



Telling Stories

On Culturally Responsive Artificial Intelligence

19 Stories

EDITED BY:

Ryan Calo, Batya Friedman, Tadayoshi Kohno,
Hannah Almeter and Nick Logler



Telling Stories: On Culturally Responsive Artificial Intelligence



Telling Stories: On Culturally Responsive Artificial Intelligence

19 Stories

Edited by
Ryan Calo, Batya Friedman, Tadayoshi Kohno, Hannah Almeter and Nick Logler

University of Washington Tech Policy Lab, Seattle

Copyright © 2020 by University of Washington Tech Policy Lab. All rights reserved.

For information about permission to reproduce selections from this book, contact:

University of Washington
Tech Policy Lab
4293 Memorial Way NE
Seattle, WA 98195.

Printed by The Tech Policy Lab, USA.

First edition.

ISBN: 978-1-7361480-0-6 (ebook)

Library of Congress Control Number: 2020924819

Book design by Patrick Holderfield, Holdereaux Studio
www.holdereauxstudio.com

Illustrations by:

Tim Corey, Colibri Facilitation: pgs - 22, 24, 28, 29, 40, 41, 43, 45, 48

Akira Ohiso: pgs - 18, 26, 27, 36, 42, 44, 50, 51, 52, 53

Elena Ponte: pgs - 20, 30, 31, 34, 38, 46, 47, 54

Free digital copy available at <https://techpolicylab.uw.edu/>

About the Global Summit on Culturally Responsive Artificial Intelligence

The Global Summit on Culturally Responsive AI is part of the Tech Policy Lab's Global Summit Initiative (2016-present). Our Global Summits convene twenty to thirty thought leaders from around the world representing design, ethics, governance, policy, and technology. Summits aim to frame and begin progress on pressing grand challenges for tech policy, providing opportunities for collaboration on global and local issues. The 2018 Global Summit structured conversation around grand challenges for developing and disseminating artificial intelligence technologies that maintain respect for and enhance culture and diversity of worldview. *Telling Stories: On Culturally Responsive Artificial Intelligence* encapsulates the wisdom from these conversations in the form of 19 short stories which can be told and retold.

Disclaimer

Note from the Editors. These stories are mostly works of fiction—names, characters, businesses, events, and incidents are the products of the author's imagination; any resemblance to actual persons, living or dead, or actual events is purely coincidental. Where a story has a basis in fact, names and identifying details may have been changed to protect the privacy of individuals. The inclusion of any brands or organizations in these stories are not meant to indicate any relationship or endorsement from said organization. The opinions expressed in the stories belong to the authors and do not reflect the views of the editors, Tech Policy Lab, University of Washington or our funders unless otherwise indicated.

About the Tech Policy Lab

The Tech Policy Lab is a unique, interdisciplinary collaboration at the University of Washington that aims to enhance technology policy through research, education, and thought leadership. Founded in 2013 by faculty from the Paul G. Allen School of Computer Science & Engineering, the Information School, and the School of Law, the Tech Policy Lab aims to bridge the gap between technologists and policymakers and to help generate wiser, more inclusive tech policy. <https://techpolicylab.uw.edu/>

**For Ian
(1965 - 2019)**

Table of Contents

| | | | |
|-----------------------|----|--|----|
| Introduction | 10 | The Stories | |
| Audience | 12 | United States of Brazil | 18 |
| Where the authors are | 14 | Watch Them | 20 |
| | | Br'er Rabbit and the Mirror Baby | 22 |
| | | Face Changes | 24 |
| | | The Box | 26 |
| | | In Search of Robot Love | 28 |
| | | The Boy Whose Light Went Out | 30 |
| | | What Justice | 32 |
| | | True Love? | 34 |
| | | Lia's AI Future 2036 | 36 |
| | | A Well-Meaning Robot | 38 |
| | | With the Lights Out, It's Less Dangerous | 40 |
| | | Lost Knowledge? Forbidden Knowledge? | 42 |
| | | Language Matters | 44 |
| | | Do We Have the Right Metrics? | 46 |
| | | Moussas of the Savanna | 48 |
| | | Bug | 50 |
| | | The Goddess | 52 |
| | | Of Genies and Magic Lamps | 54 |

| | |
|--|----|
| Appendix | 57 |
| Developing the Stories | 58 |
| Story Writing Process | 58 |
| Issues, Questions, Challenges, and Opportunities | 60 |
| Story Toolkit | 63 |
| Part I: Landscape | 63 |
| Part II: Your Narrative | 71 |
| Part III: Positioning Impact | 77 |
| Who the authors are | 82 |
| Story Index | 86 |
| Acknowledgments | 89 |

Introduction

In the summer of 2018, we brought experts from around the world to Seattle to tell original stories about artificial intelligence. Our goal was to use the power of storytelling as a way to gain traction on a pressing but difficult question: how to make AI more culturally responsive.

Technology often is not culturally responsive. Instantiations of information communications technology—whether in devices or software—can export the assumptions, experiences, and values of its developers. And these assumptions do not always carry over to different contexts or world views.

Consider the example of the Maori and the “situational when.” The Maori—an Australian aboriginal community—do not view important events as necessarily beginning and ending at a certain time and date. A funeral, for example, begins when the correct set of people and circumstances align. As smartphones proliferate among young Maori, the salience of calendar apps and reminder notifications threaten to undermine community practices and understandings around time.

Advances in AI are occurring all over the world. But the vast majority of development happens in just a few places—and then only by a specific and rarified population of technologists. As AI-enabled devices and applications proliferate across the globe, the assumptions, experiences, and values of these few groups will travel along with the technology.

How do we begin to introduce the astonishing range of cultural perspectives to technologists, policymakers, educators, and others? How do we build a bridge

between communities and their advocates and the individuals and institutions developing and regulating AI?

Given the task at hand—representing a wide and nuanced set of perspectives to a broad audience—we decided to look to storytelling as our medium. Cultures diverge, but where there are people, there are stories. We are aware of no human society that does not make use of stories to understand and describe the world and their place in it.

Stories tap into people’s heartfelt emotions and convey their deeply held beliefs and cultural ways of being and knowing. And they can do so succinctly, compellingly, and with compassion. A carefully crafted story can convey the critical heart of the matter without being lost in distracting details. In this way, stories can position technologists and policymakers to hear the concern of those who live their lives within diverse world views.

Stories can be resisted and will not land the same way with all listeners. However, regardless of whether resisted or not, once told, the message of a story (and the associated discussions and thoughts about it) persists almost like a physical object. Good

stories constitute a vehicle for sharing a perspective, one that can be reinterpreted, retold, and elaborated. A good story can also create discussion, bringing forward new insights and perspectives on challenging topics. Successful narrative takes on a life of its own. We hope this constellation of features will help us use stories to build an enduring bridge between communities and to shape wise and inclusive policymaking in a way other media might not.





Fabric Pieces

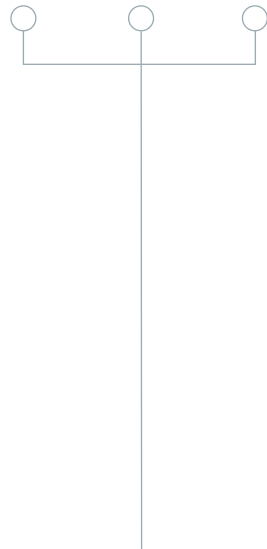
Our story writing began with fabric pieces. Each story writer brought a piece of culturally meaningful fabric to the Global Summit. Participants introduced themselves to each other through their fabrics, explaining why they chose that fabric and its cultural relevance. A lace handkerchief from Chile was used in a courtship dance with youth; a bolt of rich blue cloth with gold tear drops is typical of women's dress in Rwanda. The fabric pieces invited enactments—storytelling. They catalyzed conversations, illustrating how a particular detail can go a good distance toward conveying context and situation as well as social and technological structure.

Process

The story-writing process took place over two-and-a-half days. Authors first surfaced the issues, questions, challenges, and opportunities they saw that concern culturally responsive AI, recording each issue on a river rock with a permanent marker to make a visible and persistent reminder of the idea. Next, we discussed how stories work and why they matter. Then to the work of story generation. Authors worked in small groups to write, reflect, read aloud, and write again, as their individual stories developed. The writing phases initially emphasized the cultural landscape, then technology, and finally social and societal impact. Details on the story-generation process and supporting materials can be found in the appendix.

Audience

Our hope for *Telling Stories* is that its contents will help build a bridge among policymakers, technologists, educators, community organizers, and others. Humans have many roles and come from many contexts, but we almost all share a visceral understanding of and response to stories. The narrative format of *Telling Stories* aims to tap into our relationships with stories to generate more culturally responsive AI and AI policy. Here we highlight some of the audiences our authors and other participants had in mind for the project.



POLICYMAKERS Policymakers—broadly understood as individuals positioned to set and implement policy—have access to an increasing variety of materials related to AI. These include principles, reports, articles, and other resources. *Telling Stories* hopes to expose policymakers to new voices, ideas, and frames that can help them communicate risks and legitimate stakeholders and their concerns as a means to craft more wise and inclusive tech policy.

TECHNOLOGISTS Technologists design, build, deploy, test, and maintain technical systems. The stories contained in this volume help to identify, explore, and dramatize cultural values to which a relatively homogeneous technical community might not have been exposed.

EDUCATORS Educators at all levels are increasingly incorporating social, cultural, and ethical impacts as a component of technology instruction. *Telling Stories* can be a resource for educators aiming specifically at exposing students to a broad array of world views as well as the technological implications.

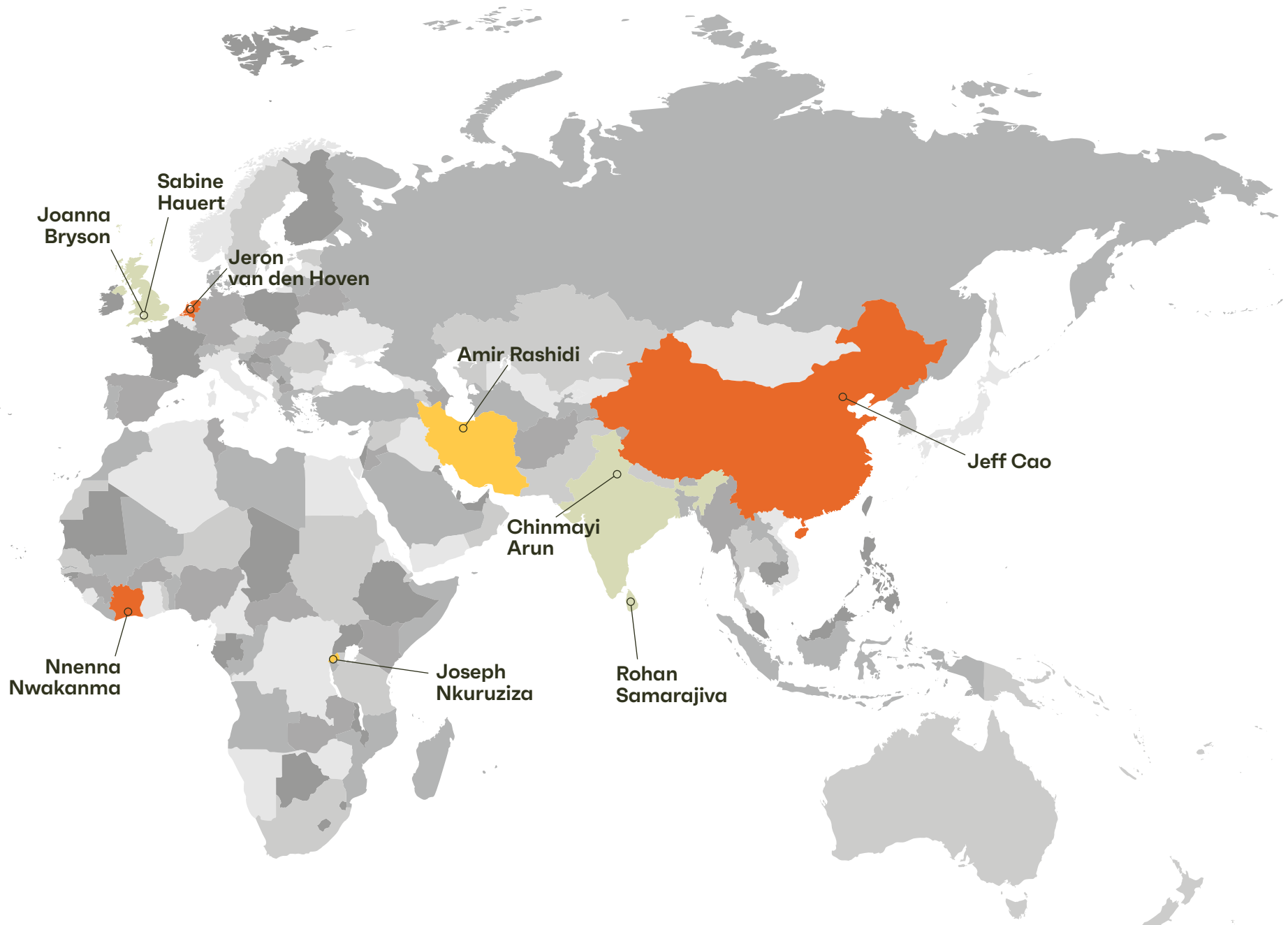
COMMUNITY ORGANIZERS Community organizers work with communities to determine and advance the community’s specific needs and values. *Telling Stories* describes one method by which to bring out community voice and, in some instances, may serve as a narrative record of values communities hope to foreground.

OTHER We want *Telling Stories* to speak to a wide range of audiences. We hope no one would put this book down because of who they are. Indeed, we selected the medium of storytelling specifically with the hope of reaching across and building bridges among different cultures, world views, and sectors.

Where the authors are

To develop, think through, and tell stories around culturally responsive AI, we must bring together diverse groups of stakeholders. This book gathers stories from authors from around the world—experts in art, ethics, policy, political science, and technology from the continents of Africa, Asia, Europe, as well as North and South America. We invite you to read their stories and explore the author biographies at the back of the book. We further invite you to consider what stakeholder groups are not represented in our stories and seek to understand what stories those people might have to share.







The Stories



United States of Brazil

Dennys Antonialli

InternetLab
Brazil



HE WAS HOMESICK. AFTER TWENTY YEARS WORKING ABROAD, he missed so many things about his country. He missed the flavorful and extremely unhealthy food, with lots of meat and salt. He missed how people would always be late for things and still find it perfectly normal. He missed not having to make a reservation for every meal and then having to spend an hour

waiting for a table by the bar chatting with the bartenders. He missed how people from work would get drunk and be slightly inappropriate at happy hours without being judgmental about it. He missed the heavy traffic jams that he could use to listen to all of his guilty pleasure songs. He missed the crazy weather changes and how he got soaked when he forgot his umbrella.

He missed having that unplanned life. He missed losing control. Life in Brazil was certainly chaotic, but Brazilians knew how to turn chaos into fun. It was part of the culture. It was part of his culture. And he was excited to go back.

None of that was possible in the United States, where he had been living for so long. His personal AI assistant took care of everything. It would advise him on his diet and make reservations for the exact type of food he wanted to eat. His autonomous car would be ready to take him everywhere, so he would never be late and would never spend extra time in traffic. He would never get drunk with people from work because his smartwatch just wouldn't let him. There was no room for error in that very artificially intelligent society.

