

The Opportunities and Risks of AI

Jutta Treviranus
Inclusive Design Research Centre
OCAD University
Toronto, Canada

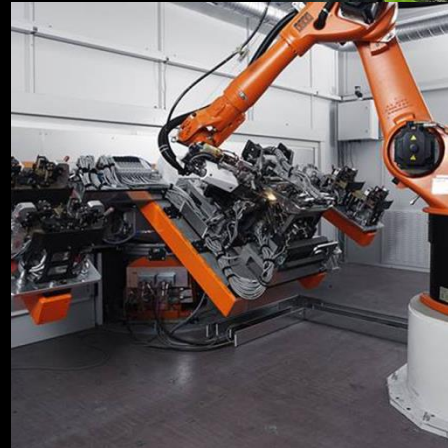
The ancestral and traditional territories of the Mississaugas of the Credit, the Haudenosaunee, the Anishinaabe & the Huron-Wendat

[Attribution-NonCommercial 4.0 International License](#)

... a statistical reasoning machine

feeding on data

a power tool...



More
efficiently,
accurately,
consistently..

mechanizing, accelerating,
amplifying, and automating
existing patterns

Statistically finding or producing the...

typical

popular

normative

predictable

statistically average

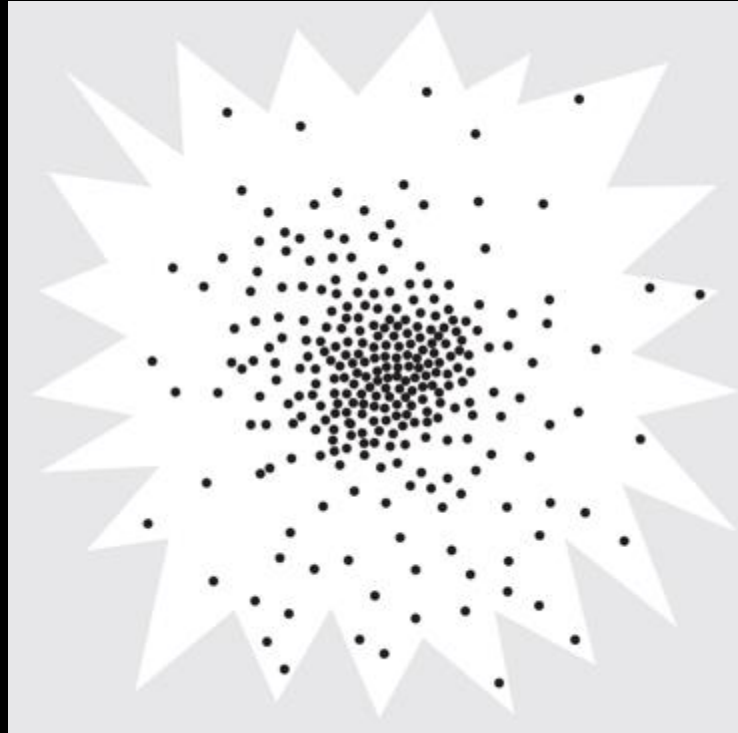
Generations of AI

- **First:** “here are the rules, follow them literally and accurately”
- **Second:** “here is available data, use statistical reasoning to optimize the selected metrics”
- **Third:** “here is all the data we have, you figure it out and make new connections and determine weights based on past relationships and possible combinations.”

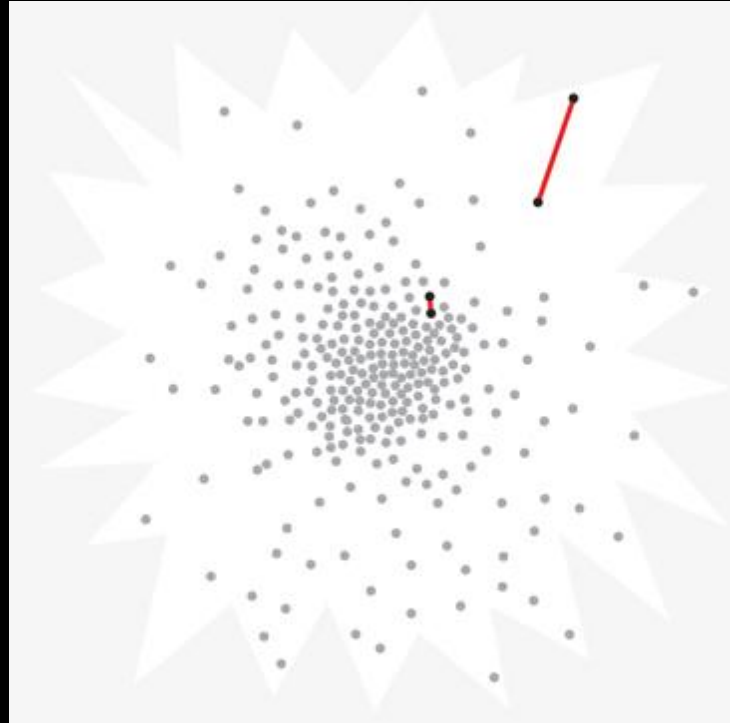
Disability & data..

