuring that powerful technologies benefit all of humanity.**

AI and other powerful technologies are raising challenges that human societies have never confronted before and amplifying long-standing challenges in sometimes deeply worrying ways. Consider the impact of the social media platforms that have emerged over just the last 15 years. These platforms operate at massive scale as a result of recommendation algorithms that are shaping perceptions about facts and the contours of public debate, and sometimes coordinating behaviour offline such as participation in political protest. Highly targeted news feeds and advertising, curated by AI-driven analysis of massive quantities of data, have changed the results of elections and potentially increased levels of depression and isolation among young people. They have created new ways in which vulnerable communities can be victimized and subjected to discrimination. The returns to increasingly massive datasets have driven the collection and aggregation of previously unimaginable amounts of information about individuals, disrupting, transforming, and potentially destroying aspects of individual privacy and security. Mass digitization is rewriting the rules of economics, politics, and social interaction.

Scholars in technology and society studies have been thinking about these transformations for many years, documenting how technology impacts culture, inequality, child development and well-being, and more. And, in the past few years, relatively robust research communities in computer science have sprung up to develop, for example, new techniques for machine learning that protect against discrimination or loss of privacy.

In addition to supporting researchers in these existing fields of inquiry, we are focusing our efforts on building new fields. We believe, for example, that there are opportunities for new domains of inquiry related to the structure and behaviour of human normative systems, the meaning of responsibility and fairness in autonomous systems, the nature of artificial cognition, the factors that determine public acceptance of and optimal design of algorithmic decision-making, and the nature of human rights when people exist in not only physical but also digital domains.

This strategic objective orients the work of the Institute to prioritizing the identification and nurturing of new, as distinct from established, fields of inquiry. Our first academic conference, for example, will focus on three nascent fields of research — computational ethics, cooperative AI, and the evolution of collective intelligence in natural and artificial communities — and building new pathways in the young but established field of fair machine learning.

#### **How will we know we're successful?**

In addition to tracking our activities and being able to account for how these activities further the goal of fostering new fields of research, we will also track these key metrics:

- **KPI: # of novel research domains contributed by U of T faculty/postdocs/graduate students in Schwartz Reisman-related topics (as determined by peer survey):**  
*Methodology: Analysis of publications and grants identified by administration for purposes of (e.g.) Times Higher Education rankings.*  
*Timeframe: Long term (5-7 years).*
- **KPI (GIFT AGREEMENT METRIC): Scholars affiliated with the Institute:** We anticipate the number of scholars affiliated with the Institute will increase twofold from an initial 25 to 50 over the Institute's first 10 years.
- **KPI (GIFT AGREEMENT METRIC): Grants:** Scholars affiliated with the Institute will secure at least \$1 million per year of grants and external funding by Year 3 of the Institute, increasing to \$1.5 million per year by the end of Year 6, and \$2 million per year by the end of Year 10. We should note that grants to humanities and social sciences disciplines tend to be much smaller than those to science disciplines.

### 3 / Regulatory Innovation

**We believe that adequately controlling the potential harm and unlocking the benefits of powerful technologies will require radically new approaches in law and regulation.**



Beneficial AI will require more than new advances in technology, it will also require new approaches to regulation and governance. This is not merely a matter of rewriting laws or updating ethical principles to apply to AI systems (such principles have generally failed to prevent many AI systems from running roughshod over human values). Instead, fundamentally new approaches to governance and regulation are needed, along with new technologies, ventures, and partnerships.

The scale, speed, and global reach of AI and other powerful digital technologies far outstrips the capacity of the systems built in the 19<sup>th</sup> and 20<sup>th</sup> centuries to govern the industrial mass-manufacturing nation-state-based economy. We fundamentally believe that new regulatory technologies are required to keep up with AI. This generates two key challenges: how to incentivize investment in the creation of these technologies, and how to ensure that these technologies are responsive to democratic oversight.

For example, reducing the harms caused by allowing misinformation, bias, or hate to spread rapidly on social media platforms will not be possible using only the traditional techniques of creating legal rights or obligations for individuals suffering harm, or public prosecutors to vindicate in court. There are millions of online posts reaching billions of people around the globe on a daily basis. Reducing harm will require technologies that automatically detect and filter out misinformation, bias, and hate before they reach a platform. Technology companies are today attempting to build such technologies. But they are doing so in ways that are governed by their private commercial interests, and largely hidden from public scrutiny. We believe these technologies should (also) be built in independent third-party organizations and companies and in ways that are governed by democratic institutions and values and subject to democratic oversight. What sets us apart from other research and advocacy organizations focused on AI is that we are (not only) exploring what are the characteristics of automated recommender or content-moderation systems that promote healthy, safe, equitable communities, we are also exploring what regulatory regimes can incentivize and support investment and deployment of such systems.