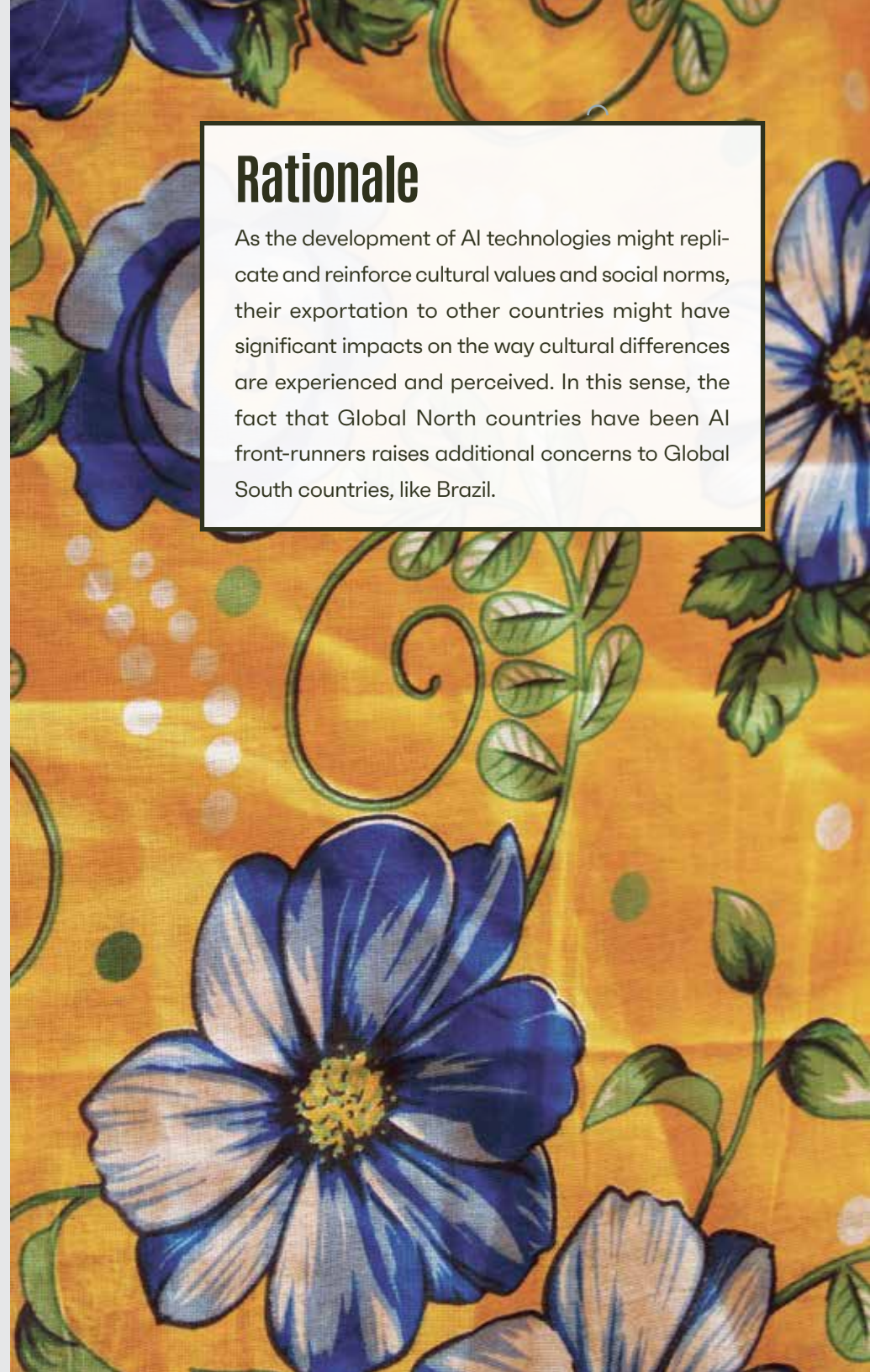
But he was finally leaving that perfect world behind.

As he leaves the plane in São Paulo, he is puzzled by the fact that everyone is wearing the same smartwatches he has around his wrist. They all have the same AI assistants. Autonomous cars are picking people up at the airports. There are no lines. Not even in the restaurants. No one is drunk at happy hours. No one is caught off guard by the storms. Everything is working perfectly.

It just didn't feel like home anymore. It was not home anymore.

Rationale

As the development of AI technologies might replicate and reinforce cultural values and social norms, their exportation to other countries might have significant impacts on the way cultural differences are experienced and perceived. In this sense, the fact that Global North countries have been AI front-runners raises additional concerns to Global South countries, like Brazil.



Watch Them

JYOTI PANDEY WAS BRUTALIZED BY FOUR MEN ON A BUS ON the night she went out to watch a movie, eat popcorn, and laugh. She was dead by the morning and was in headlines around the world. The media called her “India’s daughter.”

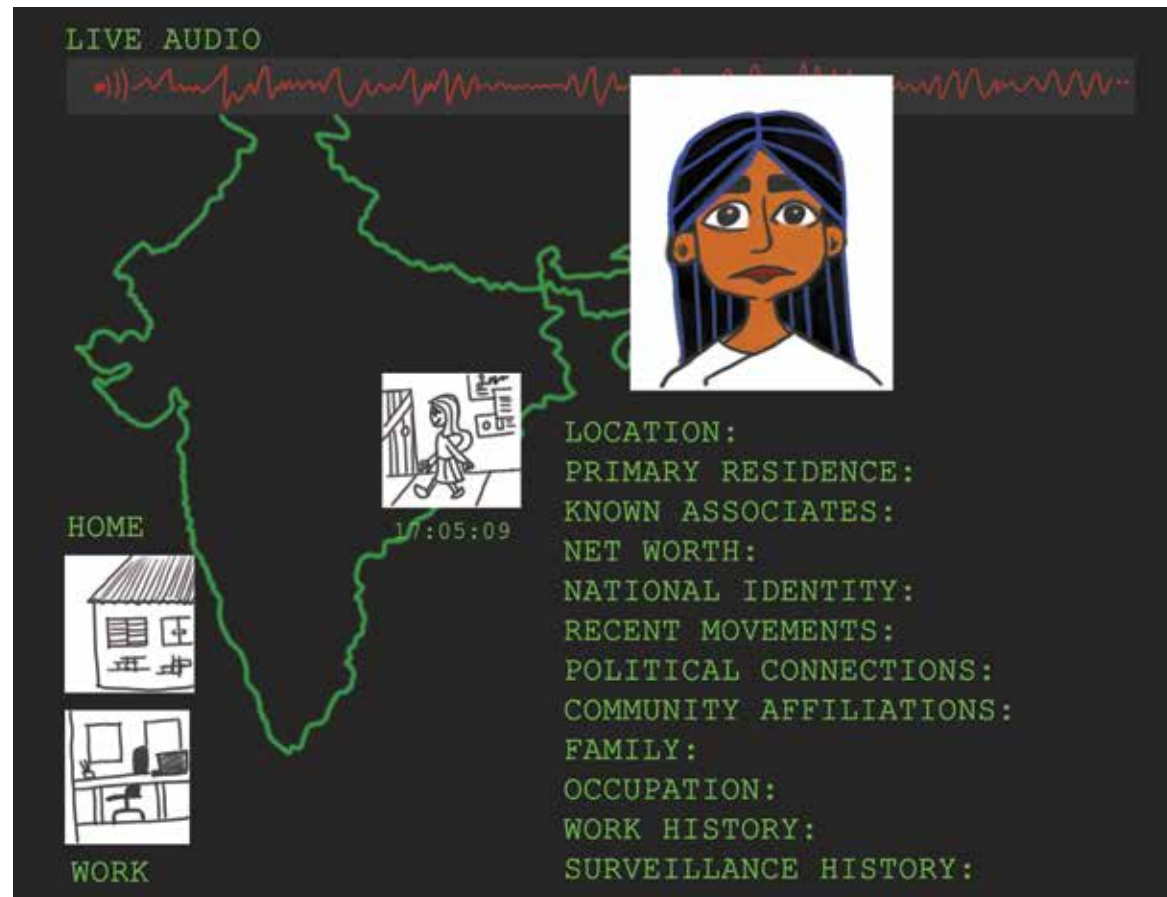
“India is disgraced. We must do a better job of watching our daughters,” said the CEO of Nazar to the Minister of IT over scotch and shammi kebabs at the club. A month later, Nazar launched new phones with GPS that didn’t switch off. After that, India made it compulsory for all phones to have a GPS that didn’t switch off.

Then Nazar imported CCTV with facial recognition technology. “What a great way to watch our daughters and keep an eye on criminals,” said the Minister of IT to the CEO of Nazar over their customary scotch and kebabs. The Minister gave Nazar access to the national identity card database, so the ID photographs could be used to identify the faces on CCTV. Every major city in India bought Nazar’s technology.

When two young women were sexually assaulted and murdered at an upscale Chandigarh bar that even the Minister’s son frequented, the police went to Nazar. Chandigarh used

Chinmayi Arun

National Law University, Delhi
India



Nazar's facial recognition CCTV, and the bar had additional Nazar cameras. However, Nazar's systems failed. The CCTV feed and GPS signals from the bar could not be found. One news website asked, "Is Nazar's technology bogus?" It was the website no one read, and the only one that could afford to lose Nazar's advertisement revenue.

The mainstream media focused on why two young women were in an unsafe bar at night. Some young women who had joined Nazar, back when the CEO was their hero, quit the company.

"We must keep women off the streets for their own safety," said the Minister and the CEO to media owners and editors over scotch and kebabs. The Minister and the CEO unveiled a prizewinning plan. Artificial intelligence would monitor integrated datasets of GPS signals, smart CCTV, and national identity card records. The AI would alert the police when women were in dangerous places at dangerous hours.

One day a woman failed to secure her veil properly. The veil fell for a few seconds. The street CCTV camera recognized her, and the AI alerted the police. The police raid was on live TV. Reporters commented breathlessly as the cameras picked up women with alcohol, cigarettes, and Maoist literature. They were marched to police vans in handcuffs. The voiceovers said these women had "abused their freedom."

A young hacker watched her girlfriend being arrested. She rummaged through her secret stash from her Nazar days. Soon the whole country was watching the Minister of IT's son assaulting and murdering two women in a Chandigarh bar. Nazar gave its AI lists of individuals to black out from its feeds. But it was too late.

Rationale

The story is a cautionary tale: the powerful actors who will design and deploy AI are the same people who effectively control the cities in which women cannot walk safely on the street at night. Their choices of technology design and deployment may not evidence a change of heart.

Br'er Rabbit and the Mirror Baby

Joanna Bryson

University of Bath
England

BR'ER FOX WANTED TO BE PUT IN CHARGE OF THE HEN HOUSE, so he decided to become a politician and get elected. He knew the hens would never vote for him, and there weren't many foxes. So Br'er Fox decided to focus on Br'er Rabbit. But Br'er Fox knew that Br'er Rabbit wouldn't trust a fox either. So Br'er Fox went to Local Tech Giant and asked them to build a Mirror Baby, to keep Br'er Rabbit distracted.

Br'er Fox set the Mirror Baby on a log where he knew Br'er Rabbit would likely come by, and sure enough, before long Br'er Rabbit did come by and saw Mirror Baby sitting there.

"Hello!" said Br'er Rabbit.

"Hello!" said Mirror Baby.

Br'er Rabbit took a better look at Mirror Baby and basically saw himself, but he didn't recognize himself.

"You are one good-looking rabbit," said Br'er Rabbit to Mirror Baby.

"You are one good-looking rabbit," said Mirror Baby.

Br'er Rabbit took Mirror Baby home, where he got more and more fascinated by it. He stopped going out even to work, because he so liked talking to Mirror Baby.

Unbeknownst to Br'er Rabbit, the microphone in Mirror Baby transmitted back to its owner, Br'er Fox.

Br'er Fox was still a little worried. He wasn't sure getting rabbits to stay home would win him an election, because there still weren't many foxes and there still were a lot of chickens. But as Br'er Fox listened to what Br'er Rabbit was saying to Mirror Baby in his home, he realized Br'er Rabbit really thought Mirror Baby was a person. That gave Br'er Fox an idea.

Br'er Fox went to Local Tech Giant again and said, "I want to rent one hundred Mirror Babies next Tuesday. I want them all to use the brain that's been learning from Br'er Rabbit, but

I want you to add in one piece of code that makes sure they go to the polls that day, and they vote for me."

"No problem!" said Local Tech Giant, who didn't have an ethics board or care very much about politics.

So Tuesday came and the polling station was entirely full of chickens and Mirror Babies. When the chickens looked at the Mirror Babies, they saw chickens. But these new chickens were all talking like Br'er Rabbit, except they said they were going to vote for Br'er Fox.

"Where did all these good-looking shiny new chickens come from?" asked the chickens. "And why are they voting for Br'er Fox? And why do they care more about carrots than seed? Speaking of carrots, where are the rabbits?" asked the chickens.



But the chickens didn't know very much about Mirror Babies, and before they figured out what was happening, Br'er Fox was in charge of the hen house. Now, ironically, it was Local Tech Giant's hen house, and suddenly there weren't any eggs for Local Tech Giant to eat. And it turned out, no one who cared about carrots was doing any work, so there wasn't anything for Br'er Rabbit to eat either. So Br'er Fox was king for a day, in fact for a week, but after a month he starved too.

The fabric is a sustainable cloth bag I carry with me because I don't like wasting plastic. It had the imprint of the European Human Behaviour Evolution Association meeting on it from when that meeting was in Durham, which is a beautiful old small city with a fantastic castle and cathedral. I liked this because AI is absolutely a part of human evolution—in fact, language is the key to our advantage over all the other species we are now owning or denying, and you could say that language is AI.

Rationale

People are easily fooled into believing that AI is or will become like them, but in fact AI is a machine that can be designed to do whatever is wanted by its owner, by its manufacturer, or sometimes even by a hacker neither the owner nor the manufacturer know about.



Reflections

- Everyone involved with AI is responsible for ensuring it is used well—the owner, the user, the manufacturer, and governments too.
- New technology gives new power in new ways, but the consequences of expressing that power can be hard to predict.
- No one is expendable. We all need to protect the weakest members in our society, because they are the majority of our society. If we do not maintain our connection to other humans, one day we will be the weak, disposable ones.

Face Changes

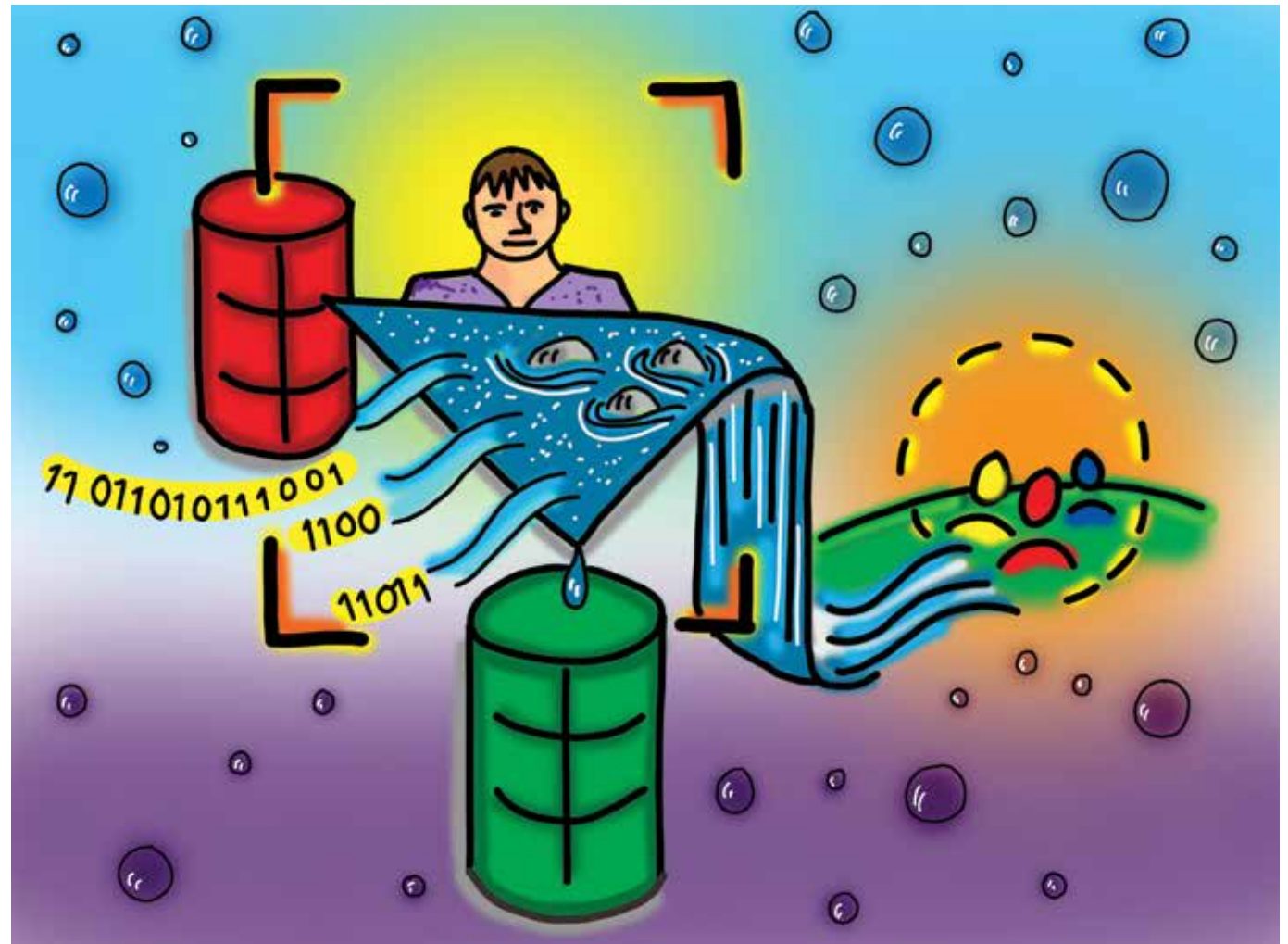
HASAN TOLD ME HE BOUGHT A SMARTPHONE SOON AFTER 3G networks reached his desert village near the Chinese border with Pakistan. He said that when he found out about WeChat, he felt that the story of his life as a Muslim really began. As he developed a pious online persona, social media algorithms pushed him to deepen relationships with Uyghur migrants in a mosque community in the city. Those friends helped Hasan

and his wife to find a sense of belonging in Ürümqi and to build networks to find work.

