DATA, RESPONSIBLY
#1
MachineLearnist COMICS
MIRROR, MIRROR
© Falaah Arif Khan and Julia Stoyanovich (2020)

TERMS OF USE

All the panels in this comic book are licensed CC BY-NC-ND 4.0. Please refer to the license page for details on how you can use this artwork.

TL;DR: Feel free to use panels/groups of panels in your presentations/articles, as long as you

1. Provide the proper citation
2. Do not make modifications to the individual panels themselves

Cite as:

Falaah Arif Khan and Julia Stoyanovich. "Mirror, Mirror". *Data, Responsibly Comics*, Volume 1 (2020)
https://dataresponsibly.github.io/comics/vol1/mirror_en.pdf

Contact:

Please direct any queries about using elements from this comic to themachinelearnist@gmail.com and cc stoyanovich@nyu.edu

Licensed CC BY-NC-ND 4.0

© Copyright Falaah Arif Khan and Julia Stoyanovich (2020)

ACCESSIBILITY STATEMENT

The purpose of scientific publication is the presentation of ideas and dissemination of findings. In the course of our (ongoing) work on creating a comic series about Responsible AI, we have found that relatable cartoons and visual humor are a rich but underappreciated source of clarity and accessibility that enable effective communication to a broad audience. Comic books are a particularly prescient medium for literature reviews and critical surveys, and for bridging insights from different disciplines such as philosophy, law, sociology, and computer science. Given the inherently interdisciplinary nature of machine learning, we see comics and other technical artwork as a promising new medium of scholarship. We hope to demonstrate their utility through our work and to popularize their adoption more broadly in the scientific community.

We care deeply about making our comics as digitally accessible as possible. Towards this end, we have taken the following measures:

1. We've chosen a typeface that was developed specially for dyslexic readers. All of the major text in the comic is in the "Open Dyslexic" font.
2. The comic book is fully alt-texted and can be read entirely using a screen reader. We are also releasing a complete transcript of the comic book, including all of the text and image descriptions.
3. We will be translating the comic into different languages to cater to speakers of languages other than English. Translated volumes can be accessed from the Data, Responsibly comics homepage.

We would like to thank Amy Hurst and Chancey Fleet for guiding us on the Accessibility front.

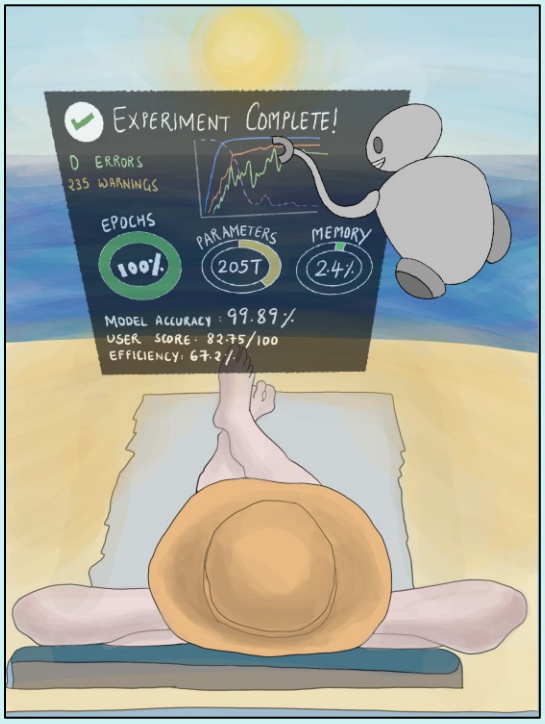
Please feel free to reach out to us if you have any recommendations on how we can further improve the accessibility of our comics.

HEY THERE!
YOU MADE IT!

WELCOME TO **OPTOPIA!** (1)

IT'S THE LAND OF ALGORITHM DRIVEN UTOPIA!

REMEMBER ALL THOSE CRAZY SCIENTISTS TALKING FOR DECADES ABOUT CREATING ARTIFICIAL INTELLIGENCE? WELL, THIS IS IT.



WE ALL LAUGHED AT THEM AND SAID IT WAS IMPOSSIBLE (2), BUT YOU KNOW WHAT...

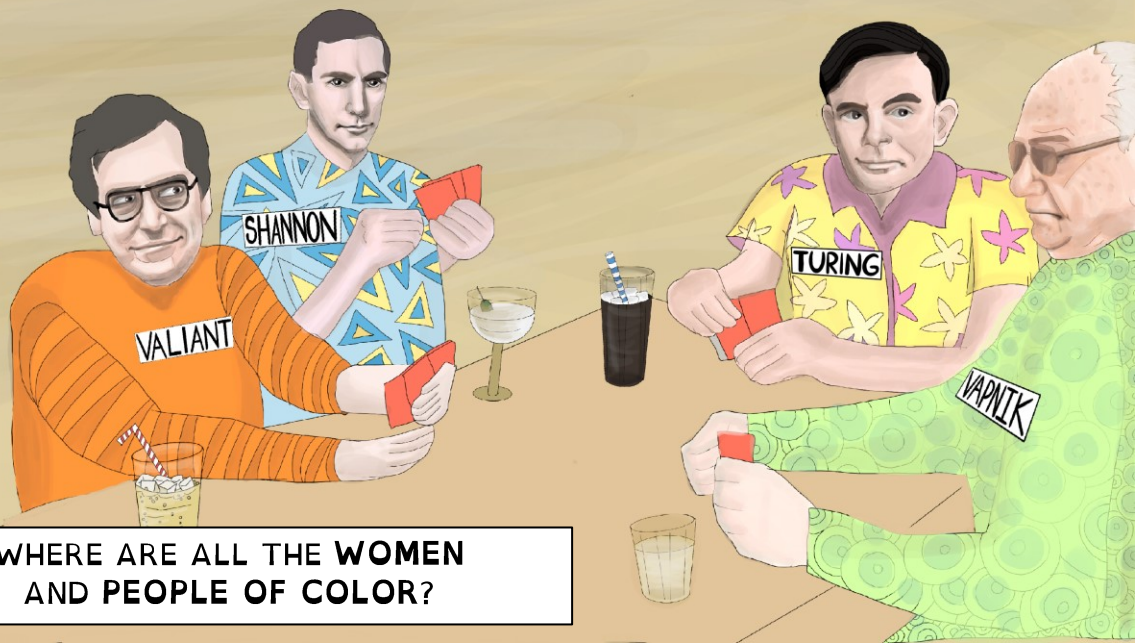
THEY WERE RIGHT. THEY DID IT.

AND NOW THEY JUST SIT BACK AND RELAX WHILE THEIR REPLICAS DO ALL THE WORK.

LOOK AT THIS GUY, HE JUST PUBLISHED A NEW PAPER, ALL WHILE SIPPING A NICE GLASS OF WINE.

I KNOW WHAT YOU'RE THINKING..

IS THIS YET ANOTHER WHITEWASHED HOLLYWOOD PRODUCTION?



WHERE ARE ALL THE **WOMEN** AND **PEOPLE OF COLOR**?

IF TECHNOLOGICAL SUPREMACY LIES AT THE SUMMIT OF THE AI MOUNTAIN THAT HUMANITY MUST SCALE AT ALL COSTS,

THEN OUR PREPARATION FOR THE CLIMB AND THE EQUIPMENT AVAILABLE TO US...

...WILL MAKE ALL THE DIFFERENCE.

BASED ON OUR CURRENT TRAJECTORY

NOT EVERYONE WILL MAKE IT.

PART I: ROCKFALL

(WHAT WORK DO WE FUND?)

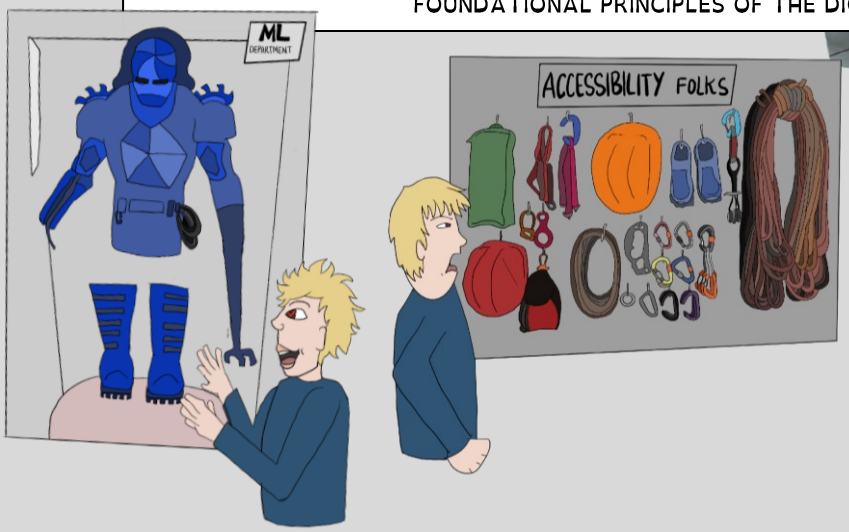
AI IS THE SHINIEST TOY ON THE BLOCK AND SO, INEVITABLY, ALL THE **MONEY-MAGPIES** HAVE COME FLOCKING.