To achieve this strategic objective, we will be supporting research and design work to develop new regulatory approaches, entrepreneurial work to foster the emergence of a robust sector in regulatory technologies, and policy work to build governments' understanding and adoption of new regulatory approaches and technologies. Key directions include the development of technical and legal approaches to building trusted computing environments, the establishment of regulatory markets that foster greater use of private-sector, technically-grounded certification and auditing services, and the acceleration of start-ups building regulatory technologies, such as software that automatically protects personal information in data transfers, or checks automated decision-making systems for bias.



In short, our goal is to provide new regulatory tools to align powerful digital technologies with democratic values and goals and promote their adoption in government.

### How will we know we're successful?

In addition to tracking our activities and being able to account for how these activities further the goal of regulatory innovation, we will also track these key metrics:

- **KPI: # of Schwartz Reisman-inspired governance mechanisms/designs implemented or piloted by governments, industry, or civil society organizations:**

*Methodology: Tracking of Schwartz Reisman-related initiatives and Schwartz Reisman media mentions in governance initiatives.*

*Timeframe: Long-term (5-7 years).*

- **KPI (GIFT AGREEMENT METRIC): Schwartz Reisman International Conference:** The conference will attract 300 registrants across multiple stakeholder groups per year; we will aim to increase this number to 500 per year by Year 5 after the Institute's establishment.

## 4 / AI for Social Good

**We believe that AI and other powerful technologies can advance human welfare by helping to improve health, education, sustainability, equality and inclusion, access to justice, and more, and that deliberate efforts are needed to ensure that AI is built for public benefit.**

Most AI technology today is being built inside large private sector technology companies and attracting billions in private investment dollars. While there is significant public benefit that comes from private sector investment and innovation, market incentives do not adequately promote the production of public goods. AI can potentially significantly improve our capacity to produce public goods, such as access to fair and neutral decision-making in governments and courts, action to mitigate the impacts of climate change, and public health, education, housing, and transportation. Ensuring that AI development includes AI for social good will require deliberate efforts to specifically promote investment in, and remove obstacles in, the public and philanthropic sectors.

A central obstacle to investment and development of AI and other digital technologies in the public and philanthropic sector is an outdated approach to data governance. This was made evident in many ways during the COVID-19 pandemic. Machine learning technologies could have helped analyze health data to improve risk modeling and treatment but governments around the world were stymied by complicated and slow regimes to ensure privacy protection. Digital contact tracing and analysis to detect outbreaks and transmission patterns were similarly blocked by our

inability to rapidly build data processing infrastructure that was trusted and understood by the public. Data sharing agreements under existing approaches frequently take months, if not years, to negotiate, and often fail entirely as legal advisors caution against running afoul of privacy legislation or relinquishing valuable intellectual property. In our work, we have heard about the same obstacles interfering with efforts to assemble data for global climate action, the treatment of chronic disease such as diabetes, and more humane systems of migration.

Other obstacles to AI for social good include a lack of appropriate regulatory infrastructure which discourages or prohibits investment (e.g., in autonomous marine vehicles that could be deployed to monitor climate change), and a lack of public sector knowledge or resources (e.g., to build AI-assisted adjudication systems).

Although these obstacles also slow innovation in the private sector, the public sector is doubly impacted by poor incentives to lobby for the regulatory changes needed to unblock AI development. In achieving this strategic objective, we are motivated to:

- ▶ Identify the dynamics and barriers that block development and deployment of AI and other technologies that benefit groups underserved by market incentives.
- ▶ Cultivate regulatory innovations that generate technology that benefits groups underserved by market incentives.
- ▶ Build AI and governance capacity in low- and middle-income countries as well as historically marginalized groups such as Indigenous and migrant populations.

#### **How will we know we're successful?**

In addition to tracking our activities and being able to account for how these activities further the goal of promoting AI for social good, we will also track these key metrics:

- **KPI: # of Schwartz Reisman-inspired governance mechanisms/designs implemented or piloted by governments, industry, or civil society organizations that specifically support AI for social good.**

*Methodology: Tracking of Schwartz Reisman-related initiatives and Schwartz Reisman media mentions in governance initiatives.*

*Timeframe: Long-term (5-7 years).*

## Conclusion

It is widely recognized that the advances in AI we have witnessed over the last few years are posing fundamental challenges to ethics, democracy, economics, and what it means to be human. It is for this very reason that the vision for the Schwartz Reisman Institute for Technology and Society is to be the world's leading institute for innovative research and practical solutions that help ensure AI and other advanced technologies benefit **all** of humanity. From a governance perspective, the Institute is overseen by a multi-divisional Council of Deans, with support from an external Advisory Committee consisting of key industry leaders and renowned subject matter experts in AI, governance, ethics, and other relevant fields.