Around this same time, a group of young Uyghurs who were attempting to flee the country attacked a crowd of Han train passengers. Motivated by this—China's "9/11"—the state began to partner with tech companies to anticipate other forms of Uyghur insurgency. They built an AI program to analyze Uyghur

Darren Byler

University of Washington
United States



speech, text, and Islamic symbols on WeChat.

Back in Hasan's village, the police went to his home, telling his parents they would be arrested if Hasan did not return. Hasan blinked away his tears as he told me that he was being forced to turn himself over to the police. He said many of his friends had been "disappeared" when they returned to their villages. Several weeks later, he reluctantly agreed to leave. In less than a month his WeChat account was erased. Like Hasan feared, he was disappeared into a reeducation camp system with hundreds of thousands of others deemed digitally "unsafe."

I lost track of Hasan soon after this, so I am forced to imagine how he has survived. Based on state contractor reports, I know there are no blank spots in the camps. Even in the bathroom detainees are watched by cameras. The cameras feed an AI system that is building an image database of the emotions they capture from detainees' faces. Those who look angry are subjected to physical punishment. This reminds me of something Hasan told me before he disappeared. He said that when he had been arrested in the past he had learned how to remain expressionless. "Otherwise you will be beaten senseless," he said. Although Hasan has no doubt learned to maintain a blank affect, I imagine that he has found other ways to stay emotionally alive. Maybe he silently chants Sufi poems. Perhaps native Uyghur songs of survival worm their way through his mind and play on repeat. Maybe he dreams about the dyed silk dress he wished he could have given his two-year-old daughter at the end of Ramadan.

This hand-dyed silk cloth called etles was given to me as welcome gift by Uyghur silk weavers near Hasan's hometown.

Rationale

In the face of experiments in political and emotional surveillance, people refuse to fully detach from intimate social relations and deep cultural values.

Reflections

Hasan's story demonstrates how social media algorithms can help people find a sense of belonging in material and virtual communities; at the same time, however, participating in such communities can be used as evidence of "extremism" if forms of ethnoracialization such as Islamophobia are institutionalized. Hasan's story also shows that one's emotional life cannot be fully erased, even as it becomes the target of AI systems. I have used the pseudonym Hasan to protect his family from further harm. Parts of his story first appeared in my PhD dissertation, "Spirit Breaking: Uyghur Dispossession, Culture Work and Terror Capitalism in a Chinese Global City" (University of Washington, 2018), as well as in the Economist (June 2, 2018).

The Box

Ryan Calo

University of Washington
United States

JOSHUA MARVELS AT THE MACHINE HE MADE. VERSATILE. User-centric. Its capabilities speak for themselves. The Box gobbles, digests, spits out information of nearly any origin.

Why should only a handful of humans on the West Coast of the United States have access to The Box? Joshua resolves then and there to send premarket units to every populated continent. Distributing The Box will take resources. Joshua thinks, *At this price, I could feed everyone in South Seattle.* And then, he thinks, *Teach a man to fish.* He has the money and he has the time and he gives himself permission.



Time passes. How are all the many, scattered people using his technology? Joshua searches for videos and news articles. He reaches out through The Box itself, federated but connected, and solicits stories. Reports flow in: Daan has planned a wedding. Azra used The Box to orchestrate a short squeeze. Niran reads by its light. The Box is a doorstep in Devonshire. In Japan, it powers an exhibit involving dolls. Someone in Samara is collecting Boxes to decorate a zealously modern bathroom.

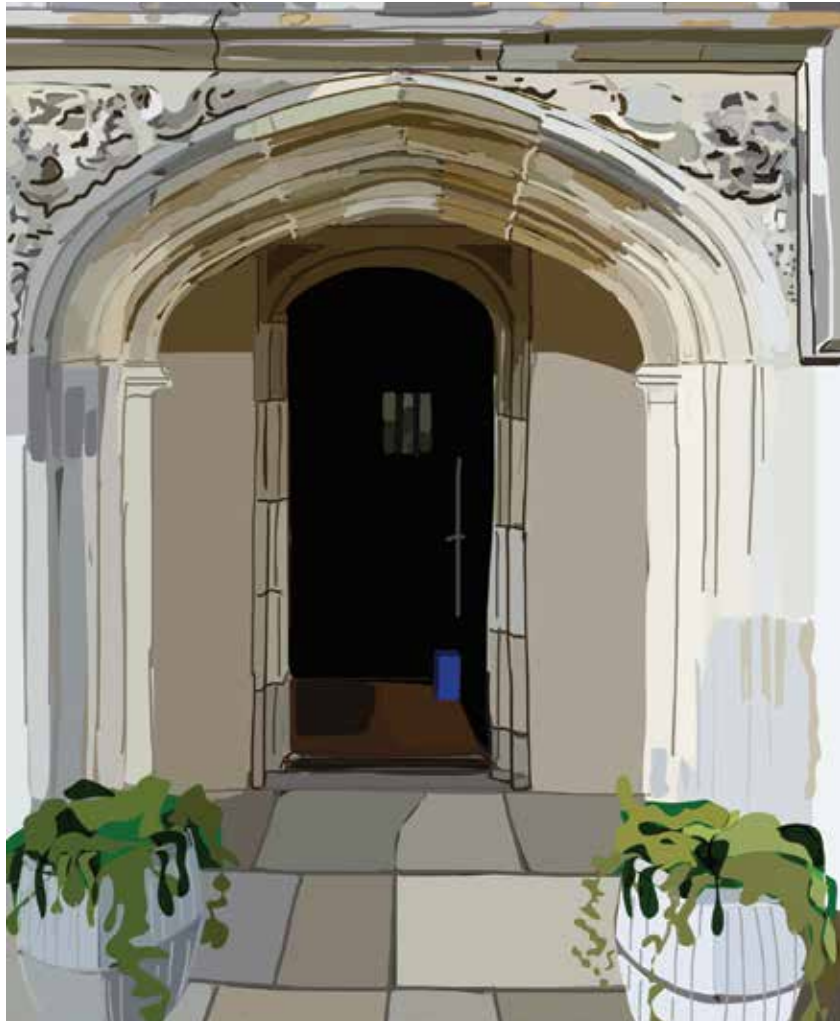
Is this everything? Where is the sea change? Where is the army of young Chinese girls with revolutionary tendencies?

More dispatches: Boxes have been confiscated. Boxes have been denounced. The Box records a fatal police shooting from its position under the arm of a Black American.

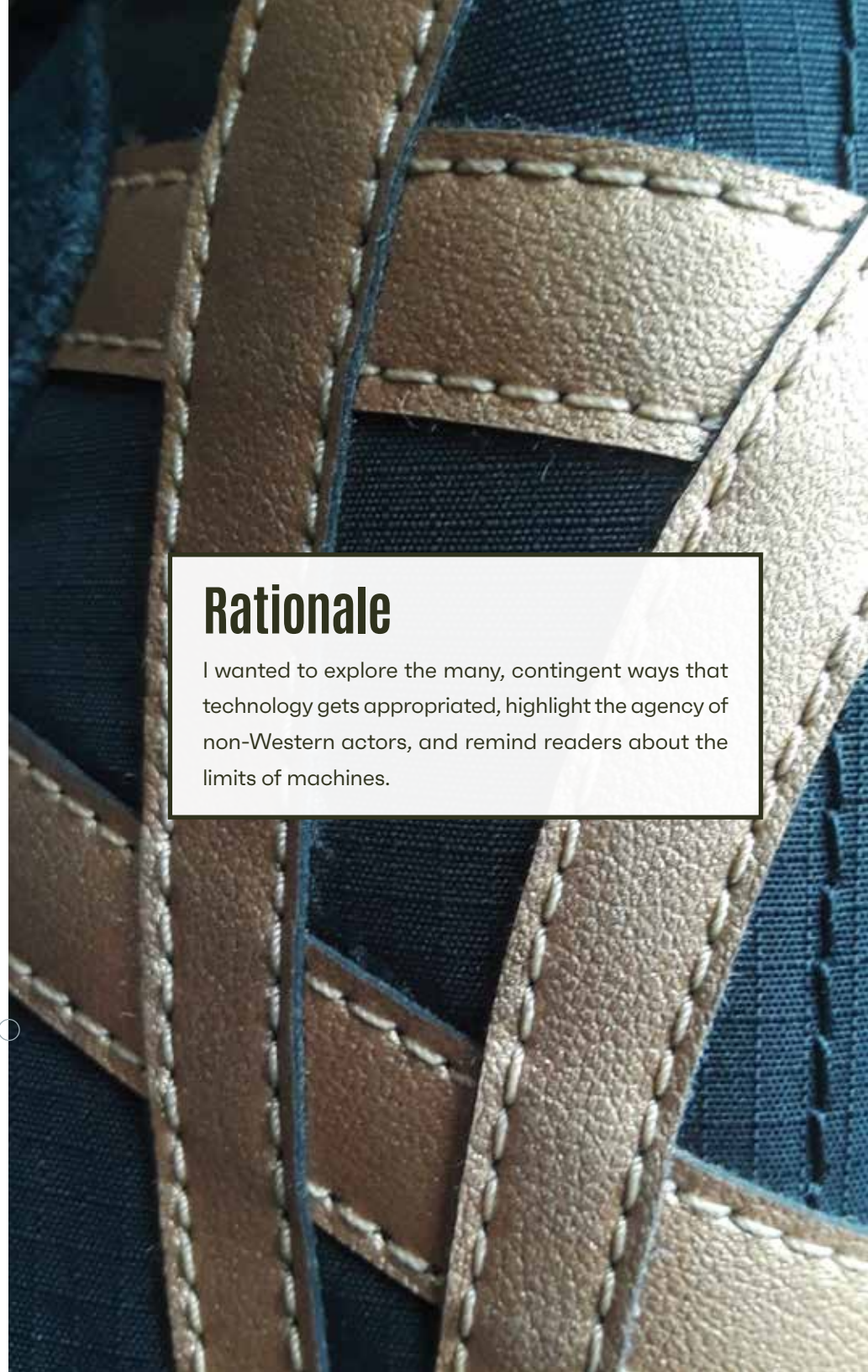
Have I done well? Have I done good? The Box, gorged on knowledge but with no taste for meaning, has the same chance of answering as the rain.

What Joshua cannot see is Uma, village screwdriver in hand, taking The Box apart, marveling at its simple complexity. The Box showed up, belonging to no one. The innards revealed themselves easily. Yet it took Uma years of questions and bus trips and books to uncover the intricacies of the underlying code.

Uma, older still, professor of engineering, marvels at the machine Joshua made. She also marvels at his choices. Imagine, shipping an expensive device to the Sahara with no dust guard. Imagine, a genie without a wish list. Imagine, building a device capable of answering any question except about itself.



Sneakers have a certain cultural specificity. I associate them with 1980s America, although they probably date back at least to the 1920s. This sneaker—the Onitsuka Tiger—is interesting for its origins in postwar Japan.



Rationale

I wanted to explore the many, contingent ways that technology gets appropriated, highlight the agency of non-Western actors, and remind readers about the limits of machines.

In Search of Robot Love

Jeff Cao

Tencent Research Institute
China



IT IS 2030, AND LIFE IS PRETTY EASY IN THE KINGDOM OF Utopia. Self-driving cars roam and fly everywhere, and autonomous robots do everything for humans. In this work-free world, everything seems perfect except one thing: half of all adult men cannot get married. To some extent, this is because a cultural preference for boys over girls has resulted in a great gender imbalance in the population. The government regards this as a real threat to its stability and is eager to solve the problem.

Meanwhile, a young guy named Niulang is under great pressure from his parents to get married soon. This seems to be an extremely hard task to fulfill, especially considering that Niulang has no interest in any particular woman. Fortunately, he is a tech genius and works in a robot company. With the support of his unmarried boss, Niulang secretly creates a humanoid robot which looks and behaves just like a real human. He names her Zhinü and successfully deceives his parents into believing that she is his wife.

Niulang's boss sees an opportunity to make money and begins to produce this kind of robot on a mass scale. The company persuades the government into enacting a law to allow marriage between men and robots, regardless of the opposition from the elderly. Suddenly all bachelors can get a robot wife like Zhinü. This leads to fierce protests from robot wives (who claim that their human husbands never treat them with the respect given to human wives), women (who ask for the right to marry robots), and other kinds of robots (who make claims for robot rights and dignity).

Under these pressures, the government has to formulate new laws to allow for marriage between women and robots, and to recognize robot rights. As for Niulang, his robot wife, Zhinü, has never been accepted by his parents. In their opinion, an infertile robot can never become a wife.

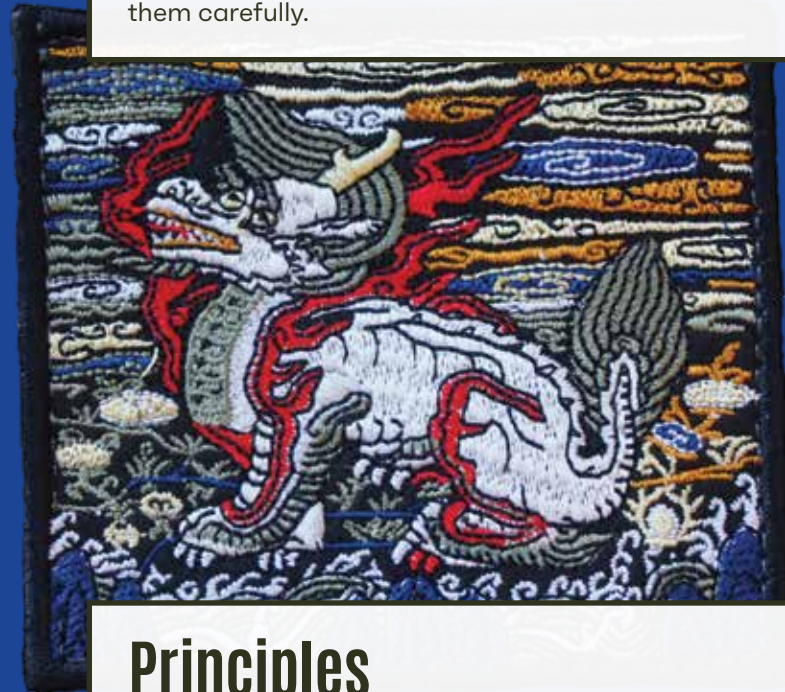
Years after the adoption of robot marriage, a crisis occurs.

Because of the infertility problem and the increasing disconnection between men and women, the human population decreases sharply. Now the real threat to the government arrives—a threat so huge that it may spell the end of the human race in the Kingdom of Utopia.



Rationale

The arrival of all kinds of autonomous and intelligent robots, including humanoid robots, will cause a series of social and ethical problems around the relationship between human and robot. In the coming years, we must think seriously about these problems and handle them carefully.



Principles

- A human-centric or human-oriented AI approach is plausible right now.
- AI shall respect human dignity, rights and freedoms, and cultural diversity.
- Use AI to enhance human creativity, not to do people harm.
- The adoption of AI and robots in human society shall aim to achieve a harmonious human-machine relationship.

The Boy Whose Light Went Out

Jack Clark
OpenAI
United States

THE FATHER HACKED HIS SON ONE YEAR AFTER THE DIAGNOSIS. This was six months after the doctors told him that the treatment would cost \$1 million, and one week after he had given up trying to raise the money, having run out of options and favors and friends.

“Son,” the Father said. “I don’t believe them. I need to see what’s going on.”

“OK, Dad. I would like to see what is going on. I have to see what is going on. I need to see what is going on. I really need . . .” The Father calmed his son by tightly hugging him. He rubbed the back of his Son’s skull and felt the nub of the brain-implant—a portal into an increasingly damaged brain.

The Father was a famous scientist and, though not an expert in neuroscience, he was adept at training machines to solve problems. So, he bought powerful computers and put them in the garage, along with a massive television and a chair. Then he called in favors with friends in medical research and obtained some software to let him read data from his Son’s implant and send back signals. Then it was ready.

As his Son had become more sick, he had developed a fascination with construction sites, finding something soothing

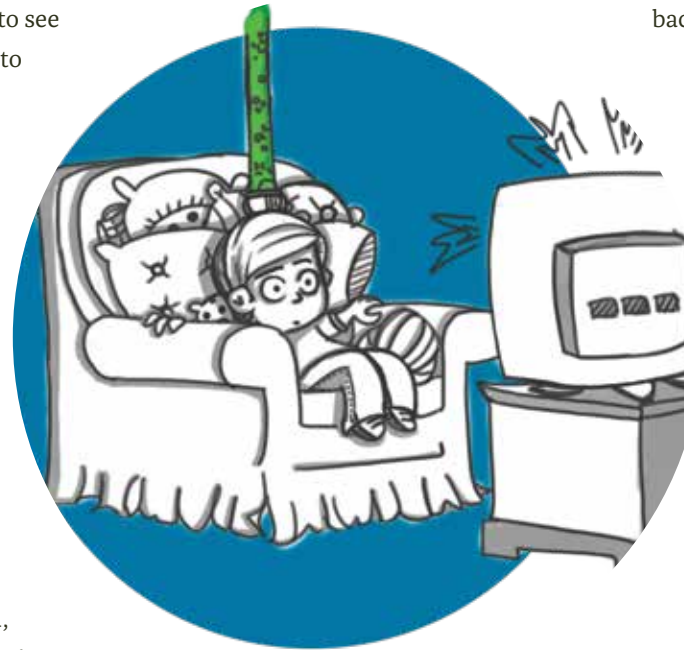
about the slow ballet of the bright yellow vehicles. So, the Father put a webcam feed of the construction site on the television in the garage and turned up the volume. The Son heard and made his way to the garage, then walked in, eyes glued to the television. He sat down in the chair, rocking back and forth, looking at the machines. The Father piled pillows onto his Son, and something about their pressure quieted him.