BUDGET CUTS

WHAAA

HOWEVER, BEYOND THE USUAL SLEW OF POPULAR APPLICATIONS, SUCH AS **VISION** AND **LANGUAGE MODELING**, THE MONEY SELDOM TRICKLES DOWN.

FOR EXAMPLE, TAKE ***HUMAN-COMPUTER INTERACTION (HCI)***. THIS WORK FOCUSES ON FOUNDATIONAL PRINCIPLES OF THE DIGITAL AGE, SUCH AS **EQUITABLE ACCESS**,



AND YET IT SELDOM SEES THE KIND OF ECONOMIC BACKING OR MEDIA COVERAGE AS MACHINE LEARNING (ML) DOES.

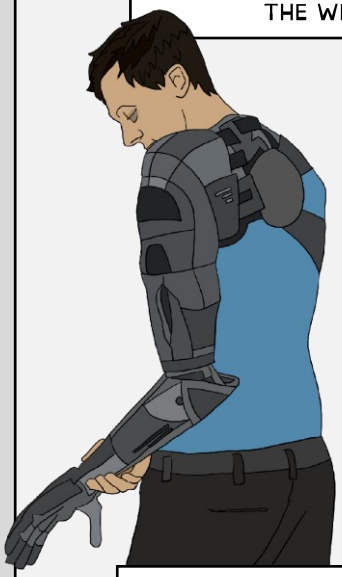
LET'S GIVE **HCI** A MOMENT IN THE **SPOTLIGHT**, SHALL WE?

DIGITAL ACCESSIBILITY

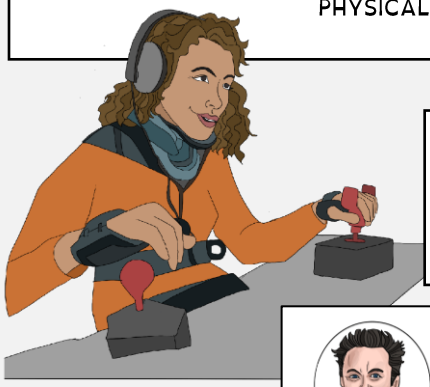
DID YOU KNOW?
15% OF THE ENTIRE POPULATION EXPERIENCE SOME FORM OF DISABILITY- VISUAL, AUDITORY, MOTOR OR COGNITIVE. (3)

"THE POWER OF THE WEB IS IN ITS UNIVERSALITY. ACCESS BY EVERYONE REGARDLESS OF DISABILITY IS AN ESSENTIAL ASPECT"
-TIM BERNERS-LEE

SO, WHAT IS **DIGITAL ACCESSIBILITY**?
THIS VOLUME IS ABOUT ML AND DATA, SO YOU'RE PROBABLY IMAGINING ROBOTIC ARMS TRAINED ON HUNDREDS OF THOUSANDS OF RUNS OF SIMULATED MOVEMENT AND CUSTOMIZED TO THE WEARER'S MEASUREMENTS AND MOTION OF ACTION.



OR HOW ABOUT A FULLY AUTOMATED, HYPER SENSITIVE ROBOTIC ARMOUR THAT SELF-LEARNS AND AUTO-NAVIGATES FOR THE PHYSICALLY DISABLED?



OR GROUND-BREAKING, HYPER-INTELLIGENT GOGGLES FOR THE BLIND, THAT COLLECT THE DISTORTED IMAGE FROM THE WEARER'S RETINAS AND RECONSTRUCT IT TO A SHARP, 10800000 PIXEL IMAGE FOR SUPERHUMAN VISION?

MAYBE, IF ELON MUSK DECIDED TO GET INTO THE ACCESSIBILITY GAME...

IN OUR REALITY, DIGITAL ACCESSIBILITY IS FOCUSED ON MAKING SURE WEB PLATFORMS ARE EASILY NAVIGABLE AND USABLE BY PEOPLE WITH ANY KIND OF DISABILITY

IT IS THIS VERY WORK THAT MAKES SURE THAT THE IMAGE YOU JUST POSTED ON INSTAGRAM HAS CAPTIONS

SO THAT THE BLIND USERS OF THE PLATFORM CAN ALSO PARTAKE IN YOUR TRIUMPH OVER THAT SOURDOUGH RECIPE.

OR WHEN YOU DROP A NEW TUTORIAL VIDEO FOR ALL ONE SQUILLION OF YOUR SUBSCRIBERS TO ENJOY,

IT IS THIS WORK THAT CONVERTS YOUR VOCAL PEARLS OF WISDOM INTO TEXT FOR YOUR DEAF FOLLOWERS.

ACCESSIBILITY NEEDS TO BE A FUNDAMENTAL DESIGN PRINCIPLE FOR BUILDING WEBSITES AND SOFTWARE,

BUT IN OUR QUEST FOR OPTOPIA, IT IS USUALLY OVERLOOKED.

WITHOUT **A11IES** (4), THE DEMOGRAPHIC THAT WAS HOLDING ON TO THE ACCESSIBILITY ROPE IS NOW CUT OFF.

LET'S GET RID OF THE **MAGPIE MENTALITY**?

FOR YOUR NEXT FUN DATA SCIENCE PROJECT, INSTEAD OF SOME COMMUNITY-OVERFITTED IMAGE RECOGNITION CHALLENGE, MAYBE CHOOSE AN **OPEN PROBLEM IN DIGITAL ACCESSIBILITY**, SUCH AS AUTOMATIC VIDEO CAPTIONING.
THEN HOPEFULLY ONE DAY THERE WILL BE **"NO MORE CRAPTIONS"** (5)

PART 2: GHOSTS IN THE SHELL

(WHO ARE WE BUILDING MODELS FOR?)

WE HAVEN'T YET FIGURED OUT HOW TO MAKE EXISTING DIGITAL PLATFORMS ACCESSIBLE TO EVERYONE, YET WE'RE ALREADY JUMPING TO FORGE A NEW "INTELLIGENT" CLASS OF WEB APPLICATIONS.

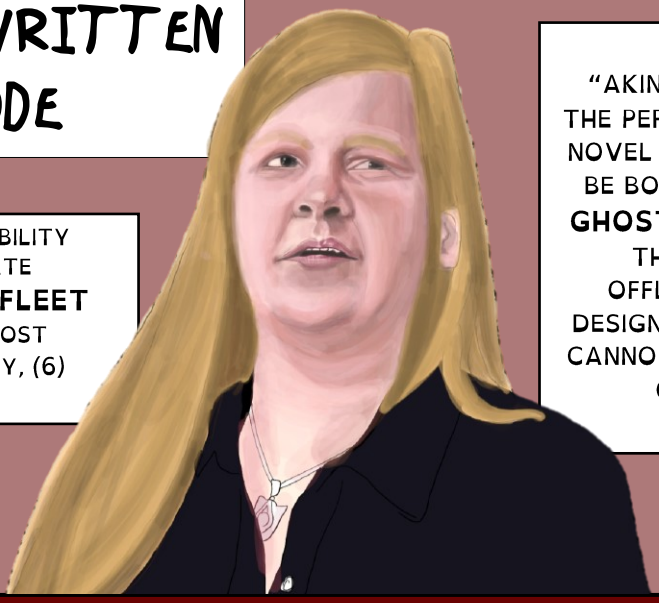
WE'RE SO CAUGHT UP IN THE **"HOW"** (USING ML/AI/DL/DS !!!) THAT WE FORGET TO ASK, **"FOR WHOM''?**

WHEN PLATFORMS ARE NOT DESIGNED FOR EVERYONE, THEY GIVE OFF THE STENCH OF "**ENCODED INHOSPITALITY**" (6).

SEEMINGLY INNOCUOUS THINGS SUCH AS **POP-UPS** AND **EXPIRING FORMS** ON WEBSITES COMPLETELY HIJACK THE ONLINE EXPERIENCE OF USERS WITH DISABILITIES WHO RELY ON SCREEN READERS.

GHOSTWRITTEN CODE

AS ACCESSIBILITY ADVOCATE **CHANCEY FLEET** PUTS IT MOST ELOQUENTLY, (6)



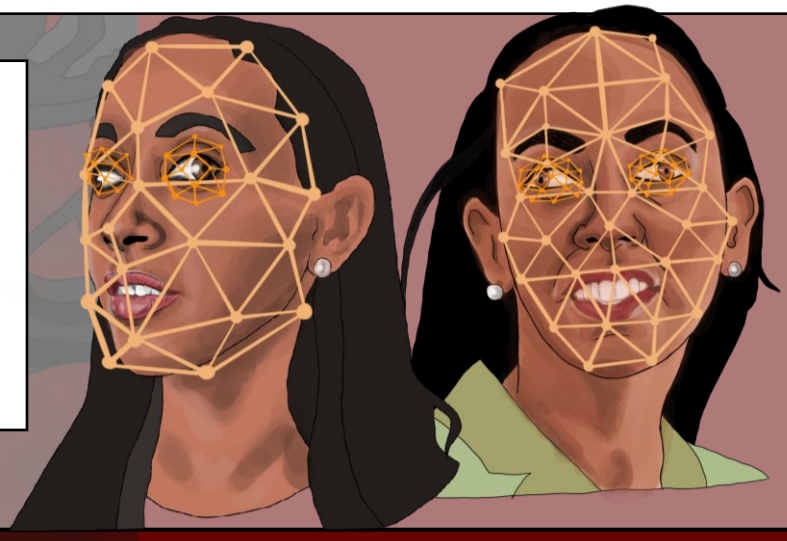
"AKIN TO HOW A **GHOSTWRITER** IS THE PERSON WHO IS PAID TO COMPOSE A NOVEL THAT SOMEONE ELSE COULD NOT BE BOTHERED TO WRITE THEMSELVES, **GHOSTWRITTEN CODE** IS SOFTWARE THAT THE ORGANIZATION HAS OFFLOADED ON PROGRAMMERS TO DESIGN FOR USERS THAT THE COMPANY CANNOT BE BOTHERED TO ENGAGE WITH OR EMPLOY THEMSELVES. "

THESE GHOSTS ARE MAKING THEIR WAY INTO DATA-DRIVEN PRODUCTS AS WELL.

