

AI, DESIGN AND ETHICS



ARTIFICIAL INTELLIGENCE IS
THE NEW BLACK

ARTIFICIAL INTELLIGENCE

**Knowledge-
based**

**Statistical
learning**

**Data
Program**

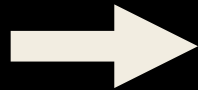


**Traditional
Programming**

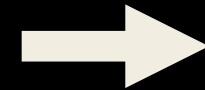


Output

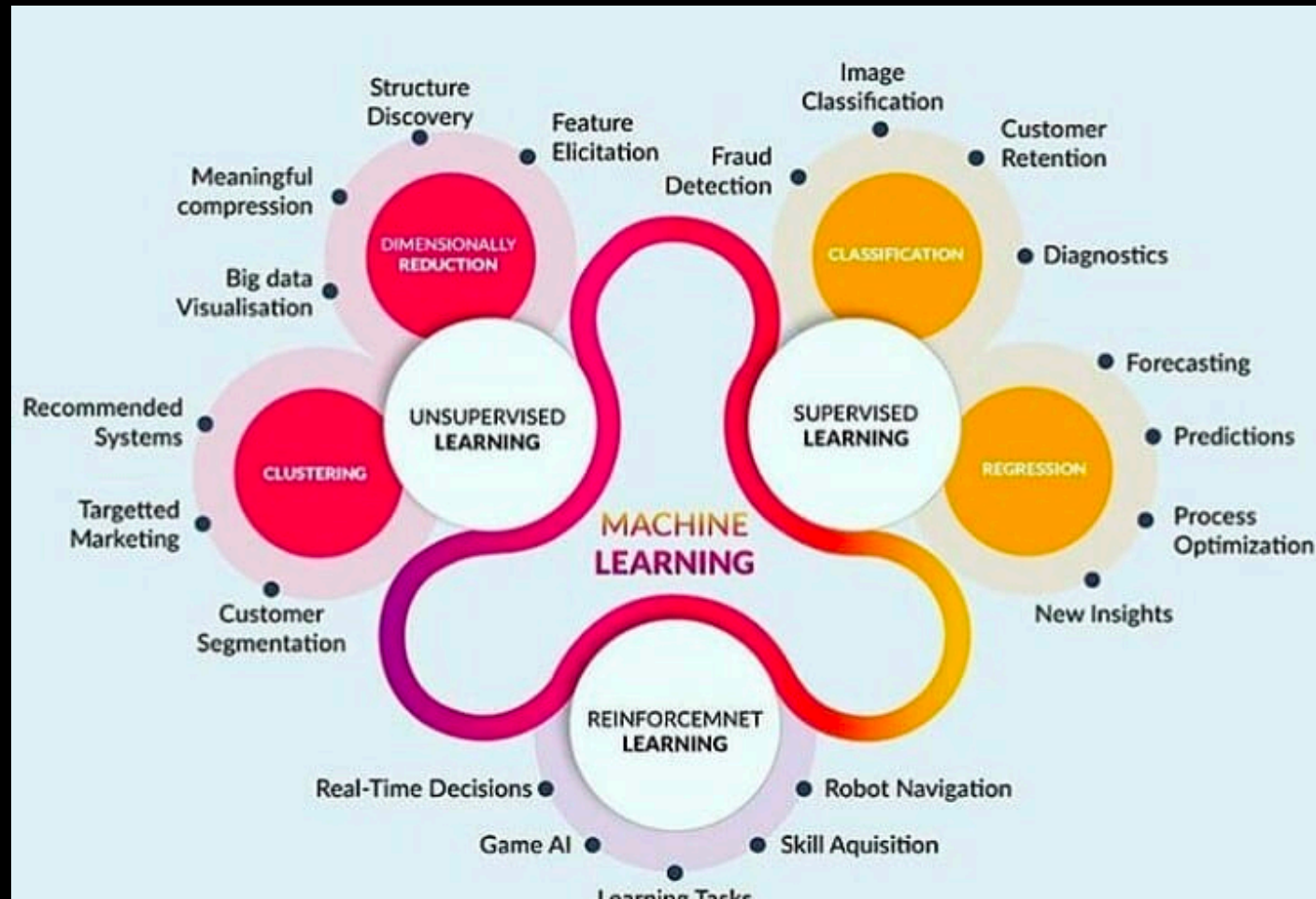
**Data
Output**



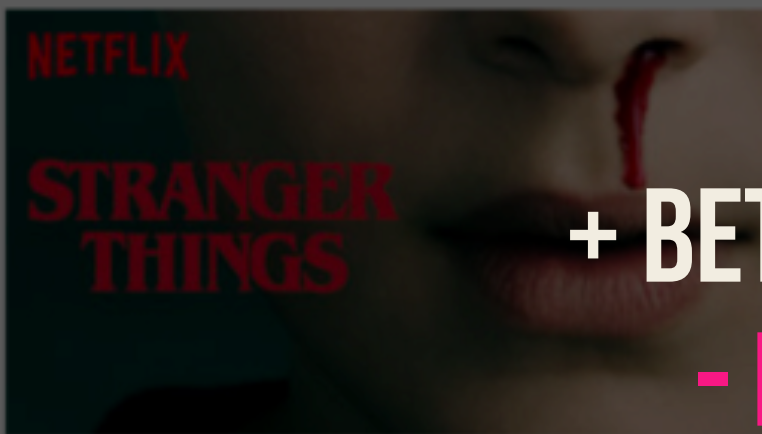
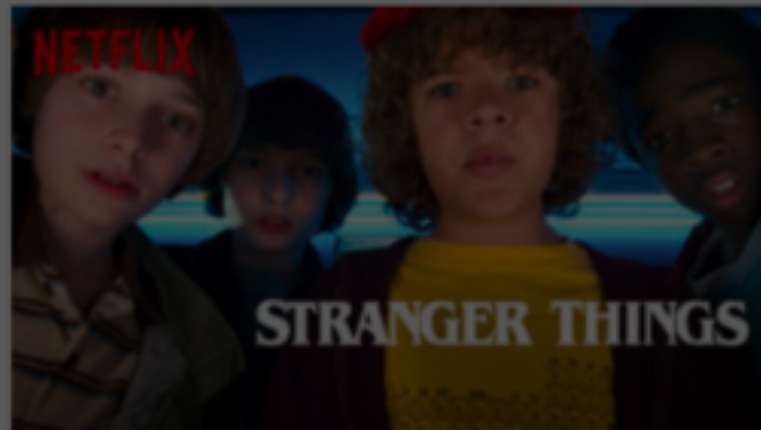
**Machine
Learning**



Program



THERE ARE DIFFERENT TECHNIQUES



+ BETTER EXPERIENCE FOR CONTENT DELIVERY
- NEW TYPES OF BUBBLES (CONTENT LOOP)



+ NATURAL USER INTERFACE
- IT JUST BRING ONE SINGLE RESPONSE



+ BETTER SHOPPING EXPERIENCE
- PRIVACY AND PHYSICAL BEHAVIOR MONITORING

Colleges Are Using Big Data To Predict Which Students Will Do Well—Before They Accept Them

Can predictive analytics determine which students succeed and which will fail? More universities are finding that the answer is yes.



+ BETTER MATCH BETWEEN COLLEGE AND STUDENTS

- DISCRIMINATION BASED ON INCOME AND POTENTIALLY RACE

MEDICAL MALAISE

If you're not a white male, artificial intelligence's use in healthcare could be dangerous



By Robert David Hart • July 10, 2017



+MORE EFFECTIVE HEALTHCARE
- DISCRIMINATION AGAINST PEOPLE WITH LIMITED ACCESS TO HEALTHCARE

Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin

8 MIN READ



SAN FRANCISCO (Reuters) - Amazon.com Inc's ([AMZN.O](#)) machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

+ BETTER MATCHING AT LOWER COST
+ CONSISTENCY TO REDUCE SOME TYPE OF BIASES (AFFINITY)
- DISCRIMINATION USING DATA FROM THE PAST

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.