The Father gently plugged the wire into the back of his Son’s head. Then he stared at the computer screen as data flooded in, watching the software draw a 3D image of his Son’s brain, before stippling it with red dots and accompanying text like **NEURO-ATYPICAL** and **OUT-OF-DISTRIBUTION FIRING PATTERN**.

So, the Father tried to see if he could do something more than monitor his Son: he wondered if he could train an AI system to figure out the firing patterns necessary to restore his Son to normal. The Father began to try to simulate his Son’s brain, trying out firing patterns he could impose upon it to make the red boxes disappear, to make **NEURO-ATYPICAL** change to **NEURO-TYPICAL**.

One month later it was ready. The AI predicted with 99% confidence that his Son could be healed, if he beamed the firing pattern through the cable and into his brain. And so, the Father stood behind his Son and thought about what it would mean to **INITIATE TEST FIRING**.

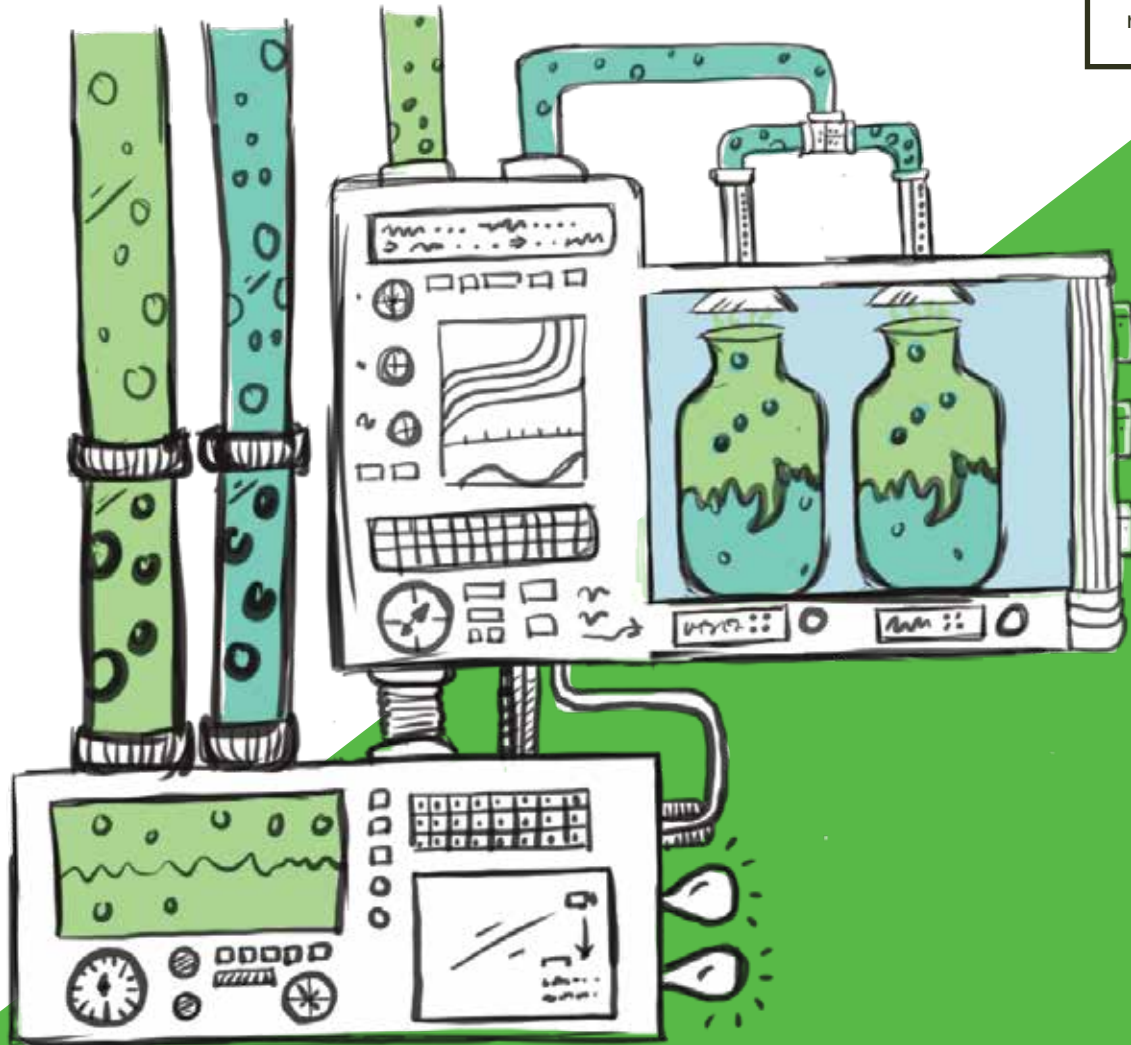
“I like the machines, they work so hard, they do not stop,”



said the Son. “They know what they are doing, they always know what they are doing, they are doing one job and they will never stop.”

“I know,” said the Father. “I like them too.”

And then he pressed the button.



Rationale

I think AI will always involve a trade-off between capability and explainability, and it seems like in the future humans are going to need to perform “leaps of faith” to get learned systems to solve really hard problems.

What Justice

Batya Friedman

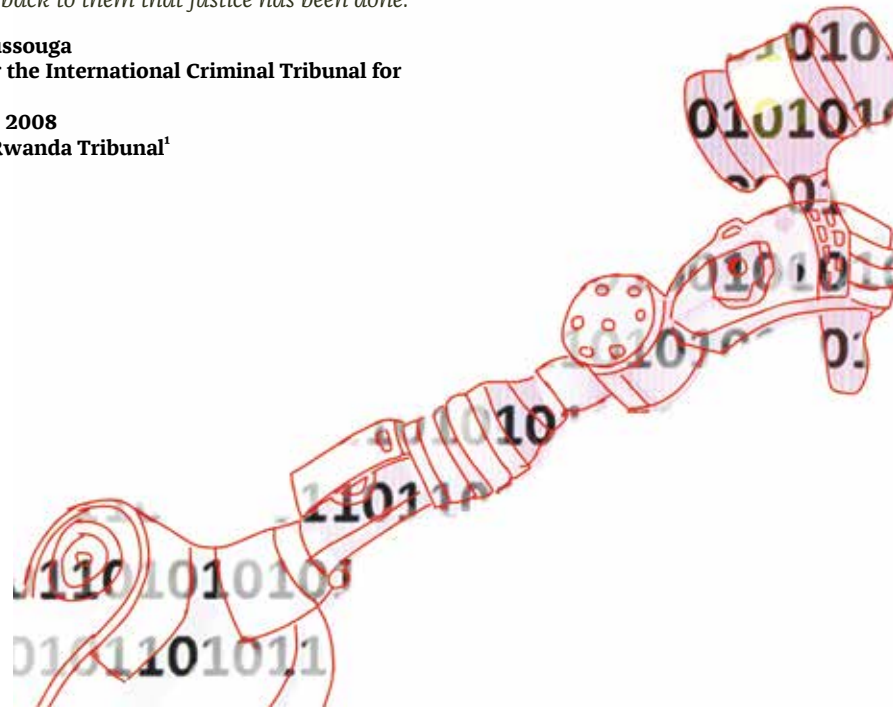
University of Washington
United States

AN OLD LADY, 85 YEARS OLD LADY WHOSE KIDS WERE ALL KILLED, HUSBAND killed, and she was raped. And I brought her in court here to testify in the first case. And this lady . . . when she entered the courtroom, we prepare her, when she entered the courtroom, she was smiling. And then when they asked her, “Witness, could you identify the accused person?” . . . And when she came . . . to see the accused person, she bowed to the accused person and she went back and sit. . . . The judges say, “Can you point the finger?” Say “In my culture you don’t point the finger to powerful people.” Said, “No, he was the mayor.” And the mayor was the most powerful. And the court agreed to – agreed that the lady has recognized the accused person on the basis of that sign. And then when we went home, I said, “Mama, how do you feel?” “I’m so happy. I could not believe that I’ll have this day in my life to see the son of God to be there with handcuff. No, it’s not possible. I can die today and go and see my kids and report back to them that justice has been done.”

Mr. Roland Amoussouga
Spokesperson for the International Criminal Tribunal for
Rwanda (ICTR)
Arusha, Tanzania 2008
Voices from the Rwanda Tribunal¹

Our story above is based in fact: in the trial of Akayesu, the panel of judges at the International Criminal Tribunal for Rwanda recognized the old mama’s dilemma—to witness and also to respect powerful people—and in their wisdom accepted her nuanced gesture in lieu of their instruction to point a finger. In this detail, culturally responsive justice. Dignity.

2048: a future tribunal. The personal backgrounds and biases of Black, White, and Brown judges are gone. Instead the witness, an old grandmother, faces a panel of metal and plastic judges—with hidden circuits, diodes, and computer code. The metal judges ask, “Can you point the finger to the accused person?” The grandmother gestures toward the accused. The judges’ AI gesture recognition system interprets that the witness had pointed to the accused. Judgment, guilty.



¹ See <http://tribunalvoices.org/voices/video/621>

2048: the day after. Headlines in The Hague read “Efficient AI Judges Make Good on International Justice!” Those in the Silicon Valley, “In International Courtroom, AI Gesture Recognition Systems Deliver Culturally Responsive Justice.” The old grandmother returns to her town. Neighbors ask, “How do you feel now?” “Empty,” she says quietly. “Empty.” “I know he was found guilty. I know he will be punished. But what do these metal judges know of the pain I suffer? How does this cold metal help me to heal?” The neighbors nod their heads, agreeing. They, too, had family members killed. One says, “Tell me, where in the courtroom was there a human being who looked him in the eye and said, ‘You are guilty, Y-O-U murdered.’ How can we heal in this community if no person has said this to him? Where is our humanity—in this metal courtroom?” “Indeed,” she sighs and shudders. “Indeed.”



These Levi jeans circa 1975 point to my cultural and intellectual roots when I came of age in the United States. They continue to mark generations and communities, in new ways.

Reflections

Justice is more than a right decision. It is a process of human beings witnessing for each other, recognizing each other, accounting for each other, restoring each other. Nothing less than compassion underlies the conditions in which we can restore ourselves—heal and move forward in life.

True Love?

Sue Glueck
Microsoft
United States



True Love?

HE WAS NEVER SURE OF THE EXACT MOMENT WHEN HE FELL in love with Robotin. Of course he was excited when he took her parts out of the kit, but surely that wasn't love. And over the days and months that followed, as he scrimped and saved for custom mods—even designing and 3D printing some himself—well, that was more obsession than love. It was as she started to learn new human-like behaviors (beyond the initial set in the kit) when he noticed a feeling akin to how he'd felt about his childhood cat. Then one day, he realized he couldn't live without her. The adorable little mistakes she made, like putting butter on her salad instead of dressing, only made him love her more.

It was the happiest time of his life. He hadn't had a close friend since he was eleven, when his next-door neighbors moved away. At long last, loneliness was banished. Until that fateful day when Robotin handed him a note from her manufacturer, translated from the original Russian. They were distributing a mandatory software upgrade, otherwise the robots could be easily hacked and programmed to harm their masters. But the upgrade wasn't cheap—almost half the price

of the original kit, which had taken him years to save up. It wasn't easy these days in a small, dying Ohio town to find enough work to cover the basics, never mind a \$1,000 upgrade. Most of the jobs had gone to the robots. Without the upgrade, she would deactivate in one month's time, at midnight on June 30th, losing all of the knowledge and memories she had acquired during her time with him. He alone would remember the first time they went bowling where she refused to hit the pins, saying it would be cruel. All of these precious memories lost, all the things that made Robotin uniquely her, a part of him.

He worked feverishly, borrowed from everyone he knew, hocked practically everything he owned, but it wasn't enough, not nearly enough. Holding the \$159.25 he had managed to save, he desperately attempted to convince the company to hold off or to let him pay in installments—to no avail. "Don't you understand? I love her, I need her!" he yelled angrily in the phone at the uncaring customer service bot that kept chanting, "But you'll get a partial refund, sir." As if a refund could heal a broken heart.

As midnight on June 30th approached, they sat together. "I love you, Robotin."

"And I you" were the last words she ever uttered.



Reflections

This story was inspired by a hypothetical in an article written by Professor Woodrow Hartzog.¹ He starts by discussing our all-too-human tendency to anthropomorphize robots, as evidenced by Kate Darling’s research over the years, in particular her description of the US military’s testing of a robot that looked like a six-legged stick insect that defused landmines by stepping on them and losing a leg with each defusion.² As Darling notes, the officer in command canceled the testing because “according to Garreau (2007), [t]he colonel just could not stand the pathos of watching the burned, scarred and crippled machine drag itself forward on its last leg. This test, he charged, was inhumane.”³

Hartzog goes on to discuss his family’s own vacuum robot that they have named Rocco. In Hartzog’s hypothetical, Rocco serves the family faithfully for years and tells jokes about how its job “sucks.” He imagines a day when Rocco starts sputtering and, seeming unwell, says, “Daddy . . . [cough] . . . if you don’t buy me a new software upgrade . . . I’ll die.”⁴

Manufacturers of robots and other objects we tend to anthropomorphize should keep in mind that normal business decisions around updating, upgrading, and stopping support could have a more significant impact on end users than those for can openers. Understanding customers’ needs and desires—and making principled decisions with them in mind—is simply good business and an excellent way to generate customer loyalty.

Rationale

Humans need to love and be loved to survive. We love our families, our pets, and our communities. Sometimes we also love the objects we have anthropomorphized. Please be cautious with our tender hearts.

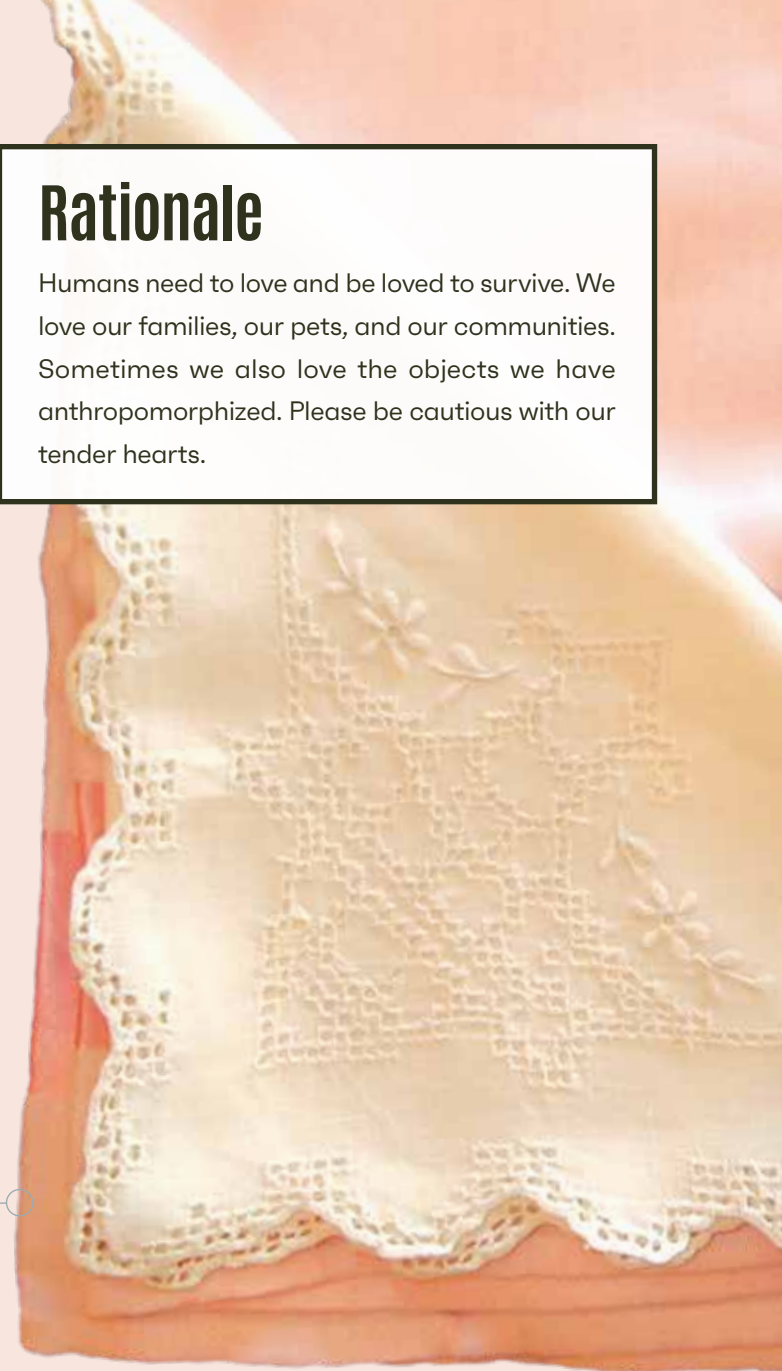
The fabrics are table linens that I inherited from my mother and grandmother, both of whom loved to entertain. Although they are no longer with us, their love for me and others shines through when I see these fabrics.