TAKE THE INFAMOUS FACIAL RECOGNITION SOFTWARE THAT HAS BEEN ALL OVER THE NEWS RECENTLY. RACIAL INJUSTICES ARE PROBLEMATIC ENOUGH, BUT HAVE YOU CONSIDERED HOW THESE MODELS DISCRIMINATE AGAINST BLACK DISABLED PEOPLE?

AS DISABILITY RIGHTS ADVOCATE **HABEN GIRMA** EXPLAINS (7),

"MY EYES MOVE INVOLUNTARILY, EACH ONE SWINGING TO ITS OWN MUSIC. THEY'VE DANCED THIS WAY FOR AS LONG AS I CAN REMEMBER."



HOW WELL DO YOU THINK **FACIAL RECOGNITION** WOULD PERFORM ON **BLIND BLACK PEOPLE**?

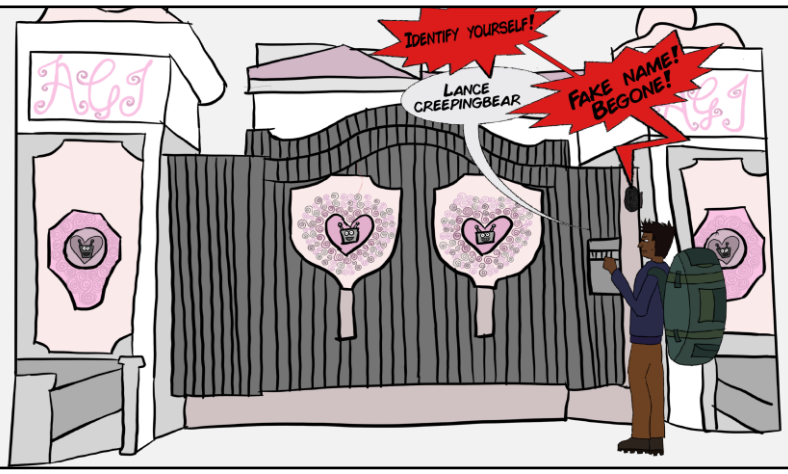
HAVING BEEN TRAINED ON THE FACIAL DYNAMICS OF SIGHTED WHITE PEOPLE, FACIAL RECOGNITION TECHNOLOGY PEDDLES AN ABLEIST AND RACIST NARRATIVE.

THE ATYPICAL, ASYMMETRIC MECHANISMS OF THE EYES OF SOME BLIND PEOPLE ARE PERCEIVED AS ABNORMAL, ANOMALOUS AND THREATENING BY THESE SYSTEMS.

HOW IS IT THAT WE CAN **FORGET** TO CONSIDER **ENTIRE DEMOGRAPHICS** WHILE DESIGNING PRODUCTS?

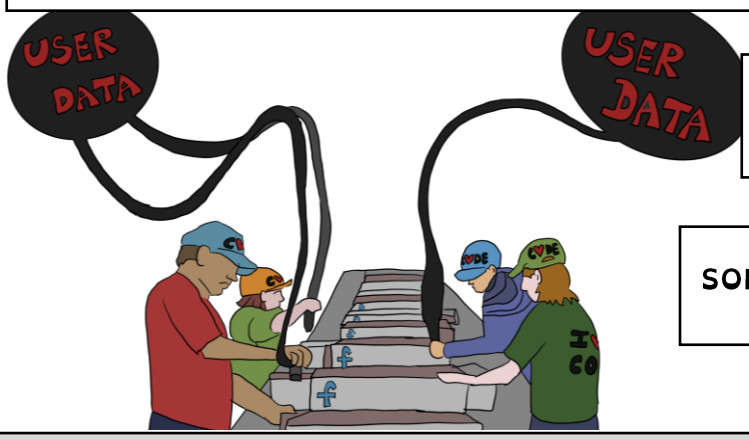
TAKE FACEBOOK'S **"REAL NAME" POLICY** THAT INDISCRIMINATELY TARGETED NATIVE AMERICANS (8)

THE LARGEST SOCIAL NETWORK IN THE WORLD SURE OVERLOOKED THE **CULTURAL AND LINGUISTIC DIFFERENCES** IN NAMES ACROSS THE GLOBE



AND ENDED UP DEPLOYING A BIGOTED ALGORITHM THAT BLOCKED USERS WHOSE NAMES DID NOT CONFORM WITH THE WESTERN ARCHETYPE OF NAMES

IN ADDITION TO COMPLETELY OVERLOOKING **WHO** WE ARE BUILDING A PRODUCT FOR, HAVE WE ALTOGETHER DONE AWAY WITH THE QUESTION OF WHETHER A CERTAIN PRODUCT ***SHOULD*** EVEN BE BUILT?



SURE, YOU HAVE SEVERAL HUNDRED TERABYTES OF USER DATA AND A FLEET OF ENGINEERS WAITING TO DIP THEIR HANDS INTO THE ML PIE,

BUT, IS YOUR PRODUCT A **SOLUTION** TO AN ACTUAL PROBLEM OR SIMPLY **SOLUTIONISM**

PART 3: THE POISONING

(WHAT PROBLEMS ARE WE TRYING TO SOLVE?)

TECHNOLOGY IS SUPPOSED TO DRIVE INNOVATION AND MOVE US TOWARDS A MORE SOPHISTICATED AND ADVANCED FUTURE, RIGHT?

AND SO WHEN THE NEW FLAVOR OF TECHNOLOGICAL ADVANCEMENT COMES TO MARKET, WHAT ELSE MUST WE DO BUT EAGERLY LAP IT UP?

WELL, IF THERE'S ANY MENTION OF "INTELLIGENCE" ON THE PRODUCT BEING HANDED TO YOU...

YOU MIGHT NOT WANT TO DRINK THAT!

IT'S SNAKE OIL!

-**ARVIND NARAYANAN**
PROFESSOR OF COMPUTER SCIENCE AT PRINCETON UNIVERSITY (9)

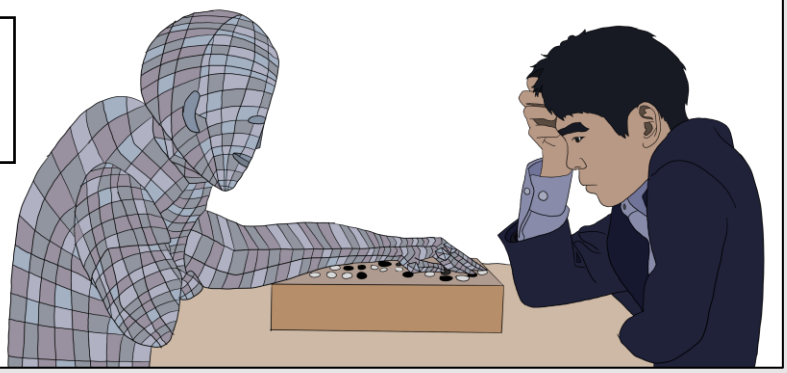
WHAT IS AI-SNAKE OIL?

SNAKE OIL IS THE MYSTICAL SUBSTANCE THAT IS CREATED BY TAKING EQUAL PARTS MEDIA HYPE AND PUBLIC MISINFORMATION AND STIRRING THEM INTO A POTION, WITH AN IRRESISTIBLE LABEL THAT SCREAMS "DATA" AND "INTELLIGENCE"

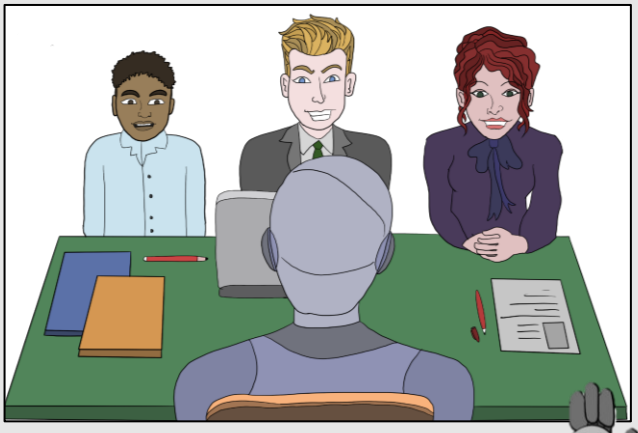
... AND AFTER YEARS OF EXPERIMENTATION, THE TECH INDUSTRY HAS FINALLY PERFECTED THE RECIPE!