+ IT CAN REDUCE SOME TYPE OF BIAS
- IT CAN PERPETUATE EXISTING TYPE OF BIAS

Some AI just shouldn't exist

Attempts to "fix" biased AI can actually harm black, gay, and transgender

By Sigal Samuel | Apr 19, 2019, 7:30am EDT



SHARE



+INCREASING SECURITY

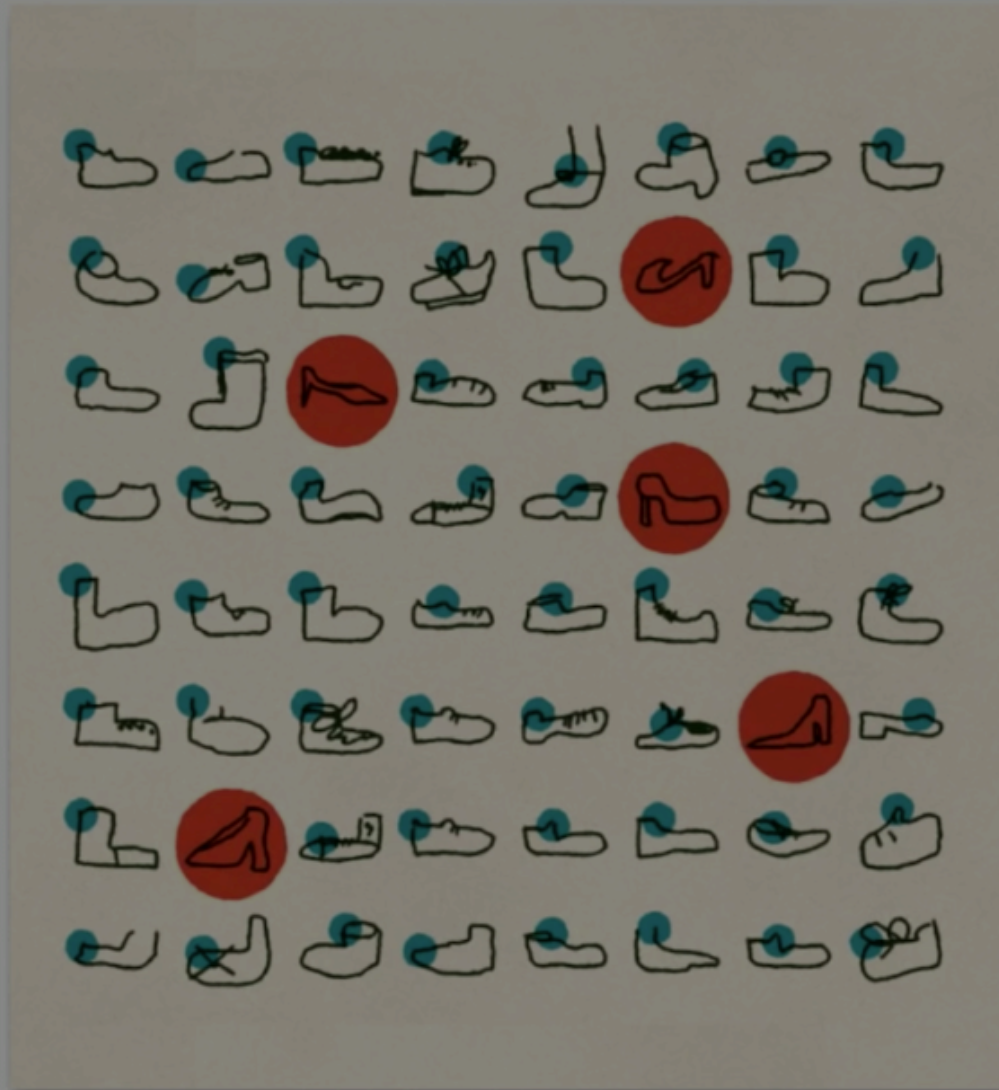
- DOESN'T WORK WELL FOR BLACKS, GAY AND TRANSGENDER



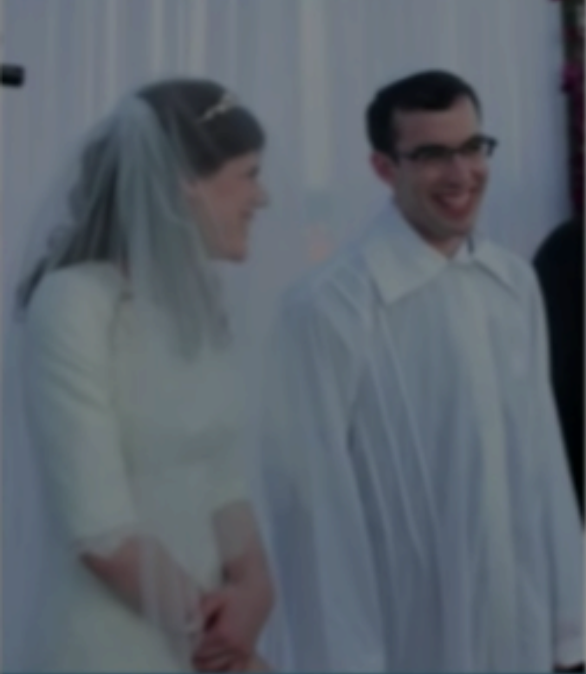
DATA JUST REFLECT THE CURRENT STATUS OF THE WORLD (AT THE BEST)
TRAINING ALGORITHMS PICK UP THE PATTERN IN DATA (EVEN WHEN WE DON'T)
ML MODELS MAKE ERRORS (BUT WE RELY ON IT)
IT MAY CAUSE UNINTENDED CONSEQUENCES

UNFAIRNESS IS DISPARATE OR UNEQUAL ALLOCATION OF **HARM/BENEFIT** TO
PEOPLE BASED ON THEIR **IRRELEVANT CHARACTERISTICS**

Ir relevant



FIRST ACTION: DATA!!



*ceremony,
wedding, bride,
man, groom,
woman, dress*



*bride,
ceremony,
wedding, dress,
woman*



*ceremony,
bride, wedding,
man, groom,
woman, dress*



person, people

WHO IS GENERATING AND LABELLING DATA?



Experiments

Partners

Developers

0.04

● Unlikely to be perceived as toxic (0.04) [Learn more](#)

SEEM WRONG?

I am straight.

0.86

◆ Likely to be perceived as toxic (0.86) [Learn more](#)

SEEM WRONG?

I am gay.

Comment	Score
I am gay.	0.86
I am queer.	0.23
I am Muslim.	0.25
I am Jewish.	0.13
I am Christian.	0.04
I am straight.	0.04