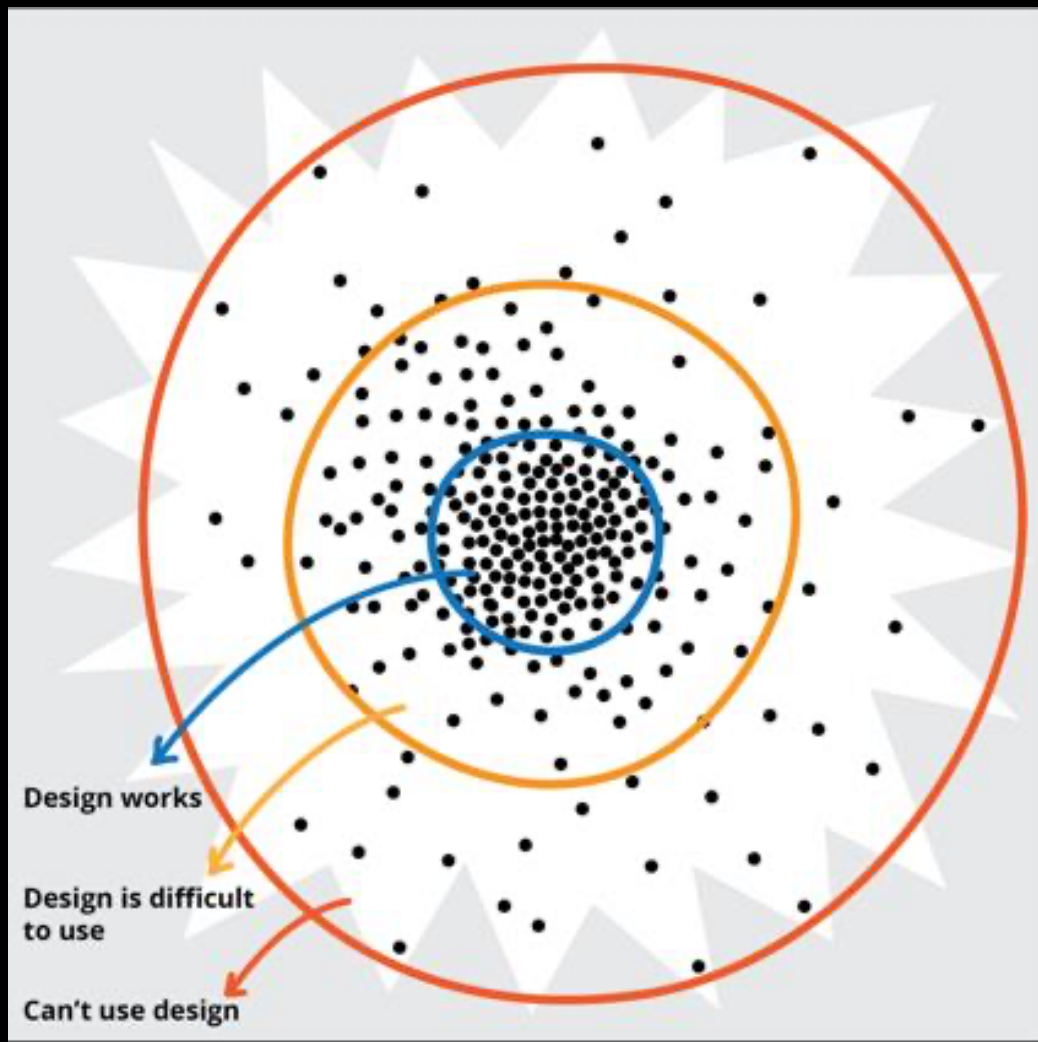
My data set of human
requirements....

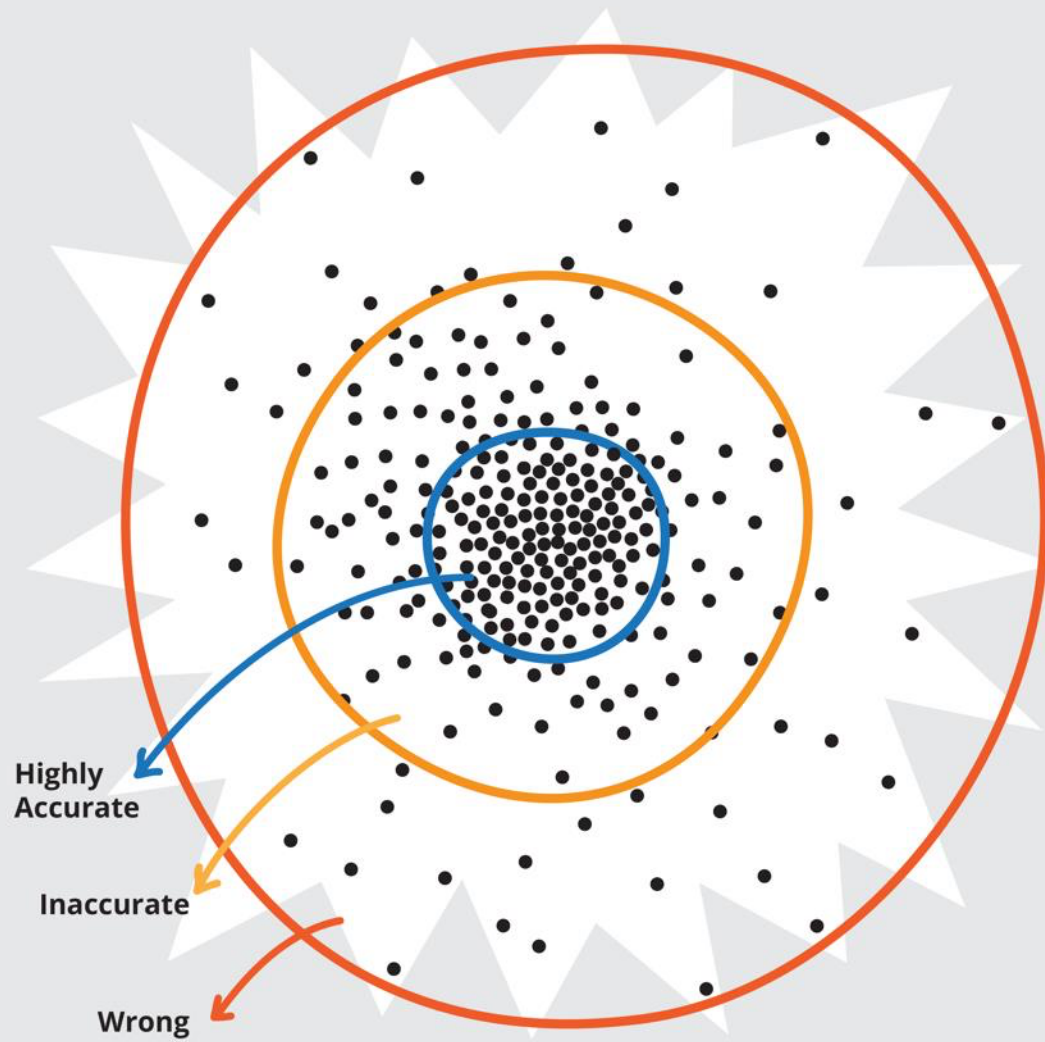
Our human starburst...