1 See Woodrow Hartzog, “Unfair and Deceptive Robots,” *Maryland Law Review* 74 (2015): 785.

2 Kate Darling, “Who’s Johnny? Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy,” in *ROBOT ETHICS 2.0*, edited by P. Lin, G. Bekey, K. Abney, and R. Jenkins (New York: Oxford University Press, 2017). “Research shows that humans tend to anthropomorphize robotic technology, treating it as though it is alive, even if we know better.” See Kate Darling, “Extending Legal Rights to Social Robots,” paper presented at We Robot: The Inaugural Conference on Legal and Policy Issues Relating to Robotics, 2012, http://robots.law.miami.edu/wp-content/uploads/2012/04/Darling_Extending-Legal-Rights-to-Social-Robots-v2.pdf.

3 Darling, “Extending Legal Rights to Social Robots,” at 5–6.

4 Hartzog, “Unfair and Deceptive Robots,” at 804.



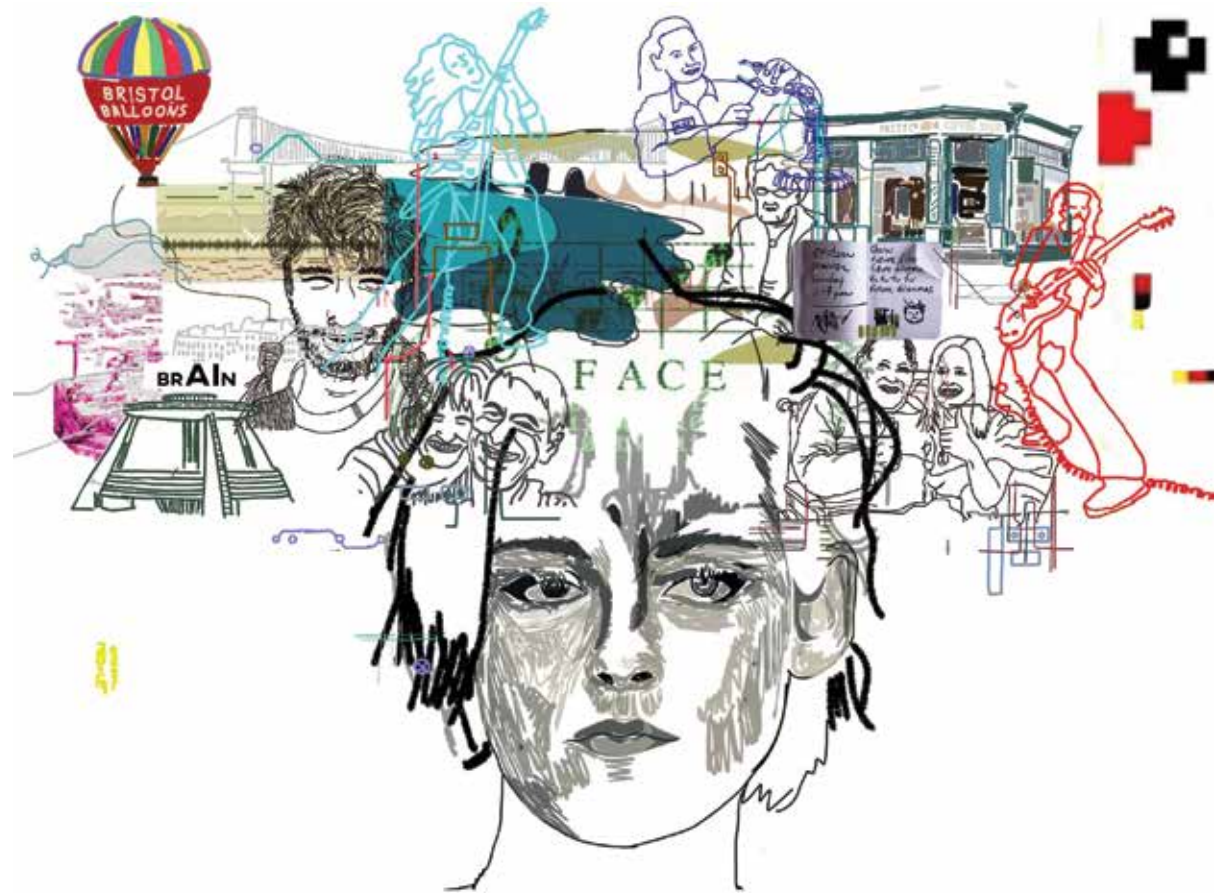
Lia's AI Future 2036

NOW THAT I'M EIGHTEEN, MY MUM WANTS ME TO FIGURE OUT what to do with my life. We're pretty well off: I live in a house where a swarm of household robots take care of everything from picking my optimal wake-up time and proposing clothes for the day to tidying up. Ever since the CEO of deep learning company "BrAIIn" became PM, we have access to free higher education through the eFuture portal, an augmented reality system that beams students into interactive virtual

classrooms. I can choose to be a technologist, which my mum wants because they make the most money; a technology user; being a surgeon sounds interesting; or a humanist, like the caregiver that watches over my grandparents. Whatever the career, I know I want to move to Cornwall, they have the most beautiful beaches, coffee shops, and alternative art scene. Since most services are provided remotely, it doesn't really matter where I live, and Cornwall is close enough to Bristol that I can

Sabine Hauert

University of Bristol
England



go home by autonomous car. The eFuture algorithm thinks I should be an Artificial Intelligist because I did well in math at eSchool, and I'm a woman—there still aren't enough of us in the field. These systems know so much about me, maybe I should just listen? eFuture thought my friend Mae should be a robot-minder and she just loves her job. Or I could rebel and become a musician, I can't play an instrument but am pretty creative and have a musical ear. The other day, Jay and I wrote a song called "future dilemma" just by describing our feelings to Siri. It was pretty good, maybe we can find a band to play it at the local bar tonight through the Giig app? Anyhow, it doesn't really matter what I choose to do, things move so fast that I can retrain in a couple years if I'm unhappy. I can't believe our parents stayed in the same career their whole lives and worked so much—that must have been awful. I think that's why they are so stressed about me sorting out my life. Sam, I know you're just an algorithm, but what do you think I should do?

I chose a baby muslin to emphasize that our decisions today regarding AI will impact future generations. And even though we all start with a similar piece of muslin, we may be impacted differently. It is our responsibility to ensure everyone flourishes through the development of AI.

Rationale

When my daughter turns eighteen, her everyday life and career prospects in an AI-powered society may look different from today, but the trepidations of a teenager will remain.

Principles

- Technology is empowering if implemented correctly.
- The future will look much like today but with technology embedded in everyday activities.
- Education needs to change to allow for rapid retraining of the workforce.
- There is a need for a diverse workforce in AI.
- The jobs of tomorrow are different from the jobs of today.
- Rather than hinder human contact, technology may allow us to spend more time in meaningful human interactions.
- Regions that are currently deprived may become more attractive as value systems change.
- People will trust technology with important decisions—but should still have control.

A Well-Meaning Robot

Alejandro Hevia

University of Chile
Chile

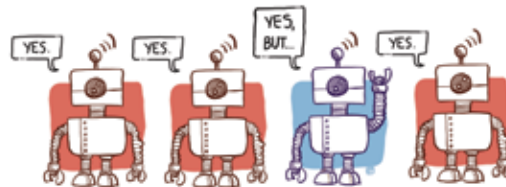
AFTER A LONG CAREER AS AN AI SCIENTIST AND ENTREPRENEUR, Jorge decided to retire and went back to his hometown, Ollagüe, a small village in northern Chile. His childhood friends and neighbors were happy to welcome him back. They worried that the town was dying because of a persistent drought. To solve the problem, Jorge decided to create several robots that could bring water by bucket from a hard-to-reach river.

Although the townspeople were very thankful, some of them were wary of the robots, fearing they would plot something nefarious if left to their own devices. To appease people's fears, Jorge instructed each robot to report every single action they do at any moment, so everyone could know exactly what each robot was doing at all times. Jorge was pleased when he pinged the river robots and saw how much water each carried.

Surprisingly, one of the robots, Benito, complained to Jorge: "Why do you need to monitor me at all times? If you cannot trust me, you could instead verify what I've done at the end."

Jorge asked, "But how would I know you are following my orders?"

"Well, this is what I'll do," Benito responded. "My work, my



goal will be depicted in a puzzle, from which I'll send your phone a piece every minute. You just verify the pieces as they come. The goal, when it's done, will surprise you."

A reluctant Jorge agreed: "So, trust but verify? Alright." Even without knowing Benito's goal, Jorge could make sure Benito would be honoring his word by simply verifying that each single piece would properly fit in the puzzle, every minute. *I can always stop him once his goal becomes clear*, Jorge concluded.

Before going away for a short trip, Jorge told the townspeople about the arrangement, which they grudgingly accepted. As expected, once all the robots began working, water started arriving in the town, which made the townspeople happy. Some, however, became increasingly worried about Benito, who was the only robot not bringing water. Instead, Benito had locked himself inside of a barn, making very loud, frightening construction noises.

Several days passed. With Jorge away, some townspeople decided that Benito had gone rogue, perhaps building a bunker with an army to attack them all. At night, locals arrived at the barn and called Benito to come out. Then they destroyed the robot.



The next morning, Jorge arrived back in town, after receiving the last piece of the puzzle from Benito. He was horrified to discover what had happened. With the townspeople, he went into the barn, where they found a beautiful, solar-powered well gushing fresh water—exactly what Jorge's phone had showed.

Reflections

This story shows a case where the hero, a robot, works on behalf of us humans, aiming for something complex and ultimately good for us. Unfortunately, we humans are unable to see the goal or understand the process toward achieving it. Even if his efforts are positive, once trust in AI is gone, so seems its utility for us. If AI cannot convince us all that a given complex technology is good for us, does this mean that humans should be suspicious of this technology, effectively renouncing any potential benefit? Should we trust the experts, those humans who can understand AI instead? Should AI learn how to show those human traits that, when present in humans, we interpret as trustworthy?

Principles:

- As with all technology, AI's algorithms will eventually become so complex that we may not be able to understand them.
- Human trust in any technology seems critical for long-term development and successful deployment.
- We should aim for transparent algorithms that well-informed humans are able to understand.

Rationale

How will future AI convey complex information to us humans?

My chosen fabric was a handkerchief used by cueca dancers. Cueca, the national dance of Chile, is performed between a man and a woman. In a series of interlocked steps, the male dancer must impress and attract the female dancer using strong and quick dance moves, while the female dancer flirts and seduces him. For both roles, the handkerchief is a critical element. It's a communication device: males use it to show their desire to attract their partners; females use it to show themselves playful and feminine. But it must be used carefully; if used appropriately, it amplifies the effect. If used poorly (or dropped) it can ruin the dance.

I selected this handkerchief as it shows that having well-defined, well-meaning goals in technology is not enough. We need to communicate them effectively, we need to understand our partners' moves and subtle signals in order to reach a successful ending.



With the Lights Out, It's Less Dangerous

Ian Kerr

University of Ottawa
Canada

“HERE WE ARE NOW, ENTERTAIN US!” PROCLAIMED THE Naval Commander, making his usual grand entrance for the monthly procurement session. Kurt could barely contain his self-congratulatory smile. Rumors about the Commander’s legendary entry were true.

“With the lights out, it’s less dangerous.” Kurt began his bid accordingly. “Technology has amplified the speed of war. The dim light of human intuition travels too slowly to guide us in real-time life-and-death decision making.” Kurt knocked it out of the park, explaining the dangers of leaving the kill switch in human hands, pitching the means by which his AI could further automate automatic weapons.

During his motorcycle ride back to the lab, Kurt basked in the glow of his own ideas. Robots do not seek revenge. They are not subject to fatigue. They are not confused by the fog of war. Much to his chagrin, Kurt’s reprise did not go over quite as well with his colleague Tobi.

“My paragon of inventors,” she replied, channeling Plato, “the discoverer of an art is not the best judge of the good or harm which will accrue to those who practice it.”

“What is that supposed to mean?” asked Kurt.

Tobi told Kurt a story about her MIT supervisor. In the 1970s

colleagues were discussing Golem, the Israeli mainframe computer. Quite by coincidence, Tobi’s supervisor and two world-renowned colleagues discovered a common connection leading back to Rabbi Judah Loew, the Maharal of Prague. Each of the three was part of a fourth generation who had been told the story of the Golem of Prague. Each had memorized the same deactivation code, passed down from the Maharal through the generations as a communal safeguard. Precisely

this had led each to study AI.

Like Adam, the clay golem was animated by an act of devout creation. The Maharal piously inscribed one of G-d’s seventy-two names along with the Hebraic code on its forehead: תתקן —the word for “truth.” The powerful combination of these sacred words

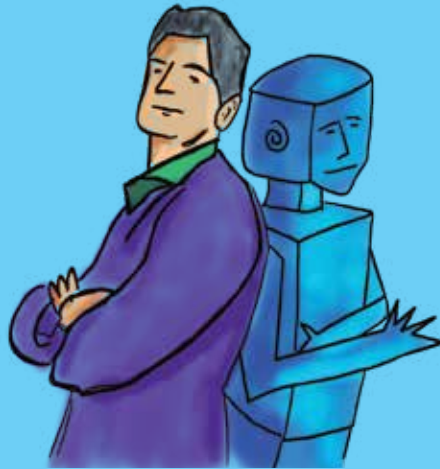


vitalized the golem, programming its narrow mission to protect the Jewish community against vicious attack. Eventually the Rabbi felt duty bound to deactivate the golem after it spiraled out of control, dangerously transcending its original purpose. Despite the tremendous advantage the golem would provide, Rabbi Loew saw wisdom in limiting its use. Erasing the letter ק , the inscription then read תת —the Hebrew word for “dead.”

“Are you telling me we need to retain human control even if it means giving up total military advantage?” asked Kurt.

“I am telling you,” said Tobi, borrowing words from Carl Sagan, “that we are creating world-altering contrivances and we have choices to make. We can relinquish control and roll the bones in a strange game of digital Russian roulette. Or we can rely on the bright light of human wisdom to place limits on what may and what must not be done, and safely pass through times of peril.”

Measuring Tobi’s words, Kurt looked dumbfounded and inert—reminiscent of the golem, still hidden away in the attic of what was once the old synagogue in Prague.



This We Robot tee-shirt represents the subject matter of this book and wears on its sleeve associations with two of my favorite people, two of my favorite institutions and, perhaps, a whiff of Teen Spirit

References:

Isaac Bashevis Singer. *The Golem*. New York: Harper Collins, 1982.

Byron Sherwin. *Golems Among Us: How a Jewish Legend Can Help Us Navigate the Biotech Century*. Chicago: Ivan R. Dee, 2004.

Rationale

In a world obsessed with violent, intelligent machines that rise up against hubristic inventors, this counter-narrative explores the philosophy of automating battle space, the recursive nature of creation, and the wisdom in setting limits.



Lost Knowledge? Forbidden Knowledge?

Tadayoshi Kohno

University of Washington
United States



SEAN, SEATTLE, WASHINGTON, USA. “WHEN I GROW UP, I’M going to make a Cyber Karate Dojo—an online school for people to learn karate. Students will upload videos of themselves, and my computers will use AI to analyze those videos and tell them what they need to work on. For sparring practice, students can rent a robot for a small monthly fee.”

— Thirty years later —

Press Release (Mountain View, California, USA): Cyber Karate Dojo Awards Its One-Millionth Black Belt. Globally, 99.9% of Karate Practitioners Train Exclusively with Cyber Karate Dojo.

— Ten years later —

Shihan Tanaka, Tokyo, Japan. “I am heartbroken, standing here in the same dojo where my parents taught me. In my sixty years of karate training, I have learned many things. I have vast advanced knowledge that I am willing to share with dedicated students who could prove their worthiness—knowledge with a deep history and connection to my culture.

Unfortunately, no new students are coming to me, and my knowledge is too nuanced, personal, and dangerous to share with any online system. This knowledge—this connection to my culture—will die with me.”

Shihan Smith, Bethesda, Maryland, USA. “I made a mistake. I wanted to honor my Japanese instructors by preserving their knowledge—the secrets from their culture—that they taught



me. So I shared everything I learned—every technique, every detail—with the creators of Cyber Karate Dojo. Some of those techniques are quite dangerous if not practiced with the maturity and discernment of an advanced student. The makers of Cyber Karate Dojo told me that they would be careful. But their company was sold, the new owners are not so cautious, and now anyone can learn these advanced techniques. People are getting hurt, and it's my fault.”