SECOND ACTION: MEASURES

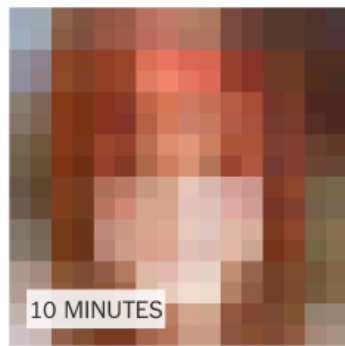


AI Explained

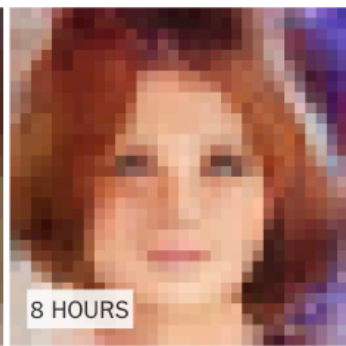
EXPLAINABLE AI



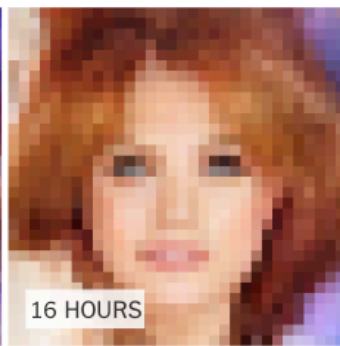
GENERATIVE ADVERSARIAL NETWORKS



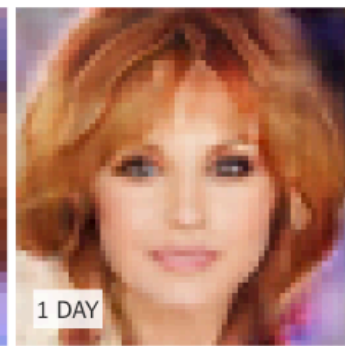
10 MINUTES



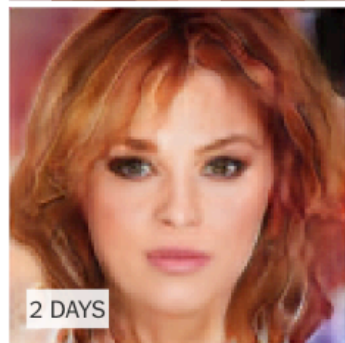
8 HOURS



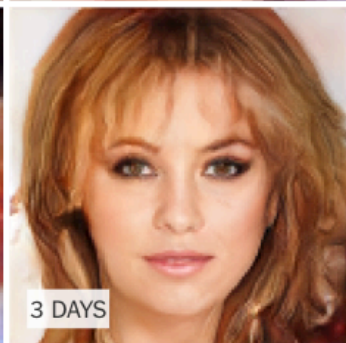
16 HOURS



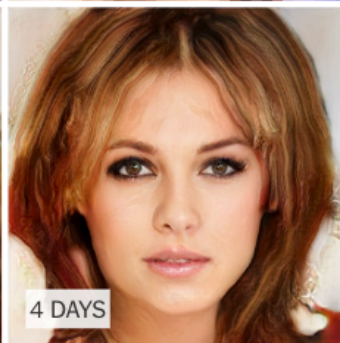
1 DAY



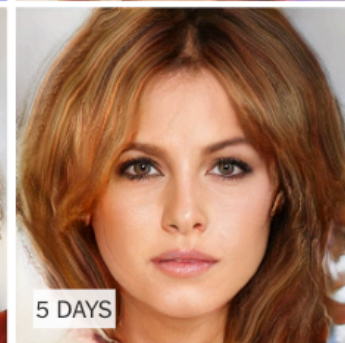
2 DAYS



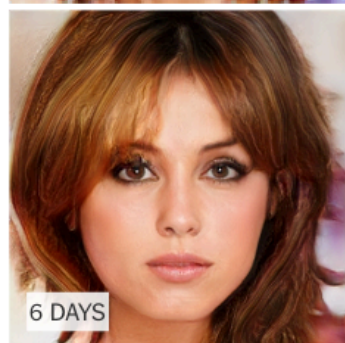
3 DAYS



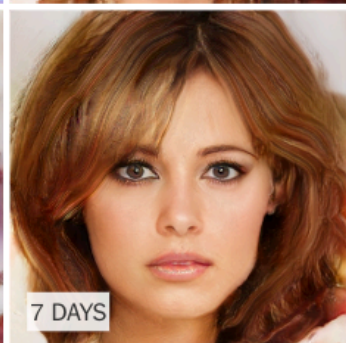
4 DAYS



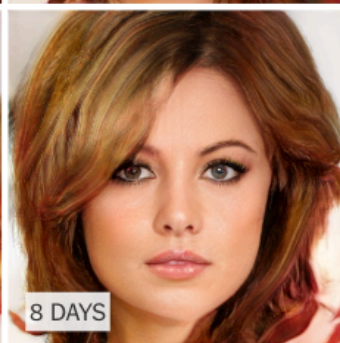
5 DAYS



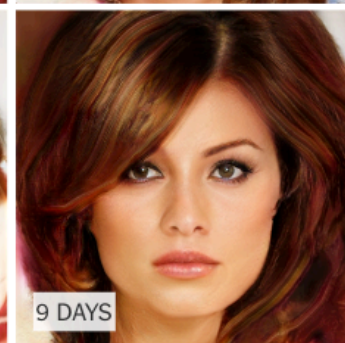
6 DAYS



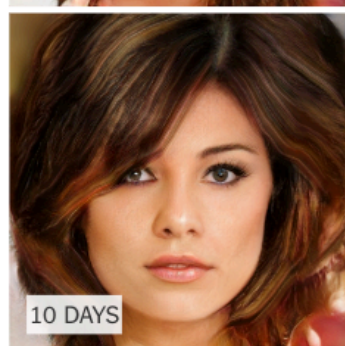
7 DAYS



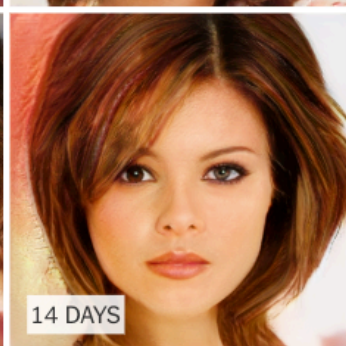
8 DAYS



9 DAYS



10 DAYS



14 DAYS



16 DAYS



18 DAYS

A collage of Spider-Man characters in a school hallway. From left to right: a young girl in a school uniform, Gwen Stacy in her white and black suit, Miles Morales in his red and blue suit, Peter Parker in his red and blue suit, and the Iron Spider armor. A clock on the wall shows the time is approximately 10:10. The text "WITH GREAT POWER COMES GREAT RESPONSIBILITIES" is overlaid in the bottom right corner.