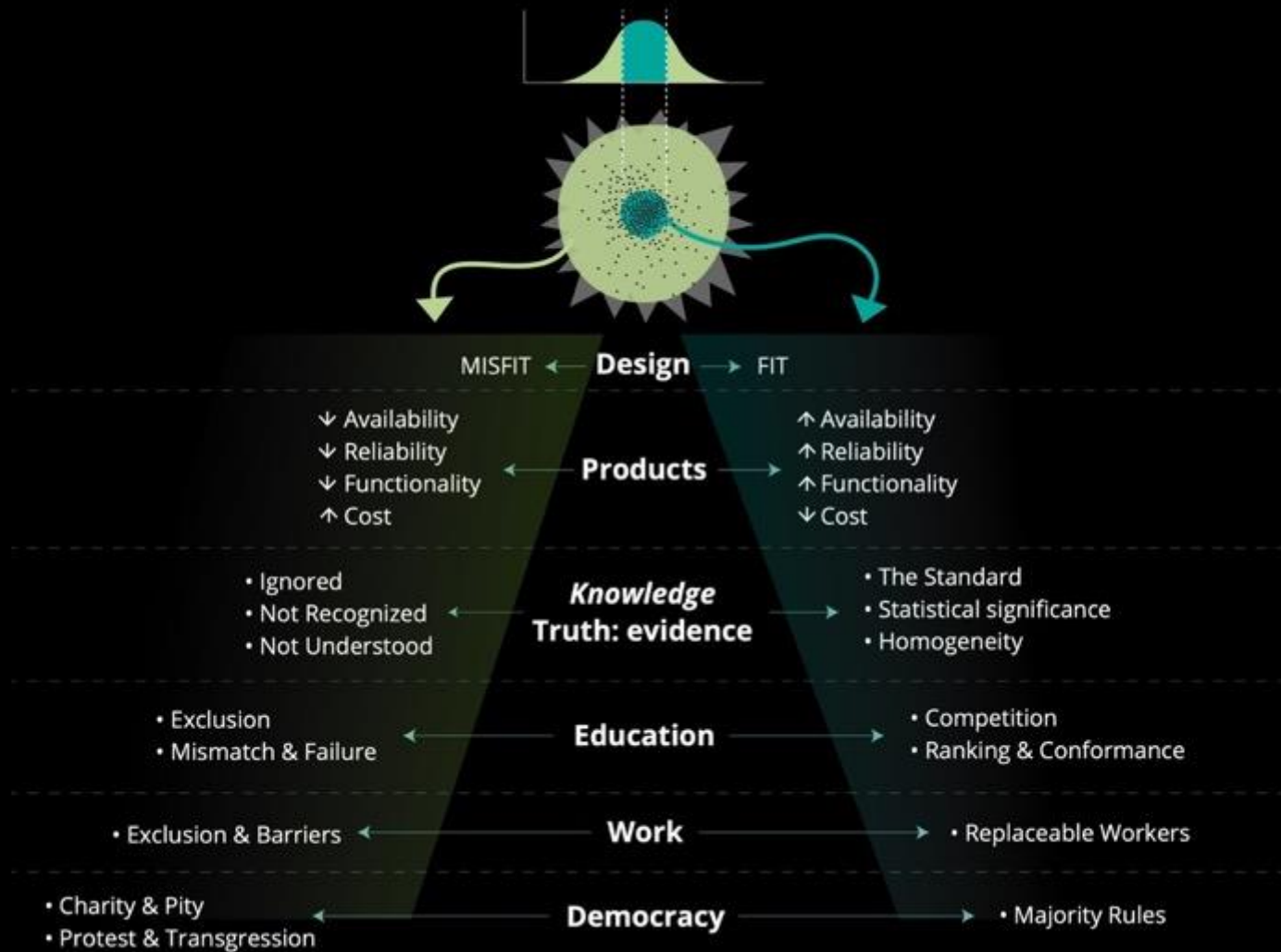


Difference...