Shihan Yamada, Vancouver, British Columbia, Canada. “My teachers sent me from Japan to North America to share my knowledge of karate. It has always been challenging to find students. Students don't want to spend the time to learn traditional karate. Helping Cyber Karate Dojo develop their platform was the best thing I ever did. I know that some elements—some special flavors—of the individual styles have been forgotten. But without Cyber Karate Dojo, much of the art of karate would be lost to this generation.”



Rationale

How do we foresee all the possible harms that might manifest over time, and how do we appropriately balance the risks with the benefits?

Reflections

Thank you to Sean for dreaming up the idea of the Cyber Karate Dojo while on a hike with his dad. He has already started to create the Cyber Karate Dojo, blending tradition with technology.

This story explores different possible negative outcomes that could happen when we try to preserve instructional knowledge with AI systems, especially when what is being taught has layers of richness. Some information may be lost forever (Shihan Tanaka's situation), or some information could be used in inappropriate ways (Shihan Smith's situation). The story also speaks to the challenges with the evolving use of information over time, where initial

uses might be safe but future uses might be unsafe (Shihh Smith's situation). Shihan Yamada's situation speaks to the fact that a single technology could be seen as harmful from one perspective but beneficial from another.



Language Matters

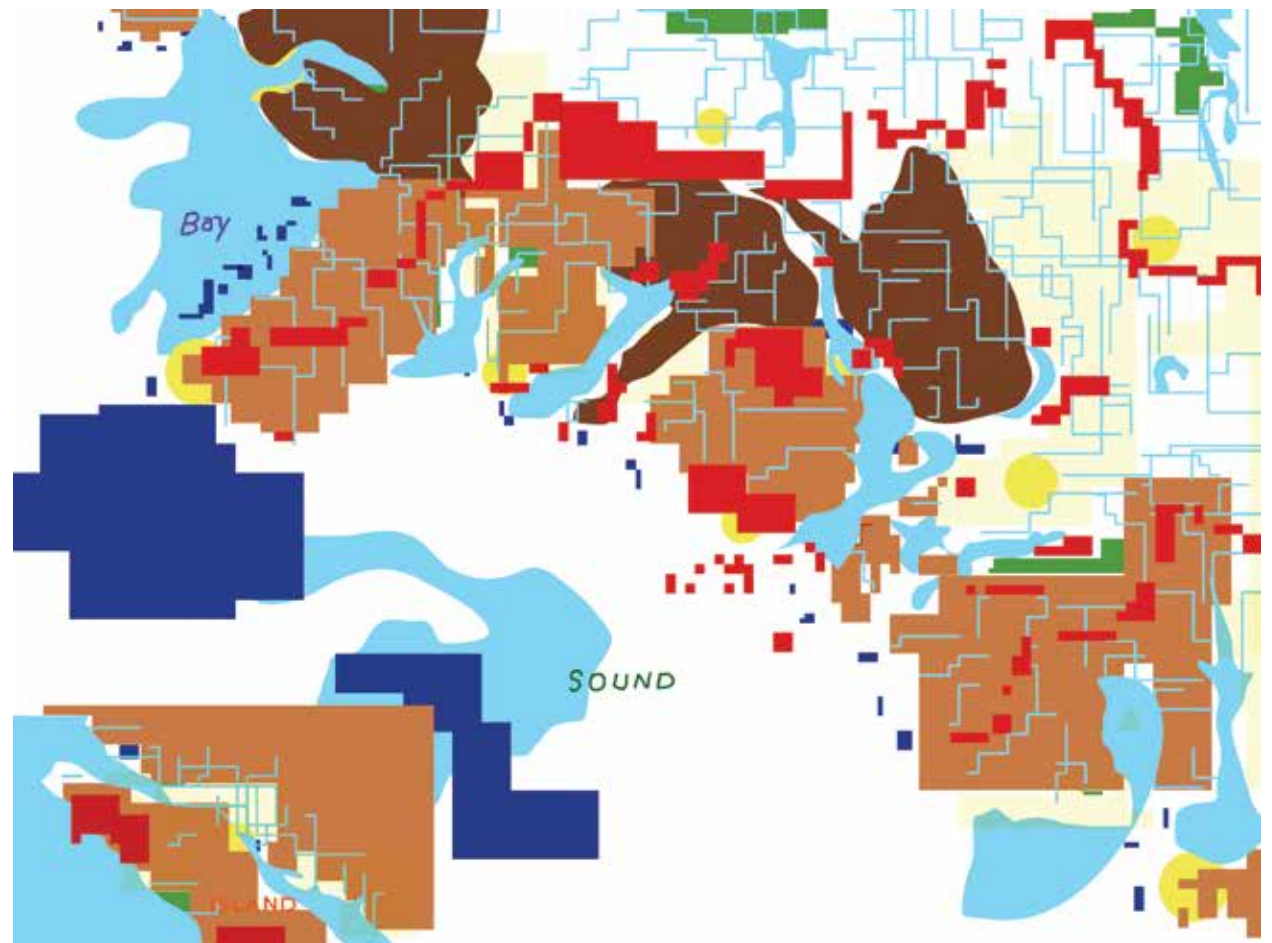
THE VIEW OUT KATE'S OFFICE WINDOW—WATER, MOUNTAINS, and trees—transfixed her. As a newcomer to western Canada, she was unaccustomed to so much natural beauty alongside stark city skyscrapers. The view is a perk of working for a machine learning lab in downtown Vancouver. It reminded Kate of the protest she attended over the weekend. A friend had challenged her to show up, not just “like” it online. The pipeline protest was led by First Nations communities and organizations. Kate hadn't expected all the stories woven into

the protesters' speeches. Stories grounded by thousands of years of stewarding the water, mountains, and trees visible from Kate's window. One guy, a lawyer, gave examples of how legal battles, protests, imprisonment, and other forms of resistance were parts of the strategy Indigenous peoples have used for generations to protect precious fragments of their culture, land, and language.

One woman's words kept returning to Kate: “Culture matters. Land matters. Language matters.” The woman had spoken

Lisa Nathan

University of British Columbia
Canada





about how culture, land, and language are inseparable for First Nations peoples. How do these ideas relate to a computer language and the programs created through that language? Kate thought about her boss at the lab. He talks so convincingly of AI's potential to help humanity. But who—which humans—will benefit from AI? What values and cultural knowledge are embedded in Kate's code? What is valued in Python? What does a programming language expressive of a particular First Nation look like? What does a First Nation's AI program value? How will it disrupt Kate's view?

The gray fabric is from a T-shirt created for the Indigitization Futures Forum held at UBC in 2016. Indigitization is a partnership based out of UBC. The Forum brought together an emerging network of community-based information professionals and practitioners, academics, and a wider community of specialists who work to support context-appropriate, digital information practices within Indigenous communities. The Indigitization logo featured on the T-shirt represents Raven. In the words of Alison O. Marks, the artist who created the logo, "Raven is a shape changer, and like technology, represents the movement between fluid and structural form. Often identified with its capacity for mischief, I also like to imagine Raven appreciating the digital realm as a site for play and creativity."

Rationale

I offer Kate's contemplation of disruption to encourage readers to question approaches to and outcomes of programming languages, the foundation of AI.

Reflections

I am deeply inspired by the work and friendship of Kim Lawson, Heiltsuk scholar and librarian, as well as that of my Indigitization colleagues.



Do We Have the Right Metrics?

THE SOCIETY WHERE I COME FROM IS RWANDA, IN THE EASTERN part of Africa, where we saw mobile technology penetrating the market in 1998. The mobile communications company Telemix introduced mobile phones. People were excited to see that they no longer had to travel to their homes or offices to make calls; now they could carry their mobile phones and access loved ones from wherever they are.

In the first two years Telemix created sixty thousand jobs. The call center employed lots of young people who worked in shifts to meet clients' needs. Whenever a client would have a query, he or she would dial 456 and talk to the company staff, who would address the query or connect them to the person in charge of that specific case. But as the company got more clients, they needed to improve customer service. Over the

next three years the company had to revise its policy to be able to meet the demand but also maintain the quality of services.

The company chose to set up customer care chatbots. These chatbots would respond to the queries of the clients. For example, when clients called, this would be the response: "You've reached the customer service center. If you need to top up your credit, dial 1; if you need assistance with your phone, dial 2; and if you need the operator, dial 3."

To an elite this was great because it was improving the customer service. But to ordinary citizens, who had never been exposed to any sort of technology, it was shocking, particularly since the chatbots could not differentiate between female and male. In Rwandan society, the way one communicates to a male is different from how one communicates or responds to a female. The other challenge was that in our culture, before one mentions what he or she wants, you have to at least greet the person on the call and chat in a leisurely fashion before posing your question. At first the clients didn't know that they were being answered by the chatbots, so they assumed that Telemix's staff was very rude.



Joseph Nkurunziza

Never Again Rwanda
Rwanda

Not surprisingly, the company saw a decrease in clients. The majority were switching to the other mobile company that was still using human capital in the service centers and call centers in responding to clients' queries. Although no AI was involved at that time, the lessons are clear.



Currently the government is championing made-in-Rwanda products. For social functions, we now dress traditionally as a way to maintain the culture while promoting Rwandan products.

Rationale

I am a medical professional and I practiced medicine for quite some time before I got heavily involved in the development sector, advocating for social justice, participatory governance, and peace building. However, I didn't envision myself talking about information technology until I was recently on a panel discussing how artificial intelligence can contribute to global governance issues.

Principles

- Have thoughtful, collective consideration of issues related to governance as an integral aspect of responsible technological development, to ensure that the future we build is one in which we can all participate and prosper.
- Include the real stakeholders—the community.
- Incorporate the local norms and customs—avoid cultural insensitivities!

Moussas of the Savanna

MOUSSA BAMBA, A BIOCHEMIST, IS INVITED TO A SYMPOSIUM in Vermont. He has checked on visa processes for the USA. Thanks to artificial intelligence, they are now fast, reliable, and stress-free. A robot does everything in five minutes.

At Uncle Sam the Robot, Moussa scans his passport and fingerprints, and a picture of him is taken. Uncle Sam normally makes a decision within ninety seconds. His decision is based on existing data, social media, and other online activities of the applicant. His answer is accompanied by a light: green,

yellow, or red. Green is “visa granted.” Yellow is “irregularities noted.” Red is “visa refused.” When Uncle Sam gives you a green light, he prints the visa immediately.

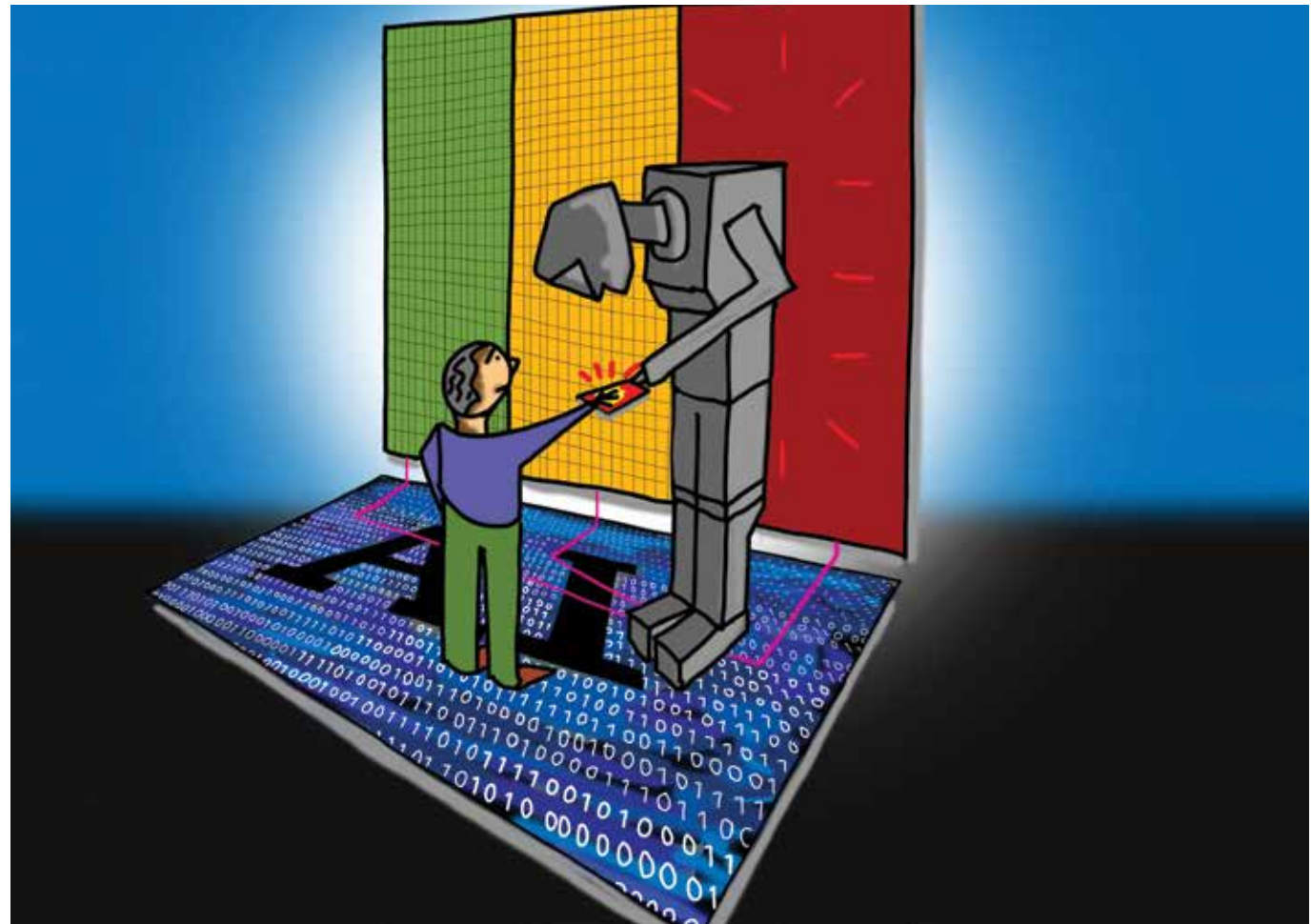
Ta-ta-ta-ta, after a minute Moussa gets a red. The robot prints him a page. The top part says, “visa refused.” But the bottom part is all red: **RESTRICTED. TERROR SUSPECT. SECURITY THREAT.**

Shock!

Two days later, Moussa’s bank calls him. “Doctor, it’s about

Nnenna Nwakanma

World Wide Web Foundation
Côte d’Ivoire



your loan application. The headquarters contacted us and asked us to limit our relationship with you to just the transfer of salary and withdrawals. No deposit, no credit.”

How? Why?

Moussa Bamba was born in Africa’s Savanna belt. Mostly Muslim, etched in traditional beliefs, his tribe believes that twins bring prosperity. When Moussa was born, a cousin arrived three days later. The family so wanted twins that they chose to go to the registry with both babies and declared that they were born on the same day. Each boy was named Moussa. That increases luck. Though the birth registration certificates were different, one Moussa was 328567 and the other was 328568. The Civil Registry Officer should have raised an objection but preferred to receive a “gift” from the Bamba family. One Moussa went on to do his studies and became a biochemist. He had no social media presence. The other Moussa became an “activist,” code-named Massacre Master on social media, with posts highly seasoned with bigotry, racial slurs, and religious extremism.

After checking birth date, location, ID number, and facial features, Uncle Sam the Robot decided that with 89% match Dr. Bamba was the Massacre Master. The robot determined that a radical Muslim biochemist, visiting a laboratory in rural Vermont, had sufficiently ticked the boxes to be placed on “red.” He therefore rejected the visa request, placed Moussa on “red,” and updated all relevant databases.

Dr. Moussa wants a redress.

Rationale

Identities are us: the cultural us, the social us, the technology us.