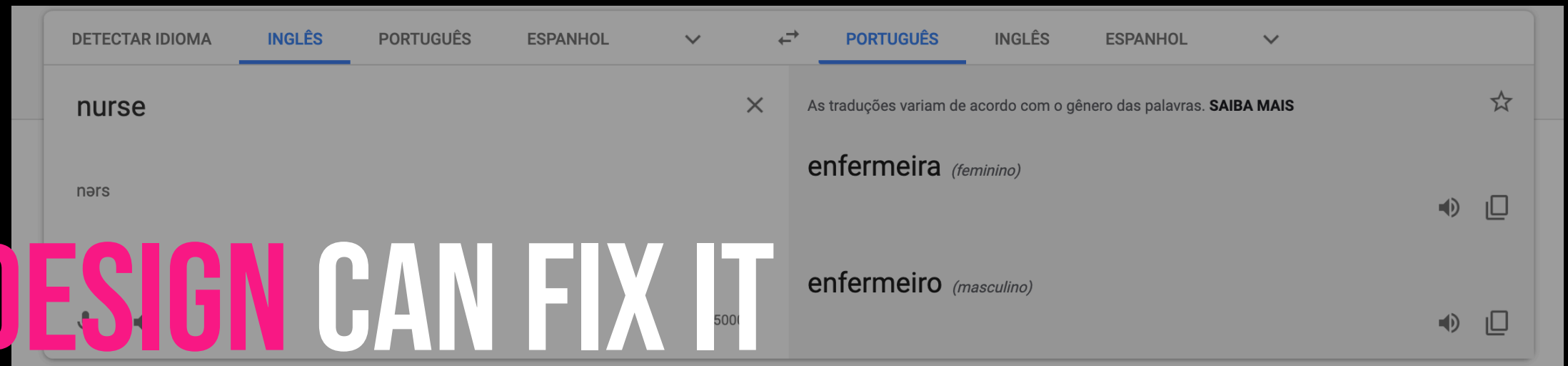
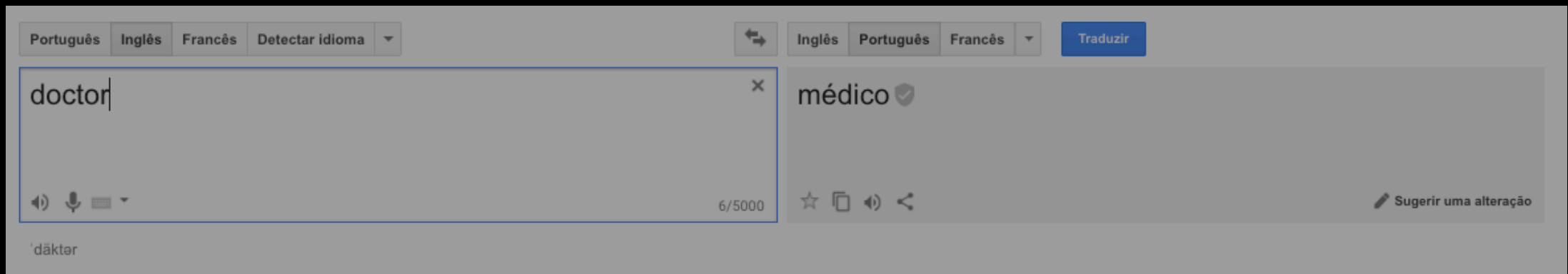
**WITH GREAT POWER COMES
GREAT RESPONSIBILITIES**

A young girl with dark hair and a pink backpack looks towards a white robot with a purple eye and a pink flower garland. The background is a blurred crowd of people at night.

**ARTIFICIAL INTELLIGENCE NEEDS A
PEOPLE-CENTERED APPROACH**



**USER NEED TO UNDERSTAND
HIS/HER ROLE IN AI SYSTEM**



Meet Q The First Genderless Voice

DESIGN CAN FIX IT

THREE MAIN TOPICS

**Quality of
Training
data**

**System
Design**

**Complex
Interactions**



GROUP
LIGENCE ARTIFICIELLE
Vers une approche humaniste ?
Une conférence mondiale
4 mars 2019
Salle de l'UNESCO, Paris, France

WORKSHOP UNESCO MOBILE LEARNING



Responsible AI

A GLOBAL POLICY FRAMEWORK

ETHICAL PURPOSE SOCIETAL BENEFIT ACCOUNTABILITY TRANSPARENCY EXPLAINABILITY FAIRNESS NON-DISCRIMINATION
SAFETY RELIABILITY OPEN DATA FAIR COMPETITION PRIVACY INTELLECTUAL PROPERTY



FIRST EDITION

ETHICAL

FIRST EDITION

FIRST EDITION

FIRST EDITION



THE IGF IS A GLOBAL MULTISTAKEHOLDER PLATFORM THAT FACILITATES THE DISCUSSION OF PUBLIC POLICY ISSUES PERTAINING TO THE INTERNET

IGF 2019 WS #179 Human-centered Design and Open Data: how to improve AI

Theme:

Data Governance

Subtheme(s):

Algorithms/Artificial Intelligence/Machine Learning

Cross border data

Innovation

Organizer 1: [Caroline Burle](#), Ceweb.br/NIC.br

Organizer 2: [Hartmut Richard Glaser](#), Brazilian Internet Steering Committee - CGI.br

Organizer 3: [Vagner Diniz](#), NIC.br - CEWEB.br

Organizer 4: [Talitha Nicoletti](#), Pontifical Catholic University of São Paulo - PUC-SP

Organizer 5: [Cappi Juliano](#), NIC.br

Organizer 6: [Nathalia Patrício](#), NIC.br

Organizer 7: [Mariko Kobayashi](#), Keio University/Mercari, Inc.