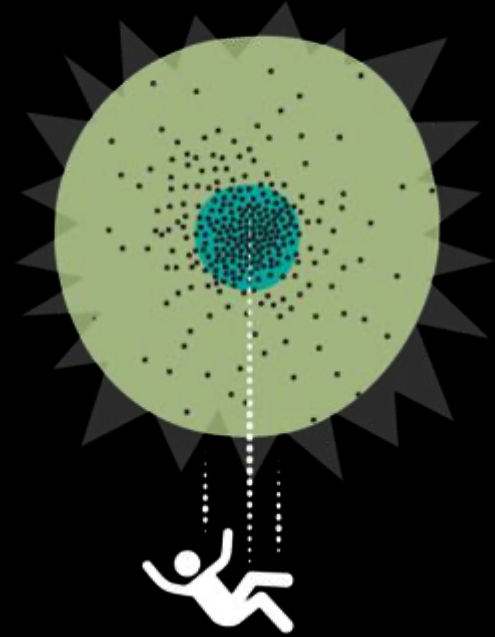




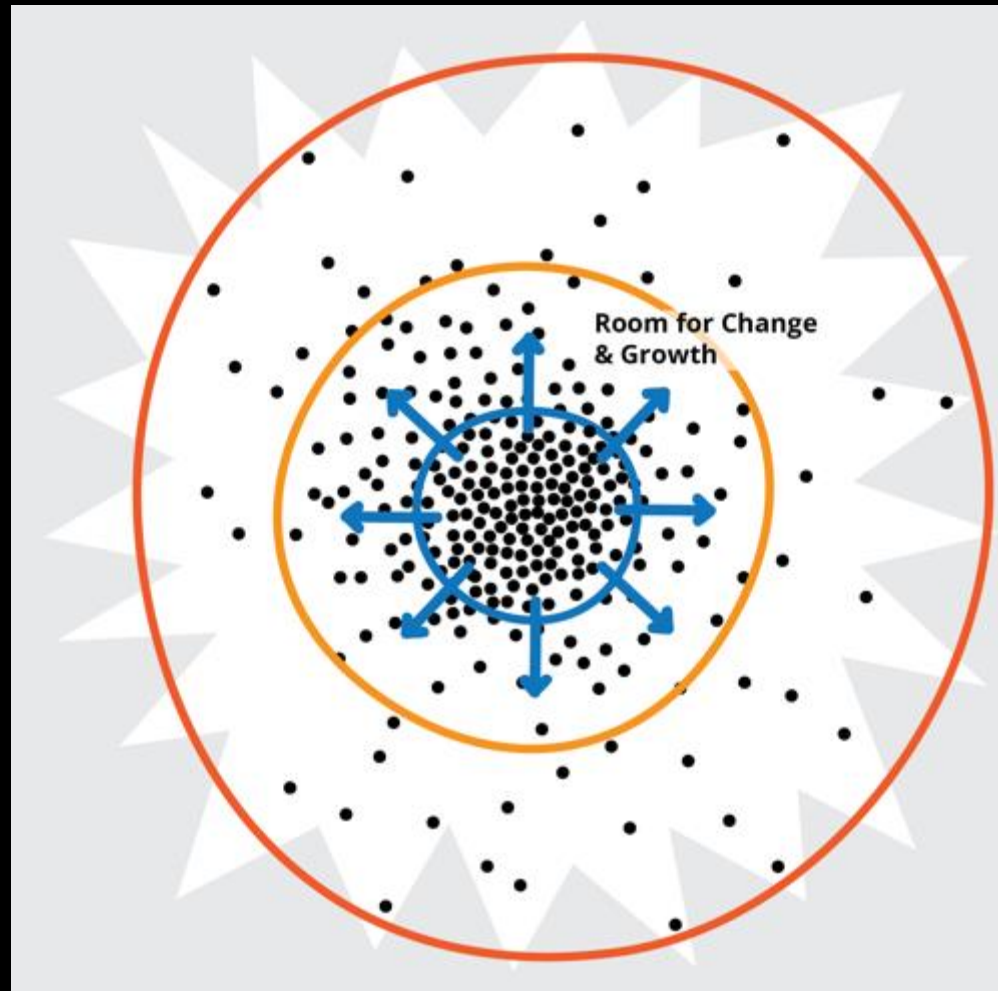




The prediction with the greatest probability

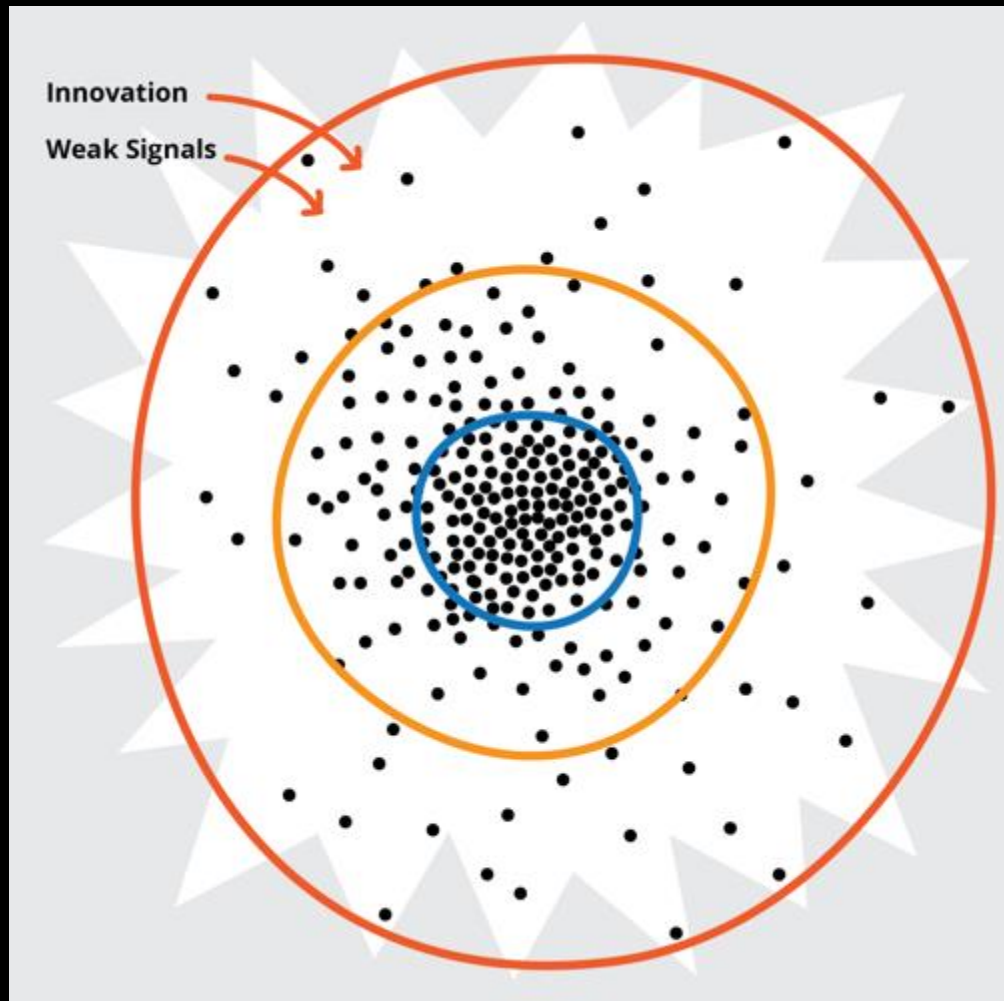


- Anytime
- When you least expect it
- Greatest Probability



Innovation

Weak Signals



If you have a disability the
opportunities and risks are at the
extremes

“Poster child of AI”



“...trust us to break things, to
make mistakes, to push forward.”

disability & technology

“for most people, technology
makes things **convenient**,

if you have a disability,
technology makes things
possible.”

To speak, read, write, learn,
effect the world,
navigate the world,
eat,
express love,
remember, plan,
breathe, ...live

Extreme Opportunities..

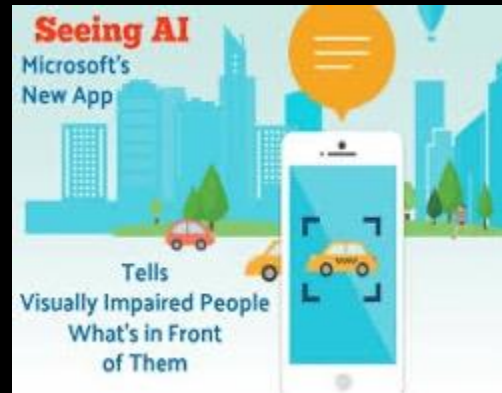
- Recognize speech, gestures, patterns
- Find a target object or pattern
- Match and label objects
- Remember forever and remind on time
- Sort possible paths to find the optimum
- Detect common mistakes and correct them

Mechanizing the Formulaic..



Current Assistive Technology (AT)..

- Image to text description (e.g., Be My AI, Seeing AI, JAWS PictureSmart AI)