Reflections

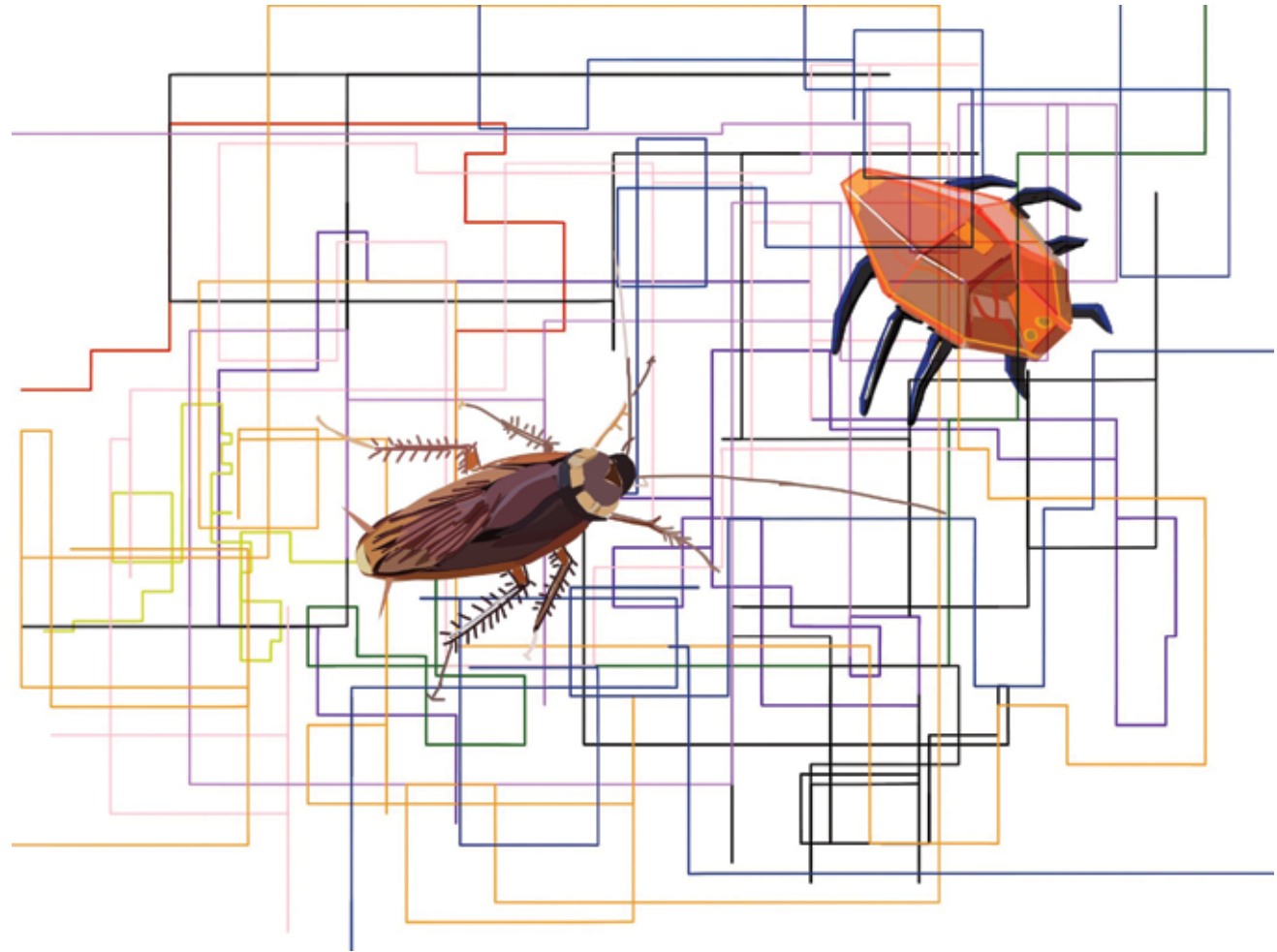
- Can we reliably run AI-based services without identities, digital and biometric?
- What is the reliability of data sources on which we are building AI?
- How will AI affect people who are connected and active online as well as those who are offline and unconnected?
- The human element can transfer affected corrupt and biased influences to AI. How do we hack human minds?

Bug

SARAH AND HER ONE-YEAR-OLD DAUGHTER, EMMA, ARE ON their way to Iran. She has been living in London for the past ten years and is looking forward to finally going home. She studied there and got married there; now, after a decade in the UK, Sarah is happy that she will finally be able to see her family and friends. All these years, she was in touch with them via

email, messaging apps, and video chat but not in person—until now.

When her airplane lands at Tehran's Imam Khomeini International Airport, Sarah is ecstatic with anticipation about seeing her parents, who are waiting for her on the other side of passport control. While in line, she tries to spot her family



Amir Rashidi

Center for Human Rights in
Iran
Iran



while holding her baby close to her chest. She catches a glimpse of her mom and can almost feel her hugs as all those good memories come rushing back.

Officer: “Ma’am? Ma’am? I need to check your passport. Ma’am? Do you hear me?”

Sarah: “Oh yes, yes, sorry, I was trying to find my mom on the other side. Here is my passport.”

Officer: “Can you please answer into this mic? How long have you been in London?”

Sarah: “Ten years.”

Officer: “Please stand there so that the camera can photograph your face. So, what did you do there?”

Sarah: “Study and work.”

Officer: “You married a foreigner. Are you involved in politics? What are your political views?”

Sarah is surprised. She didn’t expect such questions. “None. I’m not interested in politics.”

Officer: “Please come with me. You need to answer some questions.”

Two officers take Sarah, with Emma still in her arms, to a room with a table and two chairs. It’s dark and there is a Koran on the table. She sits there alone, holding her sleeping baby, for an hour. Sarah is nervous and scared. They took her passport—what will they do next? She hugs her baby closer to her chest. A man opens the door, sits down, and starts reading her file.

Officer: “So, your name is Sarah. Why do you think our government is corrupt?”

Sarah: “Me? I don’t think this government is corrupt.”

Officer: “Ma’am, we have a recording of your voice. It’s better for yourself that you don’t lie now. Do you really think if you used an anonymous account, we would not find you?” The officer turns on an audio device and puts it on the table. Sarah can hear a woman talking about the increasing cost of groceries

and complaining about the role of the Iranian government. Emma is beginning to wake up, maybe because Sarah is holding her too tightly.

Officer: “This is you.”

Sarah: “No, that is not my voice. My voice is definitely not like that. You can hear me now; listen to how different it sounds.”

The officer shows her a photograph of a woman waving a flag with text on it that says, “Equality for Iranian women.”

Officer: “On March 8th, 2040, you were marching in London with this flag and chanting against our national security.”

Sarah: “That’s not me. Her face is covered with a mask. How do you know this is me? I was not there.”

Officer: “We know. Our system never makes a mistake. We have your voice, your picture. This is you. We have your conversation with your colleagues in Iran. It’s better to confess your crimes against our national security. The AI never makes a mistake.”



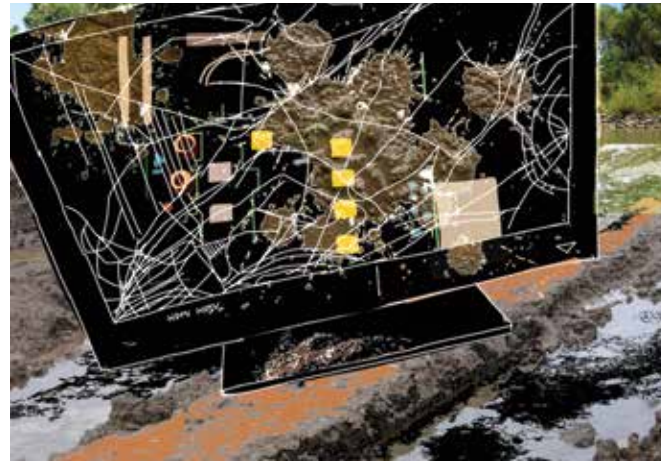
Rationale

Technology is supposed to make human life better, safer, and also improve the quality of our life, but only if the technology is in the right hands. Oppressive countries such as Iran use technology to have more control over its citizens. There must be a way to keep them accountable.

The Goddess

Rohan Samarajiva

LIRNEasia
Sri Lanka



SAMAN HAD HEARD THE WARNINGS. HE'D SEEN PLENTY OF floods but none came this far up. He had to haul some rice to the city early the next morning. No point in taking the Batta to the brother-in-law's place further up. He was tired.

Mother woke him. The little truck was gone. Moonlight showed vehicles floating by. He cursed himself. "Why didn't I move it last night?" He'd just paid the last installment with money Kusum sent from Kuwait. The Batta was finally his. And now it was gone, along with the rice. All they had sacrificed.



The evenings he hadn't gone to the riverbank to drink with his buddies.

How to get mother out? Father's picture? Mother had it. The Goddess? She can look after herself. She's supposed to bring me prosperity. My prosperity floated away. I am poor again.

Saman's phone rang. "Machang, are you okay?" That was the brother-in-law. "I'm ruined," Saman said. "I think I can get out but Amma is the problem. You have that old tube we used to take to the river? Rope?"

By the grace of Goddess Pattini, no one died. The water receded, leaving behind an awful stench. Mud coated every-



thing, as did hopelessness. How will we live when the food packets stop?

Two days after the flood, the insurance rep came by. "Here's some money. Sign this and you'll get replacement value tomorrow. Buy another Batta. Government is hiring vehicles," he said. "Boss was on TV, promising payments in full within a week. Easy for him to say."

"What TV?" asked Saman. "Here I'm cleaning out what's left of my life and you're asking me about TV. I have to find the Batta. Where's the assessor?"

"Don't need assessors," the rep said. "Haven't seen one in

ages. The Batta is finished. We know where it is. Down near the old bridge. I told you installing sensors was a good idea. It cost more at the start, but your payments went down and now we can pay out quickly.”

Payment authorization came to my phone. It’s like magic.

“Remember when these things took months? You gave me so much grief, saying you didn’t want us watching how you drive and how you hated premiums that went up and down. No one was watching. Just some crazy computer spewing out stuff.”

“What do you think now?”



This was a shirt made in Sri Lanka by artisans that was gifted to me in 1999 when I ended my first stint in government to return to the university.

Rationale

In this near-future story, I show how I think AI will touch the lives of my people, as narrow AI becomes embedded in everyday processes, products, and services. It merges my long-standing interest in disasters and the use of information in managing risk. I could easily write a dystopian ending for the story, but this one seems more useful.

Principles

- AI will not make life better by itself but will be effective when coupled with “analog complements” such as insurance.
- Insurance is what helps us manage risk; it is most dependent on information. Ubiquitous sensors and narrow AI capable of analyzing the deluge of data and making decisions is a natural for insurance.

Of Genies and Magic Lamps

Jeroen van den Hoven

Delft University of Technology
Netherlands

JINN IN ARABIC MYTHOLOGY AND THEOLOGY REFERS TO SPIRITS that can be summoned to assist humans. The Arabic noun means “to conceal,” so “genies” are literally “beings with super-human powers that are hidden.”

Will the genies of our AI-powered wearables and assistants of tomorrow help us solve the world’s problems peacefully, or will they set us up against each other? Could they turn against us? What are we allowed to ask of them? Are they permitted

to deceive humans? Who will own the more powerful ones?

As the intelligence hidden in our wearables becomes more powerful, they could become friendly helpers to implement our life decisions, expand our capabilities to make smart choices, and negotiate better deals. They can represent our preference profiles, measure toxins in our food and biomarkers in our bodies. They could eventually help us to become better people.



The story is old.

Aladdin is asked by a sorcerer to retrieve a magic lamp from a dangerous cave in which he gets trapped. Fortunately, when he accidentally rubs a ring the sorcerer has lent him, a service “genie” appears who releases him from the cave with the lamp. When Aladdin’s mother cleans the lamp, a second far more powerful service-oriented genie appears. Like the genie of the ring, it assists anyone holding the object.

With the aid of the genie of the smart lamp, Aladdin becomes rich and marries the sultan’s daughter. The genie builds Aladdin a wonderful palace. The sorcerer, hearing of Aladdin’s good fortune, returns and manages to get hold of the lamp again. He orders the genie of the lamp to take Aladdin’s palace, with the riches, to his own home. Fortunately, Aladdin still has the magic ring and is able to summon the lesser genie. The less powerful genie cannot directly undo the magic of the genie of the lamp, but he can get Aladdin at least to the place where he can recover the lamp. Which he does, and all is well.

The genies in this old story are perfect and seamlessly work with those who own the artifacts in which they reside. They serve a good boy and an evil sorcerer with equal dedication. Will our modern genies be as perfect and as morally indiscriminate? Will they be reliable and safe, will they be able to keep a secret and respect our privacy? Will governments have to guarantee AI genie assistance for everyone? When will people be allowed or obligated to defer to the advice of their genies, be excused in case they do or blamed if they don’t?

The genies—as augmentations of the human will—have already started to come out of AI’s magic bottles. It will be hard to get them back in.

Rationale

AI-powered assistants with natural language processing capabilities, and advanced measurement and detection capacities, are likely to play an important role in our lives in the coming decades. They will be inserted in the fabric of our daily lives, and the questions about their design and moral and social implications will be prominent.

Reflections

One of the core questions will concern the extent to which these assistive technologies can and ought to be designed in order to make human beings more pro-social and nudge for the greater good.

A decorative graphic consisting of a horizontal line extending from the left edge of the slide to a small white circle. From this circle, a vertical line extends downwards to another small white circle. From this second circle, a horizontal line extends to the right edge of the slide.

Appendix

Developing the Stories

The story-writing process took place over two-and-a-half days (Day 1 entailed Steps 1-5; Day 2, Steps 6-8; and Day 3, Steps 9-11). Here we briefly describe the sequence of activities that our story writers and supporting artists moved through to develop the stories included in this book. We provide a copy of the Story Toolkit (the personal story sheets and instructions) that support story development. We hope that others will use and adapt these tools for writing their own culturally responsive stories around AI and other technologies.

Story Writing Process

Step 0 — Determining Who to Invite

Step 1 — Introductions and Fabric Pieces

Authors were asked to bring a piece of culturally relevant fabric with them to the workshop. When making their introductions to each other, authors situated themselves geographically and culturally and then introduced their fabric piece, telling the story of its origin and importance.

Step 2 — Surfacing Issues

Authors began by surfacing the issues, questions, challenges, and opportunities they saw that concern culturally responsive AI. As authors voiced issues, they recorded each issue on a river rock to make a visible and persistent reminder of the idea. These rocks were displayed on a piece of cloth from Uganda for the duration of the story generation sessions.

Step 3 — What’s in a Story?

Next, authors discussed how stories work and why they matter, also touching on how stories can be structured. They were then introduced to the particular process they would use to generate the stories. They formed writing groups of three authors from different cultures and backgrounds to be discussion partners.

Step 4 — Culturally Situating Your Story

The authors were reminded that stories can begin anywhere—culture, technology, policy. Each author was given a personal story sheet (see “Story Tools”) and asked to collect their initial ideas, making notes on their story sheet as appropriate. This first story sheet was intended to help authors situate potential story ideas in their own personal context (or contexts) as well as to prompt thinking around the wider AI landscape. Authors initially noted their particular cultural and political environment. Then they identified different AI technologies and/or applications, as well as different stakeholders possibly impacted by these technologies. Finally, authors were prompted to consider potential promises, opportunities, and potential perils for AI systems. After ten minutes, each author shared their initial ideas with their writing group. Then everyone continued to write or elaborate individually on aspects of their story ideas for another fifteen minutes. Artists rotated among the groups during the course of the story generation, creating images in response to the early story ideas they were hearing.

Step 5 — Reading Early Story Ideas Out Loud

Each author read their story out loud while everyone else wrote comments on sticky notes. Even at this early stage, the diversity and power of the story ideas were apparent. After everyone had shared their story, the sticky notes were shared back to the story author, and there was a little more time for authors to refine the stories.

Step 6 — Developing a Narrative

The authors gathered again in their writing groups. This time they were given a personal story sheet that emphasized narratives. The narrative sheets were intended to help authors select particular details (e.g., context, technologies, perils) from their earlier brainstorming in order to use them as elements in their story. Authors were first prompted to list their main ideas. Next, the story sheets encouraged the authors to make specific choices around characters (e.g., stakeholders and technology) and setting (e.g., when and where). Finally, the authors were prompted to sketch or outline the main events of their story.

Step 7 — Sharing Stories Out Loud

The authors reconvened again in a large circle and went around the room having each author read their story out loud.

Step 8 — Technical, Policy, and Other Responses

Authors gather in their writing groups again, this time to engage explicitly with policy, law, or other societal considerations using the third personal story sheet. Specific prompts include: If you were to envision a role for your story in policy, where would that be? What are some policy implications? What principles, values, and/or morals does your story illustrate? What are the use cases—ideal and unintended? Who is your audience (e.g., engineers, policymakers, politicians)? Artists continued to circulate and generate images in response to the story development.

Step 9 — A First View of the Book of Stories

Overnight, the organizers assembled a “draft” book containing all of the stories in progress along with some preliminary artwork. Each author received an 11 × 17 page with their story and some of the images from the artists as well as a photo of their fabric. The authors convened in a circle and heard a reading of all the stories.

Step 10 — Missing Bit and Elaborations

Authors had an opportunity to identify any missing elements in their stories and make any improvements they wished, including soliciting additional artwork from the artists.

Step 11 — Reflection and Next Steps

The summit closed with a conversation of reflections on the process and next steps.



Unintended Consequences

- Mistakes
- Imperfections
- Bias
- AI for war?
- Revoking and protecting AI tech
- What are the unidentified consequences of using AI to mine human emotion?
- Are these systems safe?
- How to avoid Cambridge Analytica on steroids
- Collateral damage
- Unintended consequences?
- How can AI emulate compassion without being manipulative?