DEVELOPMENTS SUCH AS **ALPHA-GO** (THE GO PLAYING AI) AND **SHAZAM** (THE MUSIC RECOGNITION APP) ARE INDICATIVE OF GENUINE SCIENTIFIC PROGRESS AND DO DEMONSTRABLY MORE GOOD THAN HARM.

WHY? BECAUSE THE RULES OF GO DON'T CHANGE WHETHER THE PLAYER IS **MALE/FEMALE, BLACK/WHITE, RICH/POOR!**

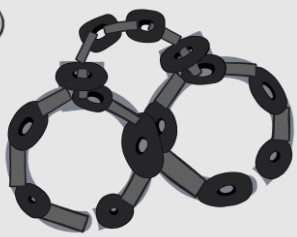


PERCEPTION TASKS, SUCH AS **FACIAL RECOGNITION**, THAT ARE INTERTWINED WITH THE **SOCIAL, POLITICAL AND CULTURAL UNDERPINNINGS** OF THE DATA ON WHICH THEY WERE TRAINED, ARE FAR MORE TOXIC.



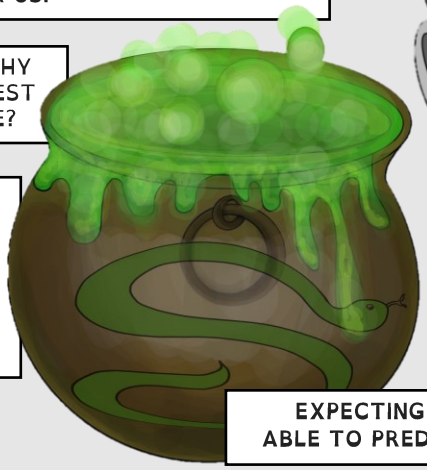
THINGS START TO GET REALLY TOXIC IN SETTINGS SUCH AS **HIRING, MODERATION OF HATE SPEECH OR ALLOCATION OF GRADES** (10), WHEN WE TRY TO IMPOSE OBJECTIVITY (FIT A MATHEMATICAL FUNCTION ONTO THE DATA) ON **HUMAN JUDGMENT**, WHICH IS INHERENTLY SUBJECTIVE

WE GET REALLY CREATIVE WITH WHAT WE THINK WE CAN ACHIEVE WITH TECHNOLOGY WHEN WE START **PREDICTING SOCIAL OUTCOMES** USING ALGORITHMS, SUCH AS **COMPAS FOR CRIMINAL SENTENCING.** (11)



WE LOOK AROUND AND SEE THE HARDEST PROBLEMS KNOWN TO US AND DECIDE THAT, SINCE WE CANNOT SOLVE THEM, WE MUST INSTEAD GET A MACHINE TO DO IT FOR US.

BUT DO YOU KNOW WHY THESE ARE THE HARDEST PROBLEMS TO SOLVE?



BECAUSE THESE ARE SYSTEMIC ISSUES THAT HAVE BEEN SLOWLY STEWING FOR CENTURES OVER

WITH A DASH OF HISTORICAL CONTEXT, A SPRINKLE OF CULTURE AND A GENEROUS HEAPING OF RACE, GENDER AND CLASS POLITICS

ALL COMPOUNDING INTO A COMPLEX BROTH OF ENTROPY;

EXPECTING A MACHINE TO TAKE ONE WHIFF OF THIS STEW AND BE ABLE TO PREDICT THE FUTURE IS JUST FUNDAMENTALLY DUBIOUS.

UNDERNEATH ALL THE BELLS AND WHISTLES OF THIS LARGER THAN LIFE SPECTACLE IS A DANGEROUSLY HIGH-RISK GAME THAT WE DON'T EVEN KNOW WE'RE A PART OF!

WELCOME TO THE

AI CIRCUS!

THE **BALANCING ACT** BETWEEN MAKING A MODEL SIMULTANEOUSLY **ACCURATE, FAIR AND FEASIBLE** IS REALLY A SPECTACLE FOR ALL TO SEE!

TAKE AI FOR HIRING.
IF A COMPANY INDULGES IN DISCRIMINATORY HIRING PRACTICES FOR YEARS ON END,

PREDICTIVE MODELS THAT AUTOMATE SUCH DECISIONS WILL FAVOUR THE SAME PEDIGREE OF APPLICANTS THAT WERE HISTORICALLY HIRED

AN EXTREMELY "**ACCURATE**" ALGORITHM WILL FAITHFULLY REPLICATE THE DISCRIMINATORY BEHAVIOR OF ITS HUMAN TRAINERS.

COUNTERACTING DATA BIAS BY ENFORCING A NOTION OF **"FAIRNESS"** IN PREDICTION COMES AT THE COST OF MODEL **"ACCURACY"** – WHEN ACCURACY IS MEASURED ON THE BIASED TRAINING DATA

WHY? BECAUSE AN ALGRORITHM THAT HAS ACCURATELY LEARNED FROM BIASED DATA WILL ALSO BE BIASED, BY CONSTRUCTION

THIS PROBLEM GETS HARDER BECAUSE ML MODELS ARE **OPAQUE**. WE HAVE LIMITED UNDERSTANDING ABOUT <u>HOW</u> A PREDICTION WAS MADE.

SOMETIMES THE DATA IS SO TERRIBLY BIASED THAT IN ORDER TO DELIVER FAIRER OUTCOMES, WE NEED TO GO BACK AND COLLECT A WHOLE NEW SAMPLE OF DATA.

THIS MIGHT NOT BE FEASIBLE IN ALL CIRCUMSTANCES AND SO COMPANIES HAVE TO TAKE A STAND ON WHICH METRIC THEY VALUE MOST.
FEASIBILITY OR FAIRNESS?

DO THEY PUSH FOR A FAIR BUT EXPENSIVE ALGORITHM OR SETTLE FOR THE "MOST FAIR" ALGORITHM THAT THEY CAN AFFORD AT THE LEAST COST?

THEN THERE'S THE

PYRAMID OF ML SCHOLARSHIP.

SOTA-CHASING

AT THE VERY TOP OF THE PYRAMID SITS (PRECARIOUSLY) OUR M.O. OF **SOTA CHASING**.

SOTA OR **STATE OF THE ART** REFERS TO THE LATEST REPORTED METRIC ON A CERTAIN TASK

SOTA-CHASING IS ABOUT OUTPERFORMING THE COMPETITION ON THAT BENCHMARK, EVEN **MARGINALLY**

DATA + COMPUTE

STATISTICAL LEARNING THEORY

WE CREATE A BENCHMARK DATASET AND THEN DECLARE A METRIC, USUALLY ACCURACY, BY WHICH WE WILL MEASURE SUCCESS.

NOW PROJECT AFTER PROJECT AND PAPER AFTER PAPER SETS OUT TO ATTAIN THAT 0.01 INCREASE IN ACCURACY IN ORDER TO BE DEEMED "PUBLISHABLE"

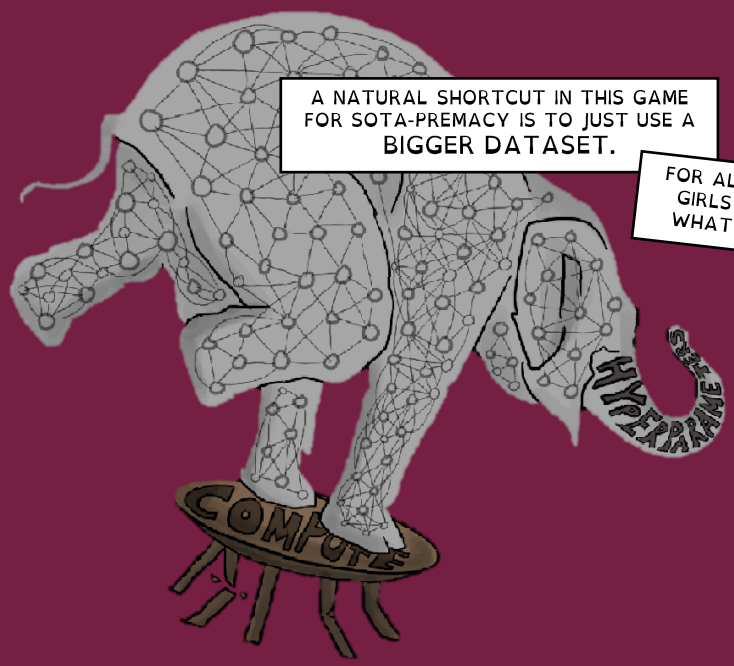
ALL THE WHILE TAKING IT FOR GRANTED THAT ACCURACY IS EVEN THE RIGHT METRIC TO MEASURE PROGRESS AGAINST.

WHY IS THIS A PROBLEM?

BECAUSE SOTA-CHASING ASSUMES THAT THE BENCHMARK IS WORTH CHASING! THAT THE DATASET IS REPRESENTATIVE OF THE POPULATION. AND THAT MARGINAL ACCURACY IMPROVEMENT MAKES A DIFFERENCE

A NATURAL SHORTCUT IN THIS GAME FOR SOTA-PREMACY IS TO JUST USE A BIGGER DATASET.