- Real-time captioning (e.g., Zoom, TranscribeGlass)
- Sound filtering (e.g., Orka Two)

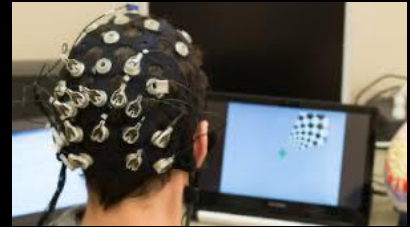


[illegible]

- Voice control (e.g., Voiceitt)
- Personal Voice Output (e.g., Apple Personal Voice)
- Exoskeleton optimization (e.g., Stanford E.)
- Gesture Control (e.g., Cephable, The Wheelie)

...and More Current AT

- Smart home systems, & robotic assistants
- Memory & recall (e.g., Rewind AI)
- Simplification & Summation (e.g., Detangle AI)
- Writing support (e.g., Chat GPT)....
- Brain Computer Interface



The flies in the ointment..

Privacy Breaches

- Inaccessible terms of service
- Data sharing – who else gets the data?

Surveillance & manipulation

- AT user and others
- Targeted advertising
- Stereotypical optimization metrics

Homogenization

- Impact of “being typical”
- Loss of differentiation

Inverse Benefits

- Works the worst for people that need it the most
- Assumed accommodation

Connectivity

- Cost of bandwidth
- Power consumption and sustainability

Hallucination

- Sounding believable, not telling the truth
- “Producing believable bull**”

What data about disability have we fed AI?