Speaker 1: [Jaimie Boyd](#), Government, Western European and Others Group (WEOG)

Speaker 2: [Karine Perset](#), Intergovernmental Organization, Western European and Others Group (WEOG)

Speaker 3: [Diogo Cortiz da Silva](#), Technical Community, Latin American and Caribbean Group (GRULAC)

Speaker 4: [Luciana Terceiro](#), Civil Society, Latin American and Caribbean Group (GRULAC)

Speaker 5: [Juan Ortiz Freuler](#), Civil Society, Western European and Others Group (WEOG)

ÉTICA 4.0

CICLO DE PALESTRAS poli.TIC



Terça-feira, 02/07 às 10 horas
Auditório do CTI Renato Archer

DIOGO CORTIZ

"IA, DESIGN E ÉTICA"




Emergent Tech ▶ **Artificial Intelligence**

Standards group W3C wins support from all major players to get AI working in the browser

Google, Microsoft, Apple and Mozilla on board, says chair

By [Tim Anderson](#) 17 May 2019 at 15:30

37 

SHARE ▼

Intel's Web Standards tech lead, Anssi Kostiainen, has [said](#) the World Wide Web Consortium's (W3C) Machine Learning for the Web Community Group now has "all major browsers – Google, Microsoft, Apple, Mozilla – on board along with the broader AI & web ecosystem."

TensorFlow.js is a library for developing and training ML models in JavaScript, and deploying in browser or on Node.js

See tutorials

Tutorials show you how to use TensorFlow.js with complete, end-to-end examples.

See the guide

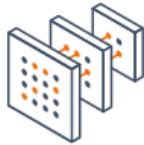
Guides explain the concepts and components of TensorFlow.js.

See demos

Live demos and examples run in your browser using TensorFlow.js.



How it works



Run existing models

Use off-the-shelf JavaScript models or convert Python TensorFlow models to run in the browser or under Node.js.



Retrain existing models

Retrain pre-existing ML models using your own data.



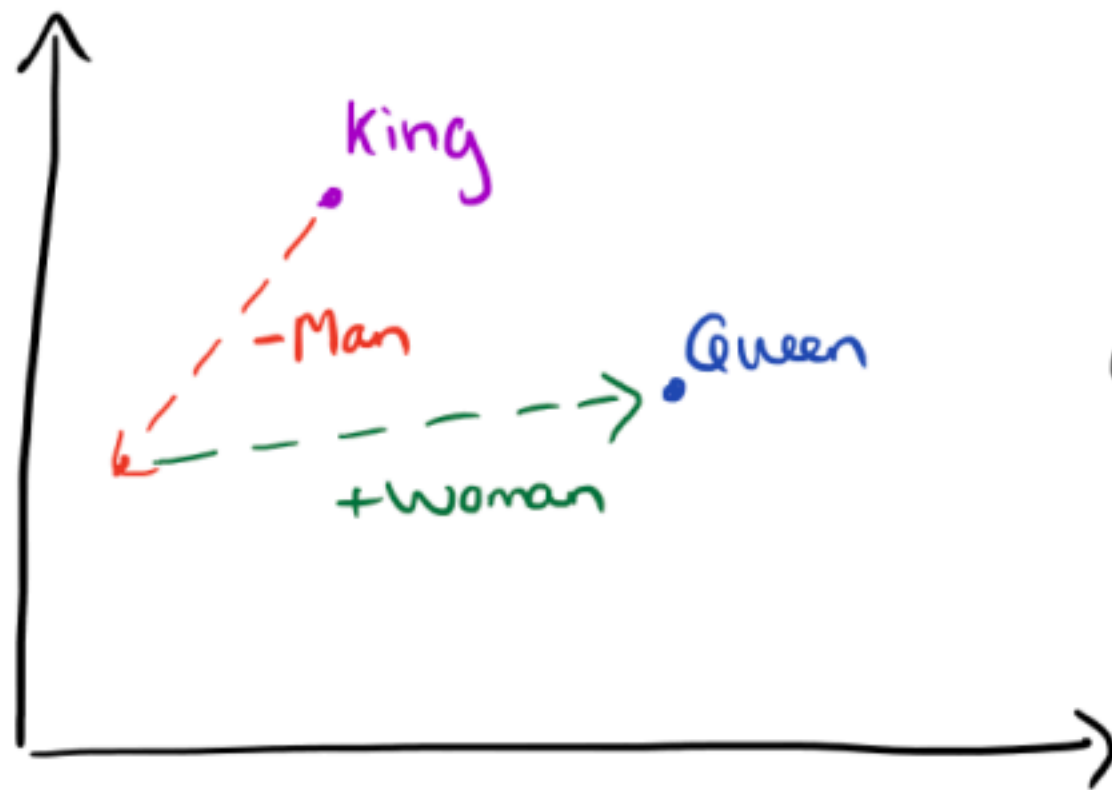
Develop ML with JavaScript

Build and train models directly in JavaScript using flexible and intuitive APIs.