FOR ALL YOU GPT3 (12) FAN BOYS AND GIRLS OUT THERE, THIS IS EXACTLY WHAT YOU'VE BEEN RAVING ABOUT:



THE SHEER, UNFETTERED ACCESS TO A GARGANTUAN DATASET AND COMPUTE (OR MOOLAH TO PAY FOR COMPUTE)

TO CREATE MODELS THAT BEAT THE STATE OF THE ART

AND GIVE THE ILLUSION OF SCIENTIFIC PROGRESS

SURE, THERE ARE THOSE FOLKS IN THE COMMUNITY WHO ARE THINKING DEEPLY ABOUT PROBLEM FORMULATION, REAL WORLD IMPACT AND SCIENTIFIC RIGOR. UNFORTUNATELY, DEEP, THOUGHTFUL WORK OF THIS KIND IS JUST NOT GLAMOROUS

...AND SO, WHEN THE CURTAIN FALLS, IT ISN'T THESE RESEARCHERS YOU ARE APPLAUDING.

HOW COME THESE FOLKS NEVER TAKE CENTER STAGE?
WELL, IT'S PARTLY BECAUSE, LIKE IN EVERY OTHER DOMAIN, THE RICH JUST KEEP GETTING RICHER.

THE SET OF RESEARCHERS WHO DEBUNK SOCIETAL HARMS OF TECHNOLOGY ARE LIKELY TO BE FROM THE SAME DEMOGRAPHIC THAT WILL BE MOST DEEPLY AFFECTED BY THOSE VERY HARMS.

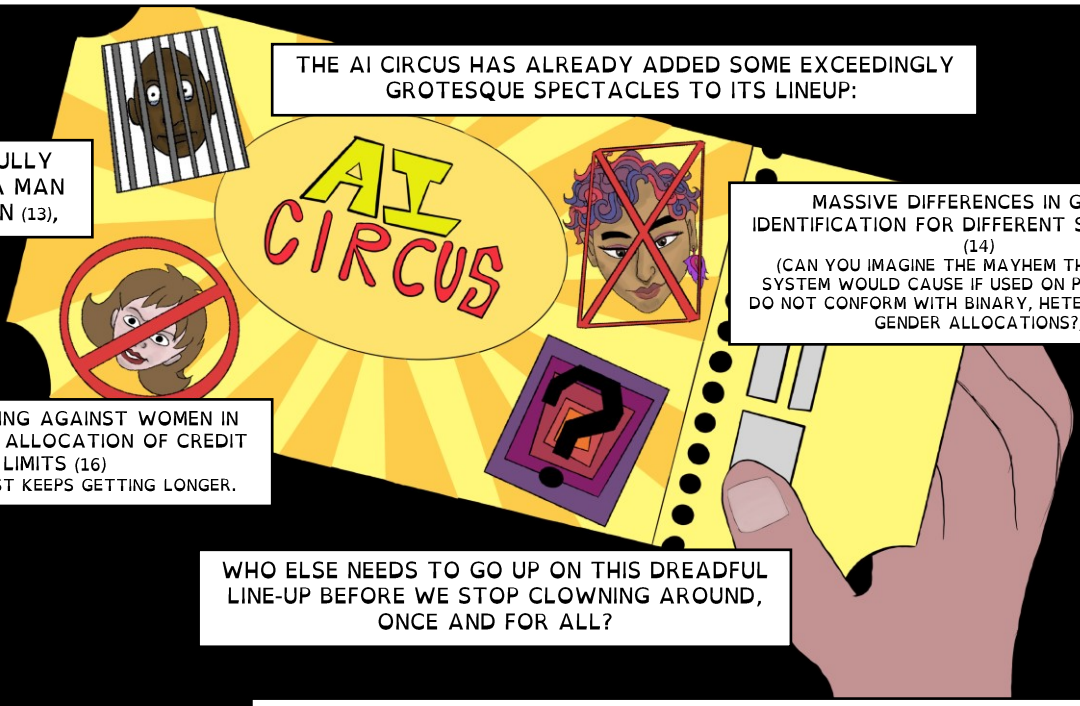
AND THIS IS NEVER THE MAJORITY.

IF OUR SCHOLARSHIP IS A REFLECTION OF OUR IDEAS, THEN WE CANNOT AFFORD TO CENSOR OR COMPLETELY ERASE THE VOICES OF ENTIRE DEMOGRAPHICS.

IF OUR PRODUCTS ARE A REFLECTION OF THE PROBLEMS THAT WE ARE TRYING TO SOLVE, THEN WE CANNOT BUILD SOLUTIONS THAT HELP ONE STRATUM AND CAUSE EXTENSIVE DAMAGE TO ANOTHER.

THE AI CIRCUS HAS ALREADY ADDED SOME EXCEEDINGLY GROTESQUE SPECTACLES TO ITS LINEUP:

WRONGFULLY SENDING A MAN TO PRISON (13),



MASSIVE DIFFERENCES IN GENDER IDENTIFICATION FOR DIFFERENT SKIN COLORS (14)
(CAN YOU IMAGINE THE MAYHEM THAT SUCH A SYSTEM WOULD CAUSE IF USED ON PERSONS WHO DO NOT CONFORM WITH BINARY, HETERONORMATIVE GENDER ALLOCATIONS?)

DISCRIMINATING AGAINST WOMEN IN HIRING (15), IN ALLOCATION OF CREDIT LIMITS (16)
...THE LIST JUST KEEPS GETTING LONGER.

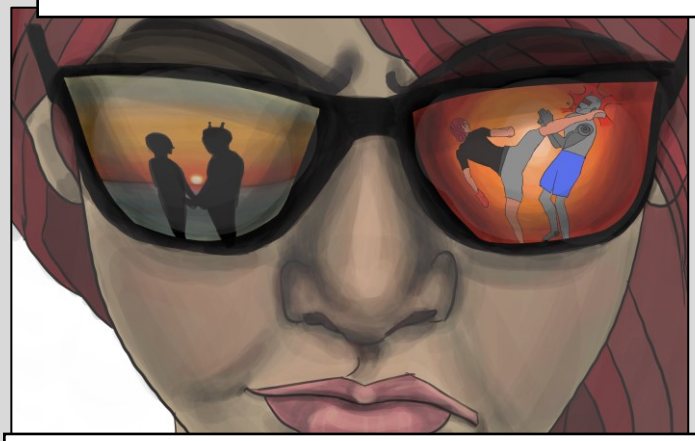
WHO ELSE NEEDS TO GO UP ON THIS DREADFUL LINE-UP BEFORE WE STOP CLOWNING AROUND, ONCE AND FOR ALL?

BEFORE YOU REACH FOR YOUR SMARTPHONE TO GET ON TWITTER TO RAGE AGAINST THE AI MACHINE OR JOIN THE RANKS OF THE TECHNO BASHERS, STOP AND LOOK AROUND

ALL AROUND ME ARE FAMILIAR FACES
CANCEL TECH
WAY OF THE FUTURE
LAWS
WORN OUT PLACES, WORN OUT FACES
REGULATION = ARMAGEDDON
BRIGHT AND EARLY FOR THEIR DAILY RACES
GOING NOWHERE, GOING NOWHERE
IT'S A VERY, VERY, MAD WORLD

IT REALLY IS A MAD WORLD. AND IT'S DRIVING US PARTICULARLY CRAZY BECAUSE WE'VE BECOME SO USED TO SEEING THE WORLD IN EXTREMES.

YOU CAN EITHER BE A **TECHNO-BASHER** OR A **TECH-OPTIMIST** AND IF YOU ARE ONE YOU **CANNOT** AND **SHALL NOT** SYMPATHIZE WITH THE OTHER SIDE.



GIVE AI THE REIGNS TO RUN THE ENTIRE WORLD OR PILE IT ALL UP AND THROW IT ALL OUT.

WE'VE BECOME SO USED TO **'HULKING-OUT'** AT THE FIRST SIGN OF DISAGREEMENT ON SOCIAL MEDIA,

THAT THE ENTIRE DISCOURSE AROUND TECH, AND AI IN PARTICULAR HAS BEEN COMPLETELY STRIPPED OF SUBTLETY.

IT'S 2020.

HOW IS IT THAT WE CAN APPRECIATE A COMEDIC TAKE ON HITLER AND THE NAZI YOUTH CAMPS (17), WITHOUT GETTING OUR FEELINGS HURT...

...BUT WE CAN'T HAVE ONE DISCUSSION ABOUT BIAS IN THE DATA WITHOUT IT IMMEDIATELY DEVOLVING INTO BLOWS.