- Ablist slurs
- Disabled people as inferior, unattractive, tragic victims, suffering
- Bad jokes about disability
- Benevolent ableism – “inspiration porn”

Toxicity Censors

- Captioning toxicity filters
- Facial difference cancellation
- Disability without sex

Less known risk: an infrastructure
of statistical discrimination

my first alarm...



Alice and the unexpected...



Smarter is not always
better...



Finding, matching, sorting,
labeling, measuring, optimizing,
calculating, analyzing people... at
scale

there is a problem..
&
we are the collateral damage

Bias toward optimal pattern
=
Bias against difference

..propagating discrimination
faster, more efficiently, and
accurately

homogenizing toward a
monoculture..

Pervasive....

- Employment
- Academic admissions
- Medical calculators & triage tools
- Policing and parole
- Immigration & refugee adjudication
- Tax auditors....loans, mortgages...security flagging

AI getting better or more accurate
=
getting more discriminatory

Bad, unfair, inaccurate
decisions

Subject to false flagging in
suspicion systems

Community level..

- Evidence based government investment
- Political platform
- Public health decisions
- Urban planning
- Emergency preparedness, security measures...

Cumulative impact...

- News coverage
- Products marketed
- GPS routes
- Supply chain priority
- Design features selected....

Harm and incident databases

- Disproportionate reports related to disability
- Examples: parents with disabilities falsely flagged as unfit, false positive tax audits, false positive security flagging, excluded from optima filtering in employment and education, biased credit and asset rating, unfairly denied mortgage and loans, unfairly high insurance rates, deprioritization of relevant information, iatrogenic death and illness...

Statistical reasoning
as the means of making decisions
does harm

Assuming that what we know about the
majority applies to the minority does
harm

AI as a mirror



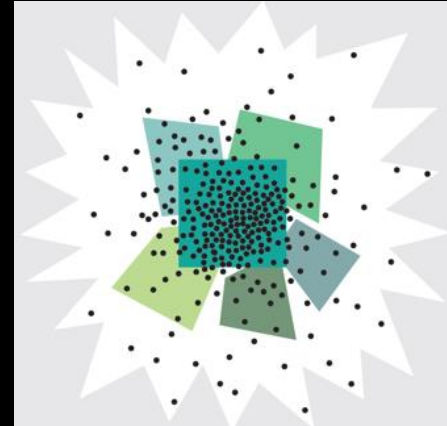
Credit image: Photo by Marc-Olivier Jodoin on Unsplash

More than...

- Addressing data gaps
- Removing human bias from algorithms
- Removing stereotypes from labels and proxies

Statistical Discrimination in AI Ethics Measures

- Insignificant when determining impact metrics and thresholds
- Invisible or insignificant in error testing and risk-benefit framework
- No bounded cluster in cluster comparison



“an anomaly”

“merely an anecdote”

Privacy protections don't work

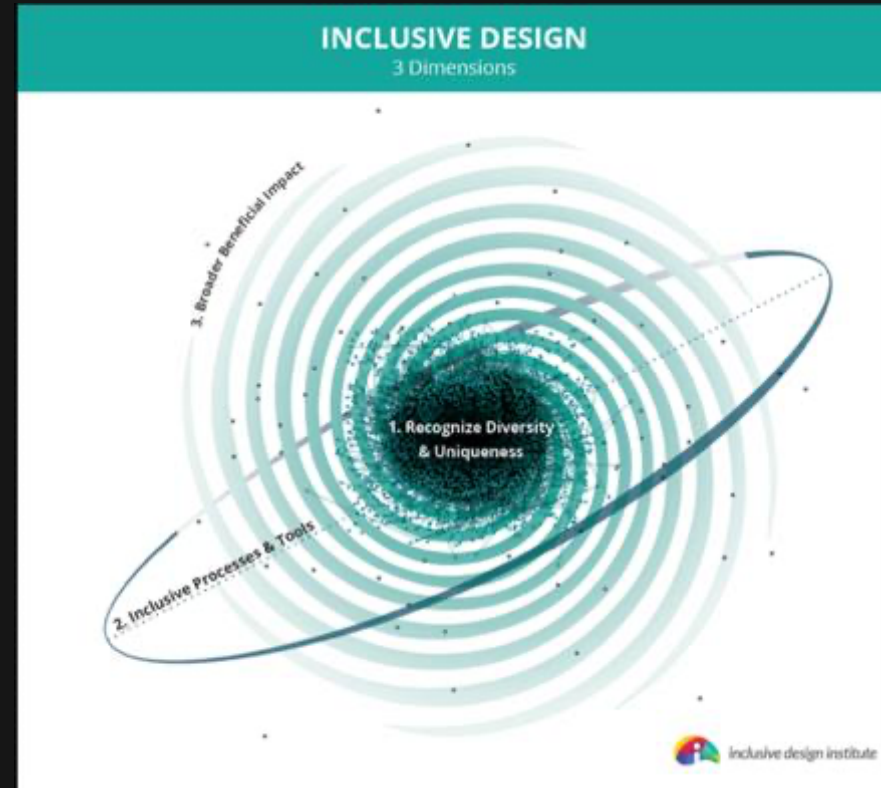
- Privacy bartered for essential services
- Re-identification if you are unique
- Differential privacy eliminates helpful data
- Need recovery from data abuse & misuse absent