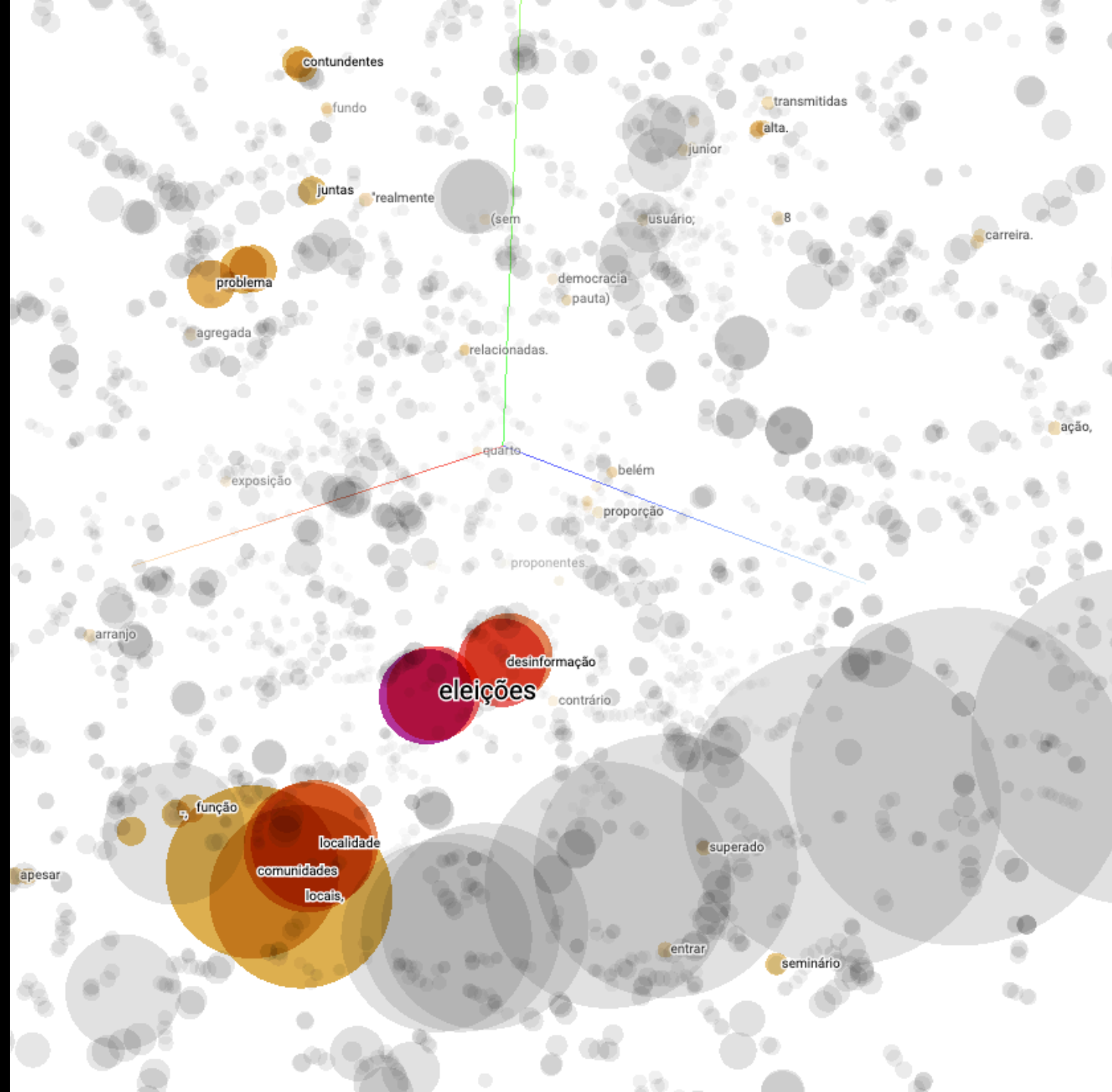
[Home](#) + [Research](#) + [Research projects](#)