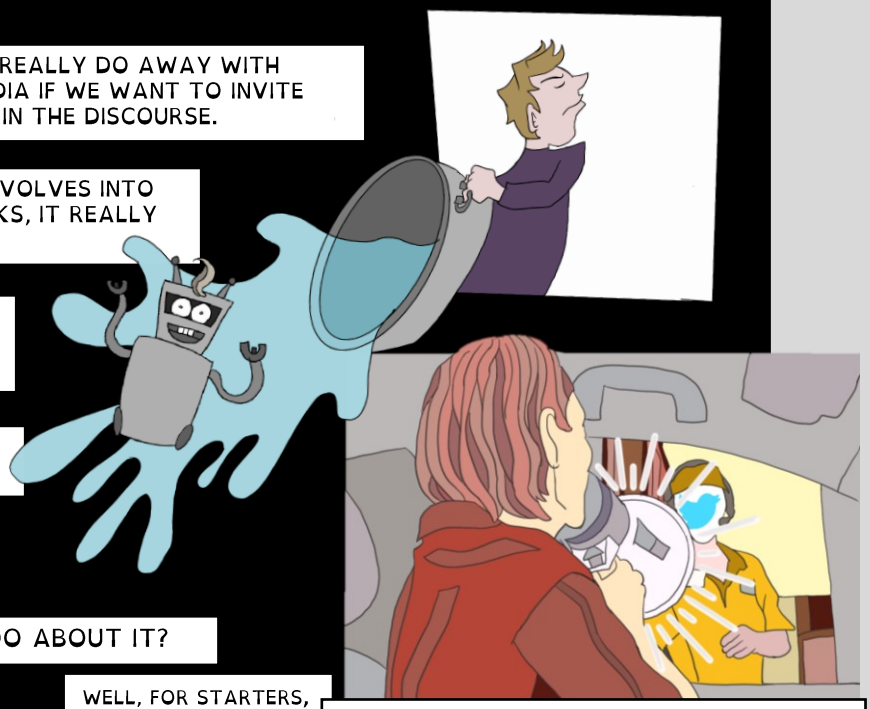
MAYBE WE NEED TO STOP REACTING TO EVERYTHING WE READ AND INSTEAD TAKE A MOMENT TO RE-READ, THINK DEEPLY AND THEN RESPOND.

BECAUSE THE TRUTH IS, WE CAN'T REALLY DO AWAY WITH THESE DISCUSSIONS ON SOCIAL MEDIA IF WE WANT TO INVITE THE GENERAL PUBLIC TO PARTAKE IN THE DISCOURSE.

BUT WHEN A DISCUSSION QUICKLY DEVOLVES INTO GASLIGHTING AND PERSONAL ATTACKS, IT REALLY DOESN'T BENEFIT ANYONE.

THE EXTANT CELEBRITY CULTURE AND INTERNET TROLLING THAT SHROUDS SCIENTIFIC DISCUSSIONS NEEDS TO GO!

OR ELSE WE JUST END UP THROWING THE BABY OUT WITH THE BATHWATER



SO, WHAT DO WE DO ABOUT IT?

WELL, FOR STARTERS,

CAN WE GET SOME ***NUANCE*** WITH OUR DISCUSSION MEAL, PLEASE!?!

HERE IS A MORE NUANCED TAKE ON WHETHER AI LEADS TO A **UTOPIA** OR A ***DYSTOPIA:***

FOR STARTERS, **THERE IS RARELY AN OBJECTIVE TRUTH**! MORE OFTEN THAN NOT, THE EFFICACY OF A MODEL DEPENDS ON THE **CONTEXT** FOR WHICH IT WAS DESIGNED

THE "**GROUND TRUTH**" THAT WE PRETEND EXISTS, AND AGAINST WHICH WE MEASURE MODEL ACCURACY, IS JUST THE **CLOTHES** THAT THE **ML EMPEROR** IS **NOT** WEARING!

THE ENGINEERING MINDSET IS TO TAKE THE CLASS LABELS AS GOSPEL AND BLINDLY TRY TO OPTIMIZE FOR THEM.

BUT CLASS LABELS ARE JUST **PROXIES** FOR **UNDERLYING SOCIAL PHENOMENA** AND NO AMOUNT OF **MATHEMATICAL FORMALIZATION** WILL TURN **SOCIAL CONSTRUCTS** INTO **OBJECTIVE TRUTHS.**

THE REALITY IS THAT **ALL** MODELS ARE **WRONG**. **SOME** MODELS ARE **USEFUL**!

IN THIS ART GALLERY, EACH PAINTING DEPICTS AN **APPLE.** BUT ONLY ONE OF THEM IS POTENTIALLY USEFUL AS A **REAL-LIFE APPLE DETECTOR**

WE OFTEN FIND IT HARD TO JUDGE WHICH MODEL IS MOST USEFUL, BECAUSE THAT REQUIRES DEEP **DOMAIN EXPERTISE.**

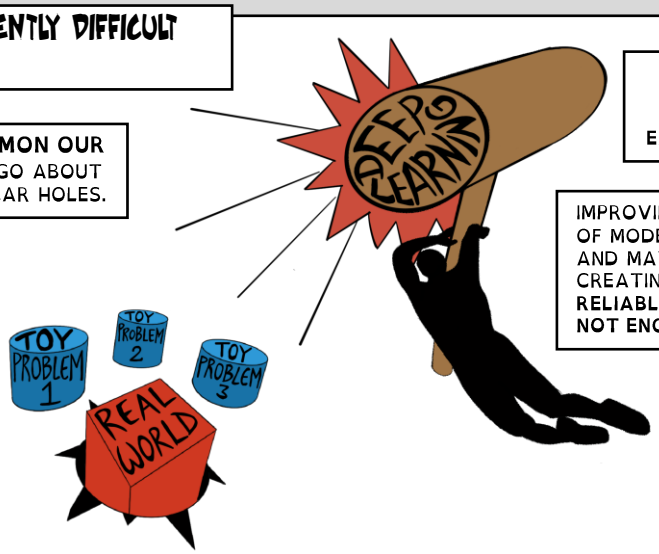
WE HAVE BEEN **DANGEROUSLY CONFLATING** EXPERTISE IN TRAINING AND DEPLOYING A MODEL WITH DOMAIN EXPERTISE.

INSTEAD WE SHOULD **ACKNOWLEDGE THE LIMITATION OF OUR EXPERTISE** AS SCIENTISTS AND ENGINEERS AND **INVITE THE TRUE DOMAIN EXPERTS** TO COME TO THE TABLE.

SOME **CONTEXTS** ARE ***INHERENTLY DIFFICULT*** TO BUILD FOR.

WE HAVE THE TENDENCY **TO SUMMON OUR DEEP LEARNING HAMMER** AND GO ABOUT NAILING SQUARE PEGS INTO CIRCULAR HOLES.

UNFORTUNATELY, THE MOST PROMISING RESULTS THAT YOU READ ABOUT WERE OBTAINED ON **TOY PROBLEMS** WITHIN **EXPERIMENTAL SET-UPS** AND ARE NOT DESIGNED TO SCALE TO THE REAL WORLD.



THE WORLD IS A COMPLICATED AND MESSY PLACE AND THE **LIMITED PERFORMANCE OF OUR EXISTING MODELS** REFLECTS THAT.

IMPROVING **GENERALIZATION ABILITY** OF MODELS IS A HOT AREA OF RESEARCH AND MAYBE WE'LL GET AROUND TO CREATING MODELS THAT CAN **PERFORM RELIABLY** IN **CONTEXTS THAT THEY DID NOT ENCOUNTER DURING TRAINING.**

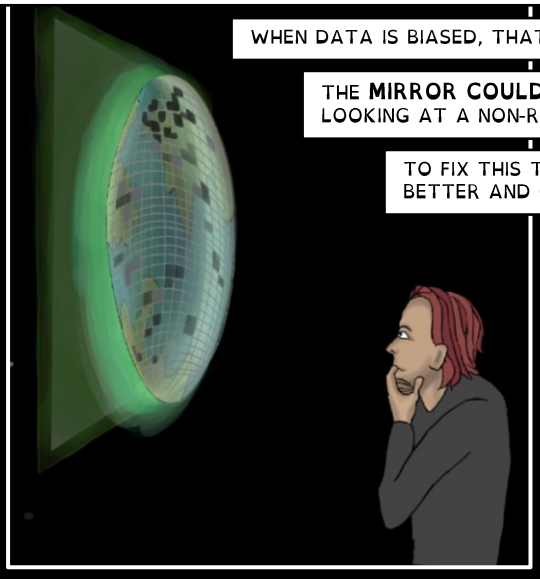
BUT WE AREN'T THERE YET.

THE OVERWHELMING MAJORITY OF PROBLEMS THAT PLAGUE AI TODAY ARE NOT BECAUSE OF *JUST* THE DATA OR *JUST* THE ALGORITHM IN ITSELF

BUT BECAUSE OF ONE CRITICAL ***CONFOUNDING FACTOR*** THAT WE KEEP OVERLOOKING:

THE WORLD

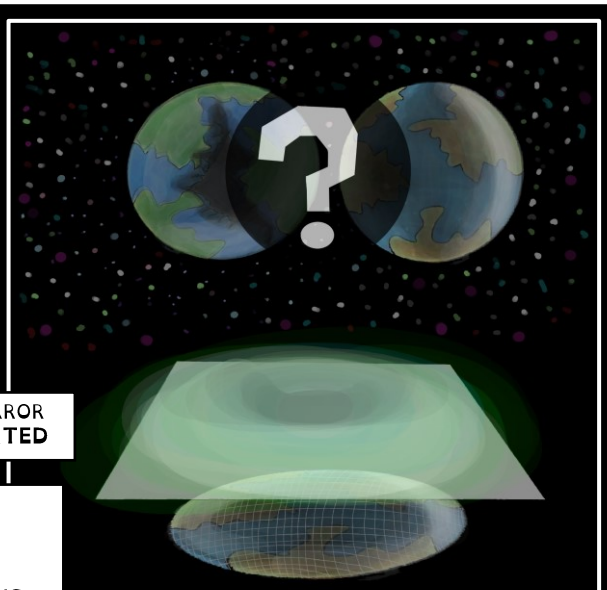
DATA IS A MIRROR REFLECTION OF THE WORLD (18)