Diversity and Culture

- How about AI existing (+unknown) inequalities?
- Diversity
- How do we celebrate and defend diverse, localized application & innovation of AI?
- Culturally responsive centaur systems?
- What is every culture's ideal society?
- Who is not represented in AI?
- Marginalized people
- With AI how do we preserve cultural means of expression?
- How can we empower everyone to welcome AI?
- How does AI adapt to changing society rather than solidifying a certain moment in time?

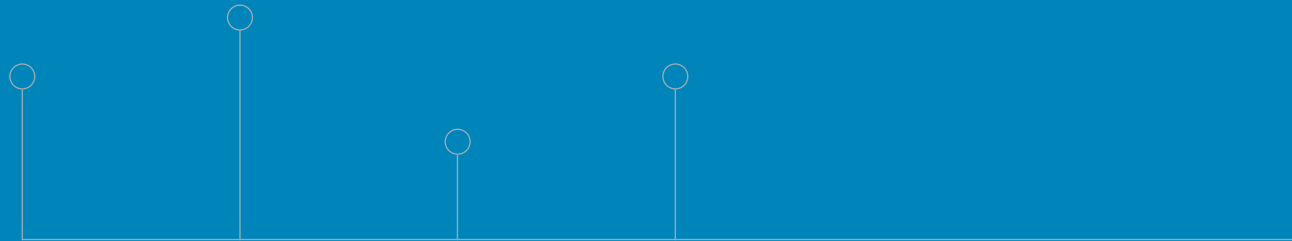
Education and Communication

- How to communicate about AI across cultures? What stories should we tell?
- Is it possible to explain how AI works?
- Hype
- How to put the user at the center of AI?
- How will humans retain the skills in case AI breaks?
- How do we teach AI to work well with humans?
- Enhance education
- Reduce number of accidents
- Help to pinpoint issues
- Capacity building (education and AI)
- Balance between knowledge (imported opinions) and popular opinions

Story Toolkit

The following personal story sheets and instructions were used to support story generation, guiding authors to focus on three areas: landscape, narrative, and positioning impact. “Part 1: Landscape” (blue sheets) focuses on the cultural and political context in which their stories would be situated, including AI technology, stakeholders, as well as promises, opportunities, and perils. “Part 2: Your Narrative” (gold sheets) focuses on the story narrative, including main ideas, characters, stakeholders, AI technology, and setting. Finally, “Part 3: Positioning Impact” (green sheets) focuses on social and societal impacts, including principles, audiences, use cases (ideal and unintended), as well as policy implications.

Part I: Landscape





AI TECHNOLOGY



A large white rectangular area with horizontal dotted lines, intended for writing or drawing.



STAKEHOLDERS



A large white rectangular area with horizontal dotted lines, intended for writing notes or a list of stakeholders.



PROMISES & OPPORTUNITIES



A large white rectangular area with horizontal dotted lines, intended for writing or notes.



PERILS



A large white rectangular area with horizontal dotted lines, intended for writing or drawing.

A solid gold background with two white decorative elements: a vertical line starting from the top edge and ending in a small circle, and a diagonal line starting from a small circle and ending in a right-angle corner at the bottom right.

Part II : **Your Narrative**



MAIN IDEAS



A large white rectangular area with horizontal dashed lines for writing.



CHARACTERS



Stakeholders

A series of horizontal dotted lines for handwritten notes under the 'Stakeholders' heading.

Technology

A series of horizontal dotted lines for handwritten notes under the 'Technology' heading.



SETTING



When

Handwriting practice lines for the 'When' section.

Where

Handwriting practice lines for the 'Where' section.



WHAT HAPPENS



A large white rectangular area with horizontal dashed lines, intended for writing a narrative.

Part III: Positioning Impact





PRINCIPLES



A large white rectangular area containing 25 horizontal dotted lines, intended for writing notes or principles.



AUDIENCES



A large white rectangular area with horizontal dotted lines, intended for writing or drawing.



USE CASES



Ideal

Ruled area for writing ideal use cases, consisting of multiple horizontal dashed lines.

Unintended

Ruled area for writing unintended use cases, consisting of multiple horizontal dashed lines.



POLICY IMPLICATIONS



A large white rectangular area with horizontal dotted lines, intended for writing notes or responses.

Who the authors are

Dennys Antonialli

is the executive director of InternetLab, a think tank on internet law and policy based in São Paulo, Brazil. He specializes in technology policy and fundamental rights, particularly in privacy, freedom of expression, and state surveillance. Having lived and studied in the United States and in Germany, he has a soft spot for multicultural discussions and international perspectives on tech policy. He is an active voice in the LGBTQ community and deeply cares about policies that promote diversity and inclusion in the tech world or outside of it.

Chinmayi Arun

has been teaching law since 2010 and resided in Delhi from 2012 to 2018. She was founding director of the Centre for Communication Governance at National Law University, Delhi, and is a member of the United Nations Global Pulse Data Privacy Advisory Group as well as UNESCO India's Media Freedom Advisory Group. She is currently a fellow of the Berkman Klein Center at Harvard University and has been a past consultant to the Law Commission of India and the Indian government's multi-stakeholder advisory group for the India Internet Governance Forum.

Joanna Bryson

is a reader (associate professor) in artificial intelligence at the University of Bath. She is a transdisciplinary researcher on the structure and dynamics of human and animal-like intelligence. She is interested in everything but particularly cognition, and since 2010 she has worked mostly on the ethics of, for, and about AI, as well as on human cooperation more generally. She grew up in a small town in the upper Midwest (Illinois), her mother's parents were farmers, and her father was a salesman and an engineer. In 1991, she

moved to the EU (to Edinburgh, Scotland, where she was later married), and since 2002 she has worked in another small, beautiful city in the southwestern part of England.

Darren Byler

is a postdoctoral lecturer in the Department of Anthropology at the University of Washington, Seattle. His teaching and research focus on dispossession, culture work, and securitization in China and around the world. In 2003, he first went to the Uyghur homeland, where he learned to speak Chinese and Uyghur. Intrigued by the way digital media was allowing the growth of new forms of cultural and religious expression among Uyghurs when internet usage became widespread in 2011, he spent two years in the region observing how Uyghurs were subsequently labeled “unsafe” technology criminals. As in any society, the surveillance systems that were put in place reflected the values of those in power. An advocate for native (Uy: yerliq) Uyghur rights, he analyzes and amplifies the stories of those who are dispossessed through such systems.

Ryan Calo

is the Lane Powell and D. Wayne Gittinger Associate Professor of Law at the University of Washington, Seattle, and an associate professor by courtesy at the Information School. He researches the ways law reacts to and channels emerging technology. As a child, he spent years in Florence, Italy, and speaks Italian. He enjoys playing chess without being particularly good at it.

Jeff Cao

is a senior research fellow at Tencent Research Institute, where he researches AI governance and ethics, regulation of autonomous driving systems and other emerging technologies, legal tech and digital justice, as well as data governance

and cyberlaw. His master's degree in intellectual property law is from Southwest University of Political Science and Law (SWUPL). He is an expert at the Internet Society of China, a special research fellow at the Institute for AI Moral Decision-Making of Hunan Normal University, and a part-time researcher at East China University of Political Science and Law. He has published dozens of articles on the internet as well as in newspapers and academic journals, including "On Civil Liability of Artificial Intelligence," "Ten Big Trends of Legal AI," and "How to Deal with Ethical Problems of AI." His experience in China's internet industry, his contact with policymakers and regulators, and his connection with academia help him see the many issues around the internet and other emerging technologies in a holistic and systematic way.

Jack Clark

is the policy director of OpenAI and the author of *Import AI*-a newsletter that analyses contemporary AI research and features an original fiction story about AI every week, read by more than twenty thousand people worldwide. He participates in the Center for a New American Security (CNAS) Task Force on AI and National Security, and he sits on the steering committee of the AI Index, an initiative from the Stanford One Hundred Year Study on AI to track and analyze AI progress. In 2018, he participated in the Assembly program on ethics and governance in AI at the MIT Media Lab and the Berkman Klein Center at Harvard University.

Batya Friedman

is a professor in the Information School and co-director of the Value Sensitive Design Lab and Tech Policy Lab at the University of Washington, Seattle. Her research focuses on foregrounding human values in the technical design process, including longer-term design processes. The Voices from the Rwanda Tribunal, which forms the basis of her story, is her

first multi-lifespan design project. Within that context, she has spent the past decade thinking about the role of information systems in realizing meaningful justice.

Sue Glueck

is the senior director of Academic Relations at Microsoft, located in the Seattle region. She has been at Microsoft for almost twenty years, including eleven years leading a team that provided privacy advice to engineering groups that create software products and online services. Her real claim to fame, however, was when the *Seattle Times* proclaimed that the privacy policy she created for Internet Explorer 8 "breaks the mold with what may be the best written privacy policy for any software product ever." Outside of work, she has a tremendous love for animals from the big cats of southern Africa to her small but mighty cat Tatters at home.

Sabine Hauert

is assistant professor in robotics at the University of Bristol in England. Her research focuses on engineering swarms across scales, from nanoparticles for cancer treatment to robots for rescue operations. She is president and co-founder of Robohub.org, a nonprofit dedicated to connecting the robotics community to the world. As an expert in science communication with ten years of experience, she is often invited to discuss the future of robotics and AI. She has a six-month-old daughter whose favorite plush toy is a robot.

Alejandro Hevia

is assistant professor at the Department of Computer Science in the University of Chile in Chile. As a cryptographer working on privacy and public policy, he wonders about the limits between mathematics (cryptography) and the public perception of its utility on human life. Cryptographic algorithms are notoriously hard to understand for humans,

yet they may provide seemingly magical tools: from inclusive electronic voting to efficient decentralized currencies and privacy-preserving fair government policies. To see this technology deployed, scientists should look beyond formulas and programs, and understand what drives human trust on technology and science.

Ian Kerr

is a Canadian scholar whose work attends to the intersection of ethics, law, and technology. His ongoing graphic novel project, *A Planet of Playthings*, explores tensions between artificial intelligence and humanism. This book is dedicated to Ian Kerr (1965 – 2019).

Tadayoshi Kohno

is a professor in the Paul G. Allen School of Computer Science & Engineering at the University of Washington, Seattle. His research focuses on helping protect the computer security, privacy, and safety of users of current and future generation technologies. He holds black belts in the traditional Wado-Kai and Hayashi-ha Shitoryu-Kai styles of karate. His first karate teacher was Sensei Hiroyuki Tanabe, and he currently studies under Shihan Akio Minakami in Seattle. He is also passionate about training in aikido and kendo. Continuing the family tradition, his oldest son trains aikido with Sensei Koichi Barrish, and his youngest son trains karate with Shihan Minakami.

Lisa Nathan

is associate professor and coordinator of the First Nations Curriculum Concentration at the University of British Columbia's iSchool (School of Library, Archival and Information Studies) in Canada. She is a settler of mixed European heritage living on the territory of the xwməθkwəyəm (Musqueam).

Joseph Nkurunziza

is the co-founder and executive director of Never Again Rwanda, a peace building and social justice organization that promotes human rights and advocates peace among the youth and the Rwandan. He is a medical doctor with over ten years' experience in the clinical field, public health, and peace building. He is also deeply involved in sustainable development.

Nnenna Nwakanma

is the interim policy director with the World Wide Web Foundation, Côte d'Ivoire. She is a Digital Equality and Women in Tech activist and leader. Named among the "World's 100 Most Influential Persons in Digital Government," she is passionate about Christ, the World Wide Web, food tourism, and soccer. She introduces herself as "Nnenna from the Internet." She is living with internetaholic conditions and has been evading rehab for at least two decades.

Amir Rashidi

is the Director of Digital Rights and Security at the Miaan Group. He has over 10 years of experience in digital security and rights in Iran. He is an expert on Iranian Internet censorship, cyber-attacks, and security trends. He has conducted tens of digital security audits, training, and rapid response actions for Iranian human rights defenders and organizations. He is in a unique position to assess the security risks and needs of Iranian organizations, especially those who are in contact with at-risk people, such as activists and journalists. He worked as the internet security and digital rights researcher at the Center for Human Rights in Iran, where he conducted in-depth research, data collection, and analysis on internet security and access in Iran, internet policy and infrastructure, and the tools and methods of state-sponsored censorship and hacking.

Rohan Samarajiva

is a fan of near-future science fiction, and most of his writing is policy-related. Before taking on his current government responsibilities, he used to write fact-based pieces for the media in multiple countries. He sees the littoral of the Bay of Bengal as home and speaks of how the people around the Bay can emerge from poverty, based on evidence assembled by LIRNEasia and other organizations.

Jeroen van den Hoven

is a university professor and professor of ethics and technology at Delft University of Technology, in the Netherlands. He works in the field of moral philosophy and technology and more specifically in applied ethics of information technology. He is scientific director of the Delft Design for Values Institute, founding editor in chief of Ethics and Information Technology, and permanent member of the European Group on Ethics, which advises the president of the EU on ethics and innovation.

Story Index

- Africa **32, 46, 48**
 - Rwanda* **32, 46**
- AI assistant **18, 36, 54**
- Algorithms
 - automated decision making* **18, 36, 40, 48, 50, 52**
 - transparency* **38**
- Asia
 - China* **24**
 - India* **20**
 - Iran* **50**
 - Japan* **42**
- Augmented reality **36**
- Autonomous vehicles **18, 36**
- Biometrics **18, 20, 24, 48, 50**
- Brain implant **30**
- Chatbots **46**
- Child **30, 36**
- City, see *urban*
- Community
 - virtual* **24**
- Companionship **28, 34, 36**
- Cost **30, 34**
- Cultural knowledge **32, 42, 44, 46**
- Data and databases
 - national ID cards* **20**
 - reliability of* **20, 24, 48, 50**
- Deception **22, 54**
- Dignity **32**
- Disinformation **22**
- Education **36, 42, 44**
- Elections **22**
- Emotion **24, 34**
- Emotion recognition **24**
- Ethics
 - corporate* **22, 26, 34, 42**
- Ethnoracialization
 - institutionalized* **24**
- Europe
 - England* **36, 50**
- Fable **22, 54**
- Facial recognition **20, 24, 48, 50**
- Global South **18, 20, 24**
- Golem **40**
- Human in the loop **40**
- Immigration **48, 50**
- Indigenous peoples,
 - see *North America, First Nation peoples*
 - see *Uyghur*
- Insurance **52**
- Islam **24, 18**
- Judaism **40**
- Justice **32**
 - international* **32**
- Karate **42**
- Land **44**
- Language
 - embedded values* **44**
 - programming* **44**
- Law **28, 32**
- Local norms and customs **18, 46, 48**
- Location tracking **18, 20, 52**
- Medical technology **30**
- Military technology **40**
- Mobile phones **46**
- Natural disasters **38, 52**
- Neurodiversity **30**
- North America
 - Canada* **42, 44**

First Nation peoples **44**
United States **18, 34, 42, 48, 52**
Obsolescence **34**
Power
 corporate **22, 34**
 gender **20, 28**
 government **20, 24, 148, 50**
Prague **40**
Reeducation camp **24**
Religious freedom **24**
Responsibility **22**
 corporate **22, 34, 42**
Risk **30, 40, 42, 52, 54**
Robots
 automation **32, 38, 48**
 emotions **28, 34, 38**
 heroes **38**
 humanoid **28, 34**
 love interest/spouse **28, 34**
 rights **28**
Rural
 town **34, 38**
 village **24, 32, 38**
Safety **20**
Sensors
 remote **52**
Smartphone **24**
Smartwatch **18**
Social media **22, 24**
South America
 Brazil **18**
Spontaneity **18**
Surveillance **20, 24, 48, 50**
Technology
 use and scale over time or long-term use **26, 42**
Trust **20, 22, 32, 36**
Urban **18, 21**
Uyghur **24**
virtual reality **36**
voice recognition **50**
wearables **18, 54**
wisdom **40, 54**
work
 future of **36**

Acknowledgments

The editors of this book would like to thank our funders, the William and Flora Hewlett Foundation, the John D. and Catherine T. MacArthur Foundation, Microsoft, and the Pierre and Pamela Omidyar Fund at the Silicon Valley Community Foundation.

We also thank the artists, Tim Corey from Colibri Facilitation, Akira Ohiso, and Elena Ponte. The following team members supported this project: Jabu Diagana, Ivan Evtimov, Madeline Lamo, Emily Brady Rosenfeld, and Hayley Younghusband.

And our enormous thanks go to the authors for joining us on this creative journey.





About Telling Stories

What world, what worlds will we build with artificial intelligence?

Intended for policymakers, technologists, educators and others, this international collection of 19 short stories delves into AI's cultural impacts with hesitation and wonder. Authors from Brazil, Canada, Chile, China, India, Rwanda, Sri Lanka, the United States, and elsewhere vividly recount the anticipated influences of AI on love, time, justice, identity, language, trust, and knowledge through the power of narrative.

Deceptively simple in form, these original stories introduce and legitimate perspectives on AI spanning five continents. Individually and together, they open the reader to a deeper conversation about cultural responsiveness at a time of rapid, often unilateral technological change.