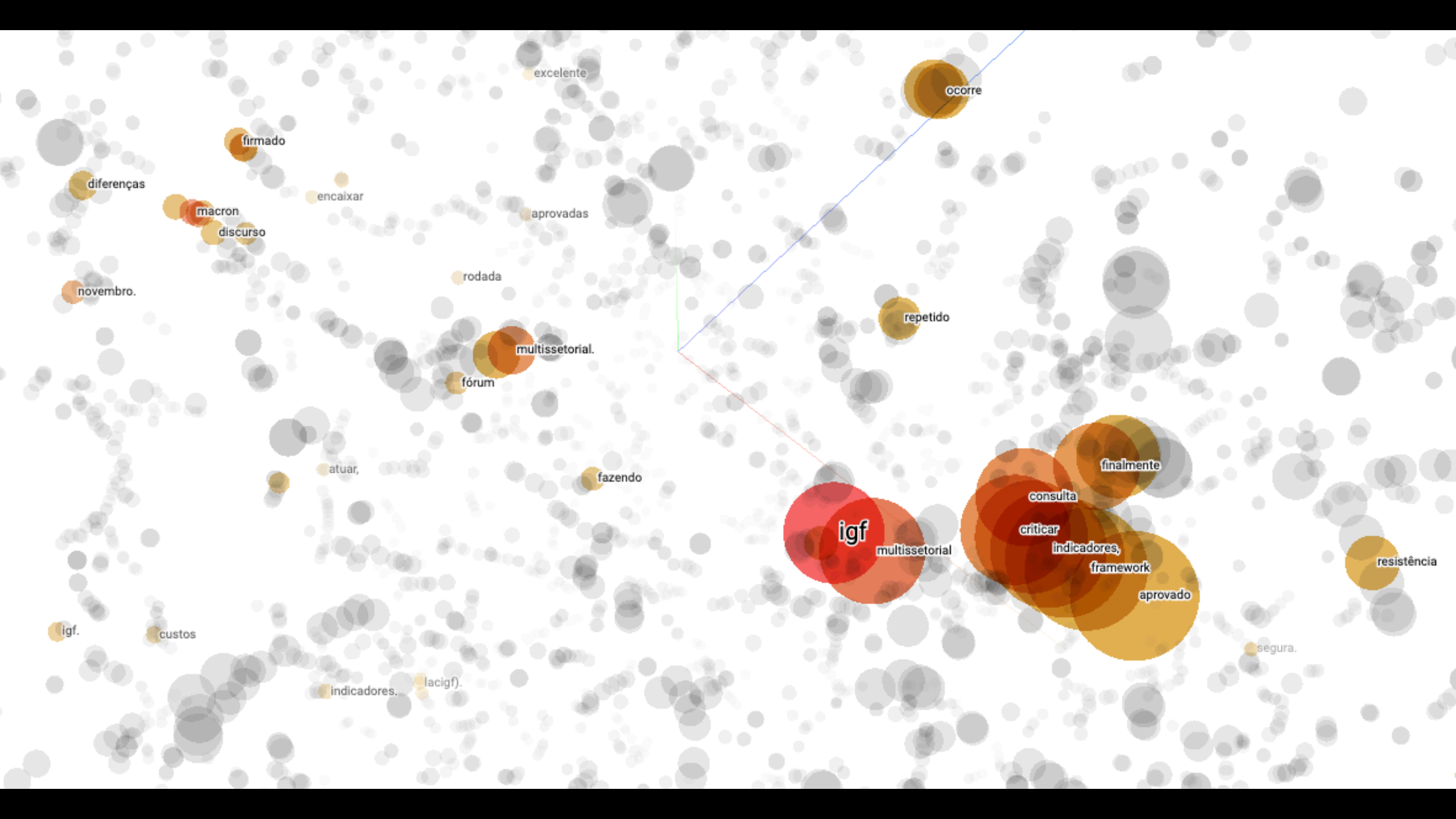
Citizen participation and machine learning for a better democracy

Using machine learning and natural language processing to overcome information overload in citizen participation platforms



Vector
Composition









European
Commission

CORDIS
EU research results

English

Search

HOME

RESULTS PACKS

RESEARCH*EU MAGAZINES

NEWS & EVENTS

PROJECTS & RESULTS

ABOUT US

Sign in

HORIZON
2020

Fake News discovery and propagation from big Data ANALysis and artificial intelligence Operations

Fact Sheet

Objective

Fake News are now a hot issue in Europe as well as worldwide, particularly referred to Political and Social Challenges that reflect in business as well as in industry. Europe is lacking of a systematic knowledge and data transfer across organizations to address the aggressive emergence of the well-known problem of fake news and post-truth effect. The possibility to use cross sector Big Data management and analytics, along with an effective interoperability scheme for all our data sources, will tackle this urgent problem, generating new business and societal impacts involving several stakeholders: a) Media Companies: news agencies, broadcaster, newspapers, etc, b) Governmental institutions and organisations, c) The overall industrial ecosystem, d) The entire society. The aim of FANDANGO is to aggregate and verify different typologies of news data, media sources, social media, open data, so as to detect fake news and provide a more efficient and verified communication for all European citizens. European tradition in democracy, journalism and transparency should play a worldwide example in fast changing society, where all citizens appears completely overwhelmed by the new technologies and by the new social challenges. The FANDANGO project aims to break data interoperability barriers providing unified techniques and an integrated big data platform to support traditional media industries to face the new "data" news economy with a better transparency to the citizens under a Responsible, Research and Innovation prism. This goal will be validated and tested in three specific domains Climate, Immigration and European Context, these are typical scenarios where fake news can influence perception with respect to social and business actions and where news can be verified and validated by trustable information, based on facts and data.

Programme(s)

Project information

FANDANGO

Grant agreement ID: 780355

Status
Ongoing project

Start date	End date
1 January 2018	31 December 2020

Funded under:
H2020-EU.2.1.1.

Overall budget:
€ 3 583 125

EU contribution
€ 2 879 250



Coordinated by:
ENGINEERING - INGEGNERIA INFORMATICA
SPA



THANK YOU!

FACEBOOK.COM/DIOGOCORTIZ

DIOGO@NIC.BR