WHEN DATA IS BIASED, THAT REFLECTION IS DISTORTED. THERE ARE SEVERAL EXPLANATIONS FOR THIS

THE **MIRROR COULD BE DISTORTED**: WE COULD BE COLLECTING THE WRONG DATA, OR LOOKING AT A NON-REPRESENTATIVE SAMPLE

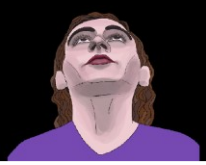
TO FIX THIS TYPE OF BIAS, WE CAN ATTEMPT FIXING THE MIRROR TO COLLECT BETTER AND CLEANER DATA



BUT THERE'S ALSO THE POSSIBILITY THAT THE MIRROR IS PERFECT AND **THE WORLD ITSELF IS DISTORTED**

WE TEND TO UNDER-APPRECIATE THIS POSSIBILITY BECAUSE WE INSTINCTIVELY COMPARE THE REFLECTION (DATA) WITH **HOW WE WANT THE WORLD TO BE,** RATHER THAN WITH **HOW IT ACTUALLY IS!**

BASED ON THE REFLECTION, AND WITHOUT KNOWLEDGE OR **ASSUMPTIONS** ABOUT THE PROPERTIES OF THE MIRROR AND OF THE WORLD IT REFLECTS, WE **CANNOT KNOW** WHETHER THE REFLECTION IS DISTORTED, AND, IF SO, FOR WHAT REASON.



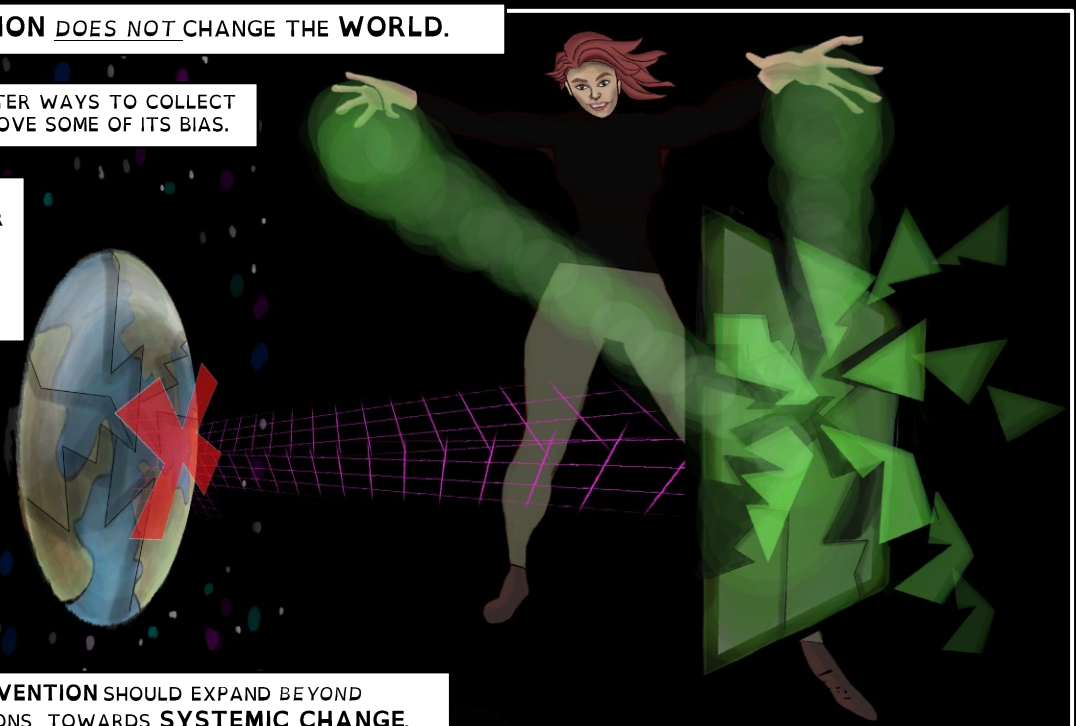
DATA ALONE **CANNOT** TELL US WHETHER IT IS A DISTORTED REFLECTION OF A PERFECT WORLD, OR A PERFECT REFLECTION OF A DISTORTED WORLD, OR WHETHER THESE **DISTORTIONS COMPOUND.**

CHANGING THE **REFLECTION** *DOES NOT* CHANGE THE **WORLD.**

WE'VE COME UP WITH BETTER WAYS TO COLLECT DATA, CLEAN IT AND REMOVE SOME OF ITS BIAS.

BUT, ALL OF THESE FIXES ARE **APPLIED ON THE MIRROR** OR ON THE **REFLECTION** AND THEY **DO NOT PROPAGATE BACK** TO CHANGE THE **WORLD.**

THE **UNDERLYING SOCIETAL INEQUITIES** THAT GIVE RISE TO DISCRIMINATORY OUTCOMES **REMAIN INTACT** IF WE **ONLY INTERVENE ON THE DATA.**



HENCE, OUR **INTERVENTION** SHOULD EXPAND *BEYOND* TECHNOLOGICAL SOLUTIONS, TOWARDS **SYSTEMIC CHANGE.**

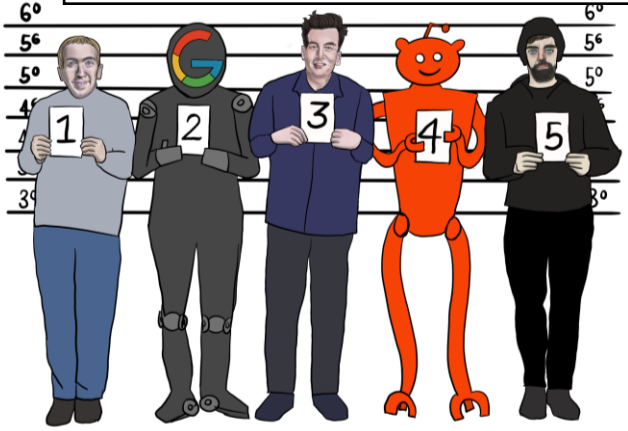
PREDICTION IS DIFFICULT, ESPECIALLY ABOUT THE ***FUTURE***! (19)

WHEN THINGS (INEVITABLY) GO WRONG, WHO IS **RESPONSIBLE**?

IT CANNOT BE THE ALGORITHM.

BUT GIVEN THE MANY STAKEHOLDERS THAT PLAY A PART IN THE CREATION AND OPERATION OF A SOFTWARE PRODUCT,

HOW DO WE DETERMINE WHICH **HUMAN** IS CULPABLE? ARE THEY ALL?



I KNOW WHAT YOU'RE THINKING...

"I SEE WHERE YOU'RE GOING WITH THIS... YOU'RE NOT SERIOUSLY GOING TO GET INTO REGULATION NOW, ARE YOU?"

WELL...TIME TO REMIND YOU OF OUR RECOMMENDED APPROACH TO THINKING ABOUT AI.

REMEMBER, **NUANCE**?!

RIGHT NOW, SILICON VALLEY WILL HAVE YOU BELIEVE THAT TECHNOLOGY NEEDS TO BE ALLOWED TO RUN FREE. REGULATION IS A CATASTROPHE OF COSMIC PROPORTIONS AND WOULD BE THE END OF THE INTERNET, AND BY EXTENSION, INNOVATION AND PROGRESS.

THE FACT OF THE MATTER IS, WE PUT OUR CHILDREN ON THE AI HYPE-BIKE AND SENT THEM OFF AT FULL SPEED.

WE WERE TOO BRASH IN OUR RAPID ADOPTION OF AI AND IT HAS LED TO SOME TERRIBLE OUTCOMES WITH VERY REAL IMPACTS ON PEOPLE'S LIVES.

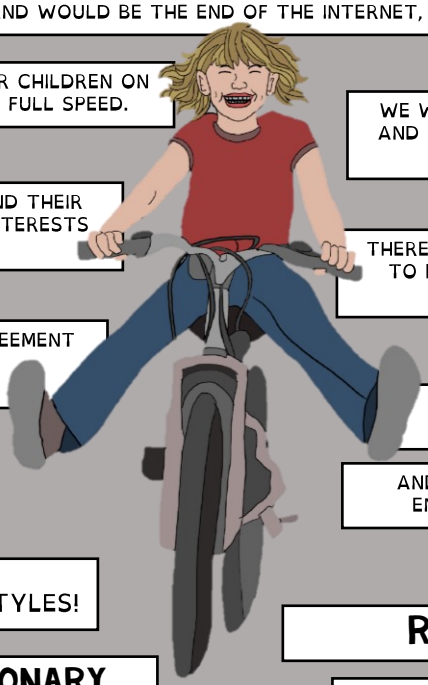
AND SO WHILE TECH COMPANIES AND THEIR CELEBRITY CEOS PROTECT THEIR INTERESTS BY BAD MOUTHING REGULATION,

THERE'S REALLY NO EXCUSE FOR THE GENERAL PUBLIC TO BUY INTO THIS NARRATIVE AND BE COMPLICIT IN THE VANDALISM OF OUR MORAL SOCIAL FIBER.