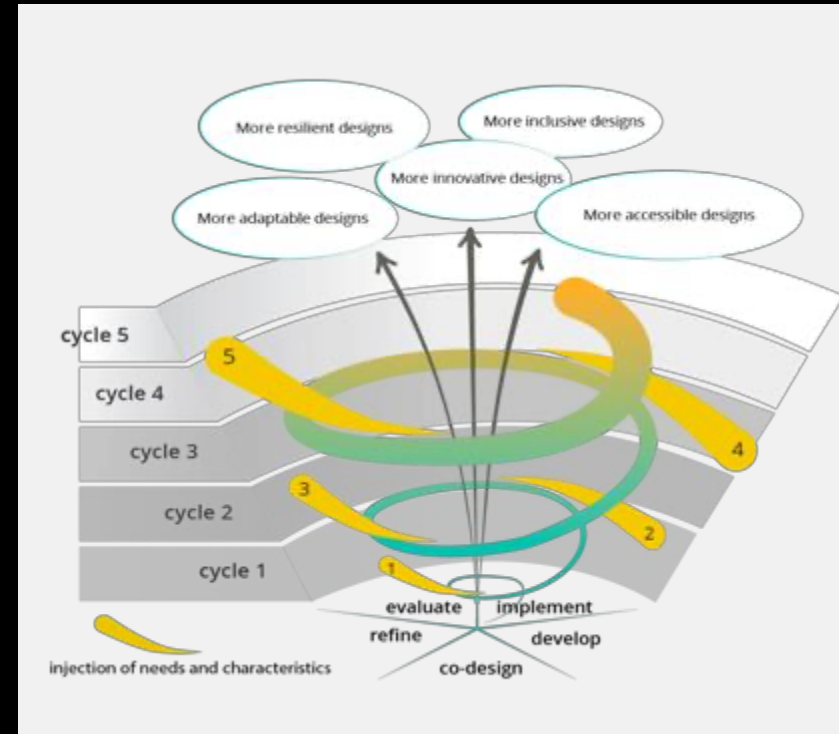
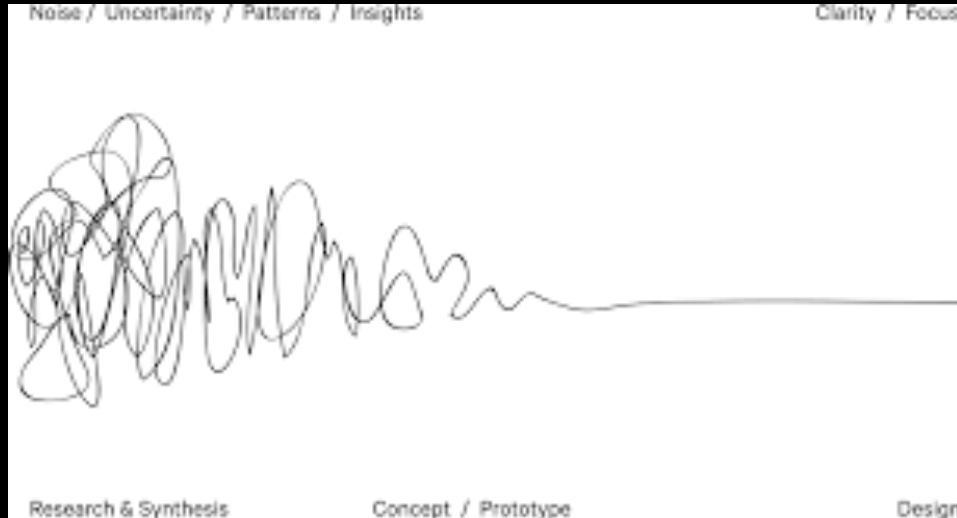
How are we addressing
trustworthy AI?

Inclusive Design..

1. Recognize diversity, variability, & self-knowledge in integrated design
2. Inclusive process – co-design – diverse team- “who is missing?”
3. Benefit-for-all - intervention in complex adaptive system, culture change

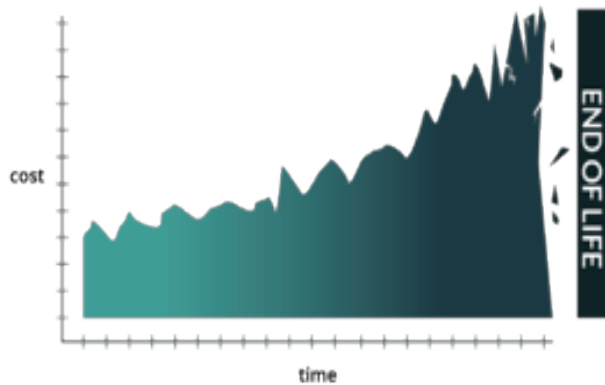


The Virtuous Tornado and the Design Squiggle

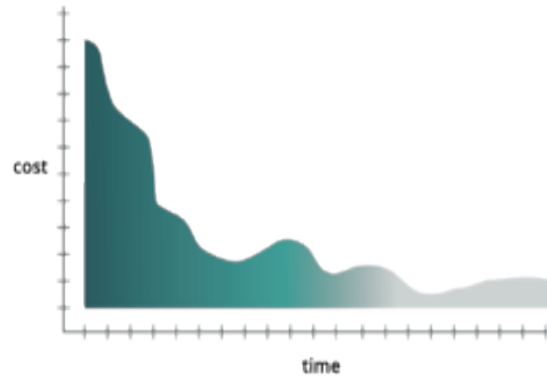


What about cost?..