With its mission to deepen our knowledge of technologies, societies, and what it means to be human by actively integrating research across traditional boundaries and building human-centered solutions that really make a difference, the Institute is reframing the conversation around AI, so that instead of talking only about what we want from technologies such as AI, we are talking also about how we are going to govern a world infused with AI. This is reflected in our four strategic objectives to: (1) increase the range and depth of interdisciplinary research on the AI and society ecosystem at U of T and beyond; (2) build new fields of research to deepen our understanding of powerful technologies and human relationships to technology; (3) transform the world's approach to regulating powerful technologies (regulatory innovation); and (4) work with industry, governments, and civil society organizations to maximize the benefits that new powerful technologies can bring to people across income, class, gender, ethnic, and global divides (AI for social good). In addition to our Gift Metrics, we have also identified a number of KPIs to measure our success in achieving these objectives.

Finally, it is clear that AI poses some of the deepest challenges we now face. Achieving the promise of AI – and avoiding its negative consequences – requires us to build the technological, legal, and regulatory infrastructure to ensure that AI advancements are responsible and aligned with human goals, ethics, and values – and, conversely, to ensure that our existing legal and regulatory systems do not either unintentionally or unnecessarily block AI innovation. Only by addressing these two challenges can we hope to ensure that our latest, and most powerful, invention – intelligence that lives in machines and not only brains – serves the global goal of human societies that are vital, peaceful, inclusive, and just.

## Appendix A: Landscape of Comparable AI and Society Institutes

- [Center for Human-Compatible Artificial Intelligence \(Berkeley\)](#)
- [Human-Centered Artificial Intelligence \(Stanford\)](#)
- [Future of Life Institute \(Cambridge, Massachusetts\)](#)
- [Schwarzman College of Computing \(MIT\)](#)
- [Berkman Klein Center for Internet & Society \(Harvard\)](#)
- [The AI Now Institute \(NYU\)](#)
- [Center for Security and Emerging Technology \(Georgetown University\)](#)
- [The Alan Turing Institute \(UK\)](#)
- [Centre for the Governance of AI – Future of Humanity Institute \(Oxford\)](#)
- [The Ethics in AI Institute \(Oxford\)](#)
- [Oxford Internet Institute \(Oxford\)](#)
- [Center for the Study of Existential Risk \(Cambridge, UK\)](#)
- [Leverhulme Centre for the Future of Intelligence \(Cambridge, UK\)](#)
- [Humanising Machine Intelligence \(Australian National University\)](#)

# Appendix B: Council of Deans & Terms of Reference

## Terms of Reference

The president of the University will appoint a Council of Deans that will provide oversight for the Schwartz Reisman Institute for Technology and Society. This Council of Deans will be chaired by the president or one of U of T's vice presidents, and will include representatives from the Division of the Vice-President & Provost, the Division of the Vice-President, Research & Innovation, and representative deans. It will appoint the director of the Schwartz Reisman Institute (and holder of the Schwartz Reisman Chair in Technology and Society) and is responsible for approving the Institute's strategy and budget.

## Members (2020-2021)

1. **Ted Sargent**, Vice President, Research & Innovation (Co-Chair)
2. **Melanie Woodin**, Dean, Faculty of Arts & Science (Co-Chair)
3. **Adalsteinn (Steini) Brown**, Dean, Dalla Lana School of Public Health
4. **Jutta Brunnée**, Dean, Faculty of Law
5. **Kenneth Corts**, Interim Dean, Rotman School of Management
6. **Wendy Duff**, Dean, Faculty of Information
7. **Alexandra Gillespie**, Vice President & Principal, UTM
8. **Wisdom Tettey**, Vice President & Principal, UTSC
9. **Trevor Young**, Dean, Temerty Faculty of Medicine
10. **Dexter R. Voisin**, Dean, Factor Inwentash Faculty of Social Work
11. **Christopher Yip**, Dean, Faculty of Applied Science & Engineering
12. **Cheryl Regehr**, Vice President and Provost
13. **David Palmer**, Vice President, Advancement (invited guest)

## Appendix C: Gift Agreement Metrics

**Publications:** Within the first three years of the Institute's operations, scholars affiliated with the Institute will author or co-author 50 papers and one book per year in the Institute's technology and society theme. At least 10 of the publications per year will involve co-authorship of a scholar in a social sciences or humanities discipline working with a scholar in a science discipline. The Institute's goal will be to increase the annual number of publications from this baseline to 35 per cent by the 10th year of the Institute.

**Scholars Affiliated with the Institute:** We anticipate the number of scholars affiliated with the Institute will increase twofold from an initial 25 to 50 over the Institute's first 10 years.

**Grants:** Scholars affiliated with the Institute will secure at least \$1 million per year of grants and external funding by Year 3 of the Institute, increasing to \$1.5 million per year by the end of Year 6, and \$2 million per year by the end of Year 10. We should note that grants to humanities and social sciences disciplines tend to be much smaller than those to science disciplines.

**Courses:** We will develop one course per year.

**Other Events:** Six international speakers per year, drawing an average of 100 participants per event. 50 local events per year, drawing an average of 50 participants per event.

**Schwartz Reisman International Conference:** The conference will attract 300 registrants per year; we will aim to increase this number to 500 per year by Year 5 after the Institute's establishment.