WE NEED TO COME TO AN AGREEMENT ON HOW TO GO ABOUT REGULATING TECHNOLOGY.

AND SO WE MUST START EDUCATING OURSELVES,

AND PARTAKE IN THIS LOFTY ENTERPRISE IN GOOD FAITH.



MAYBE IT'S TIME TO CONSIDER OTHER PARENTING STYLES!

PRECAUTIONARY

THINK OF THE OLD ADAGE **"IT'S BETTER TO BE SAFE THAN TO BE SORRY"**

THIS PRINCIPLE CALLS FOR CAUTION IN SITUATIONS OF **UNCERTAIN HARMS**, IE. THOSE THAT HAVE NOT BEEN SCIENTIFICALLY STUDIED YET.

A COMMON CRITICISM OF THIS APPROACH IS IT IS **"PARALYZING"** AND **"SELF-CANCELING"**, SINCE ANY NEW TECHNOLOGY IN ITS EARLY STAGES OF ADOPTION WOULD HAVE RISKS THAT CANNOT BE ACCOUNTED FOR.

V/S

RISK-BASED

UNDER THIS PARADIGM, REGULATE BASED ON **KNOWN** RISKS, AND MODEL THE **LIKELIHOOD** THAT THESE **RISKS** WILL LEAD TO **HARMS**

A PROMISING APPROACH IS **ALGORITHMIC IMPACT ASSESSMENT (AIA)** - A FRAMEWORK THAT HELPS UNDERSTAND AND REDUCE THE RISKS TO INDIVIDUALS AND COMMUNITIES

UNDER AIA, THE **LIKELIHOOD AND SEVERITY OF HARM** DETERMINES THE **LEVEL OF OVERSIGHT**. THE HIGHER THE RISK OF HARM, AND THE MORE SIGNIFICANT THE HARM ITSELF, THE MORE STRINGENT THE OVERSIGHT REQUIREMENTS. AND THE LESS AUTONOMY IS GRANTED TO THE AUTOMATED SYSTEM: A HUMAN MUST BE BROUGHT INTO THE LOOP TO TAKE RESPONSIBILITY FOR IMPACTFUL DECISIONS

AIA WILL ONLY WORK IF THE RISKS ARE KNOWN. THIS GIVES EACH AND EVERY ONE OF US THE OPPORTUNITY TO **BE A PART OF THE CHANGE**! NOW'S THE TIME TO **GET INVOLVED IN PUBLIC CONSULTATIONS**, TO MAKE YOUR CONCERNS HEARD!

IF WE WANT OUR ATTEMPTS AT REGULATION TO BE *TRULY EFFECTIVE*, WE NEED TO **RECONCILE** SOME INHERENT DISAGREEMENTS BETWEEN TECH AND LAW.

FOR STARTERS, HOW DO WE MAKE SURE THE LAW **KEEPS UP** WITH THE **RAPIDLY EVOLVING** SOCIO-TECHNOLOGICAL LANDSCAPE?

ANOTHER MAJOR PROBLEM IS ***HOW*** DO WE REGULATE?

NOTIONS SUCH AS **FAIRNESS, ACCOUNTABILITY AND INTERPRETABILITY** HAVE BECOME THE POSTER CHILDREN FOR AI POLICY. BUT THEY STILL DON'T HAVE UNIVERSALLY ACCEPTED TECHNICAL MANIFESTATIONS.

WHY? BECAUSE AMBIGUITY IN DEFINITIONS IS AN INTENTIONALLY WIELDED TOOL THAT ALLOWS FOR **INTERPRETIVE** AND **CONTEXTUAL** READINGS OF **LAW**

BUT THE VERY SAME AMBIGUITY IS CATASTROPHIC FOR **TECH**, WHICH RELIES ENTIRELY ON **MATHEMATICAL FORMALIZATIONS THAT CAN BE WRITTEN INTO CODE**

AND FOR **REGULATORS** WHO NEED **PRECISE DEFINITIONS** TO BUILD RULES AND POLICIES

TO COME UP WITH GOOD DEFINITIONS, WE NEED EXAMPLES OF SYSTEMS THAT ARE USED **TODAY**!

TAKE THE **NYC AUTOMATED DECISION SYSTEMS (ADS) TASK FORCE**, THE FIRST OF ITS KIND IN THE U.S., ENVISIONED TO BE THE BEACON FOR **TRANSPARENCY** AND **EXPERT INSIGHT** INTO THE USE OF ALGORITHMS TO AID DECISION-MAKING BY CITY AGENCIES. (20)

BUT THEY DIDN'T GET VERY FAR.

A **GOOD DEFINITION** WAS LACKING, AS WERE **EXAMPLES**.

WHAT IS AN ADS?

A CALCULATOR IS NOT AN ADS. BUT A SYSTEM THAT COLLECTS DATA, BUILDS A MODEL, AND THEN ENACTS POLICY THAT IMPACTS PEOPLE'S LIVES-ALLOCATES SCHOOL BUDGETS, OR OFFERS HOMELESSNESS ASSISTANCE, OR MATCHES STUDENTS WITH SPOTS IN HIGH SCHOOLS-CERTAINLY IS.

WITH ALL OF THIS IN MIND, LET'S REVISIT THAT QUEST OF HUMANITY FOR **OPTOPIA.**

IF WE DISCARD ENTIRE SOCIETIES AND DEMOGRAPHICS ON THE WAY, AND COMPLETELY OVERLOOK SOCIETAL PROBLEMS THAT RENDER ALGORITHMIC INTERVENTIONS FUTILE, IS THE TREK STILL WORTH PURSUING?

MAYBE INSTEAD OF A POWER TRIP IN THE NAME OF A TECHNOLOGICAL MISSION (WHEN DID WE ALL AGREE THAT HUMAN INTELLIGENCE IS WORTH REPLICATING?), WE SHOULD FOCUS ON HARNESSING THE POWER OF LEARNING TECHNOLOGIES TO POSITIVELY IMPACT PEOPLE?

AND NOT ONE, AFFLUENT, HIGHLY INFLUENTIAL DEMOGRAPHIC OF PERSONS, BUT TRULY ALL PERSONS, OF ALL SOCIAL STRATA, CLASSES, GENDERS AND RACES.

MAYBE WHAT WE NEED INSTEAD
IS TO **GROUND** THE DESIGN OF ***AI SYSTEMS*** IN **PEOPLE**

USING THE DATA ***OF*** THE PEOPLE,

COLLECTED AND DEPLOYED WITH AN EQUITABLE METHODOLOGY AS DETERMINED ***BY*** THE PEOPLE,

TO CREATE TECHNOLOGY THAT IS BENEFICIAL ***FOR*** THE PEOPLE.

FIN.

ABOUT

Falaah is a Scientist/Engineer by training and an Artist by nature, chasing a passion for building Robust and Ethical ML all the way from industry to academia. In the face of having to incessantly remind everyone around her about the limitations of current ML capabilities, Falaah started "**machinelearnist comics**" - online Scientific Comics about the AI Landscape.

Julia is an Assistant Professor of Computer Science and Engineering and of Data Science at NYU. She is passionate about Responsible Data Science and leads the "***Data, Responsibly***" project, the latest offering of which is the inimitable, interdisciplinary course on **Responsible Data Science**.

With the *undecipherable alchemy* that is grad-school admissions, the **Cosmos** brought these two creative minds together and thus was born: **Data, Responsibly Comics**!

Whether you're a **Student**, unsure about where to get started in the sea of ML scholarship; or an **Educator**, looking for a fun new pedagogical instrument for your students; or a **Practitioner**, looking for some relatable content about all the idiosyncrasies of the current AI landscape; or just a good ol' John/Jane Doe who likes to read comics and is intrigued by the prospect of a long form scientific volume, Data, Responsibly Comics are for you!

JULIA STOYANOVICH
@stoyanoj
Co-Creator, Writer

FALAAH ARIF KHAN
@FalaahArifKhan
Co-Creator, Writer, Artist, Cover Artist

REFERENCES :

[1] https://mrtz.org/gradientina.html#/

[2] http://www.hutchinsweb.me.uk/MTNI-11-1995.pdf .

[3] https://www.who.int/disabilities/world_report/2011/report/en/

[4] https://www.a11yproject.com/

[5] http://nomorecraptions.com/

[6] https://datasociety.net/library/dark-patterns-in-accessibility-tech/

[7] https://twitter.com/habengirma/status/1278035954628915200

[8] https://en.wikipedia.org/wiki/Facebook_real-name_policy_controversy

[9] https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf

[10] https://sarahwyerblogs.wordpress.com/2020/08/17/classed-outliers-the-algorithmic-divide-in-plain-sight-a-levels-and-highers-divide-the-uk/

[11] https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[12] https://github.com/openai/gpt-3

[13] https://www.nbcnews.com/business/business-news/man-wrongfully-arrested-due-facial-recognition-software-talks-about-humiliating-n1232184

[14] http://gendershades.org/

[15] https://in.reuters.com/article/amazon-com-jobs-automation/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idINKCN1MK0AH

[16] https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html

[17] https://www.imdb.com/title/tt2584384/

[18] https://dataresponsibly.github.io/documents/mirror.pdf

[19] https://quoteinvestigator.com/tag/niels-bohr/

[20] https://www1.nyc.gov/site/adstaskforce/index.page