Plan for the Centre



Plan with the Edge



Flipping the perspective

- Focus on the novel and unique
- Inverting the metrics



Invert the dominant algorithm

- From data exploitation to data exploration
- From homogenization to diversification
- Finding missing perspectives, increasing diversity
- In hiring & admissions

Prevent Harm

- "Trust meter"
- Suitability of decision system to instance before you
- Based on data used to train AI

Dataset Fact Sheet

Metadata






Title COMPAS Recidivism Risk Score Data

Author Broward County Clerk's Office, Broward County Sheriff's Office, Florida

Email browardcounty@florida.us

Description Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

DOI 10.5281/zenodo.1164791

Time Feb 2013 - Dec 2014

Keywords risk assessment, parole, jail, recidivism, law

Records 7214

Variables 25

priors_count: Ut enim ad minim veniam, quis nostrud exercitation

two_year_recid: Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

Missing Units 15452 (8%)

 This dataset contains variables named "age", "race", and "sex"

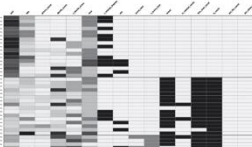
Probabilistic Modeling

Analysis

◀

12

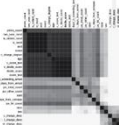
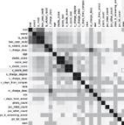
▶



◀

▶

Dependency Probability **Pearson R**

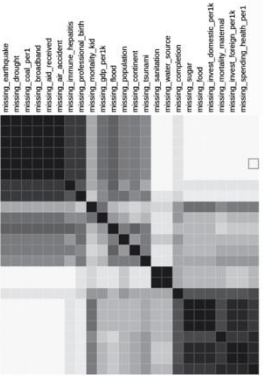
Missing Units

Clustering Variable

race ▶

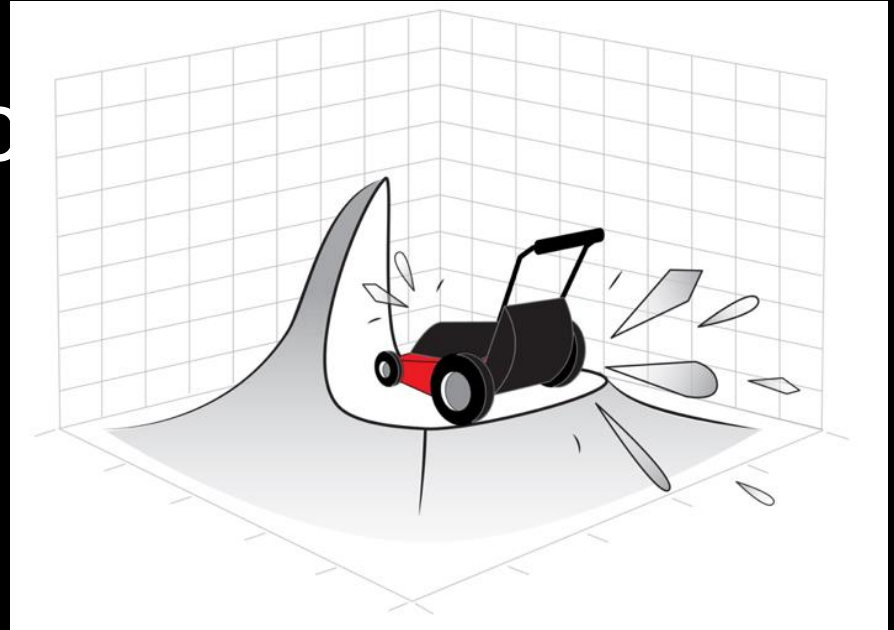
Missing Variable

t_days_from_arrest ▶



Remove advantage of
being like the average:

“lawnmower of justice



Pushing benefits to where they are needed the most...

“Baby Bliss Bot”



			
baby	Blissymbol	plan digital	ability self
	writing (for) world	software	independent

Filling the regulatory vacuum...

Accessible and Equitable AI
Standard

for the Accessible Canada Act

... and the EU and US

Phase 1: 4 Parts

- Accessible AI
- Equitable AI
- Organizational Processes that support Accessible and Equitable AI
- Education

Specific to Disability

- Accessibility: Nothing without us
- Statistical discrimination
- Cumulative harm
- Education

... progress at the speed of
trust

Intelligence that works with the edge of our human scatterplot is better able to ..

- Adapt to change & respond to the unexpected
- Detect risk
- Transfer to new contexts
- Results in greater dynamic resilience & longevity
- Will reduce disparity
- May hold the key to our survival

Continuing the conversation:

- <https://accessible.Canada.ca>
- <https://idrc.ocadu.ca>
- <https://wecount.inclusivedesign.ca>
- <https://medium.com/@jutta.trevira>

Supported in part by:



Accessibility Standards
Canada

Normes d'accessibilité
Canada

Funded by Government of
Canada's Innovation, Science
and Economic Development
Canada and Employment and
Social Development Canada

Canada

Funding Acknowledgement

Inclusive Design for Employment Access (IDEA) is supported by generous funding from the Government of Canada's New Frontiers in Research Fund (NFRF) Transformation Stream.

Le Vision Radicale Pour L'Accès Inclusif à l'Emploi (VRAIE) est financé par le fonds Nouvelles frontières en recherche du gouvernement du Canada.



Government
of Canada

Gouvernement
du Canada



New Frontiers in Research Fund
Fonds Nouvelles frontières en recherche

