Coexister avec nos artifices

Introduction à l'épistémologie du numérique

Modern Times, Charlie Chaplin, 1936

Boris Beaude - STSLab - Université de Lausanne

Master en humanités et cultures numériques - UNIL - 2021

Coexister avec nos artifices

Lee Sedol vs AlphaGo (DeepMind) - Mars 2016

Boris Beaude - STSLab - Université de Lausanne

Coexister avec nos artifices

James Bridle - Autonomous Trap 001 - Magic Salt Circle - 14 mars 2017

Quelle nouveauté ?

Problématique du changement

- Expertise -> Expérience
- Puissance de traitement
- *Données*
- *Algorithmes*
- Modèles économiques

Sous quel forme ?

Problématique de l'existence matérielle

Androïde / Humanoïde

Problématique de l'anthropomorphisme et vallée de l'étrange…

Cyborg

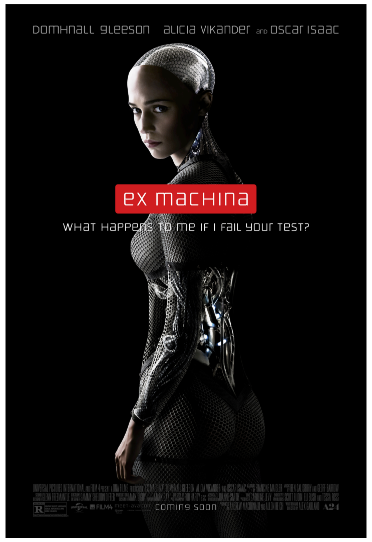
Problématique du transhumanisme…

Bot

Problématique de l'algorithme…

Hybride et distribuée

Problématique de la cybernétique…



Intelligence artificielle ?

Intelligence produite par l'activité humaine

Spécialisée

Intelligence artificielle faible

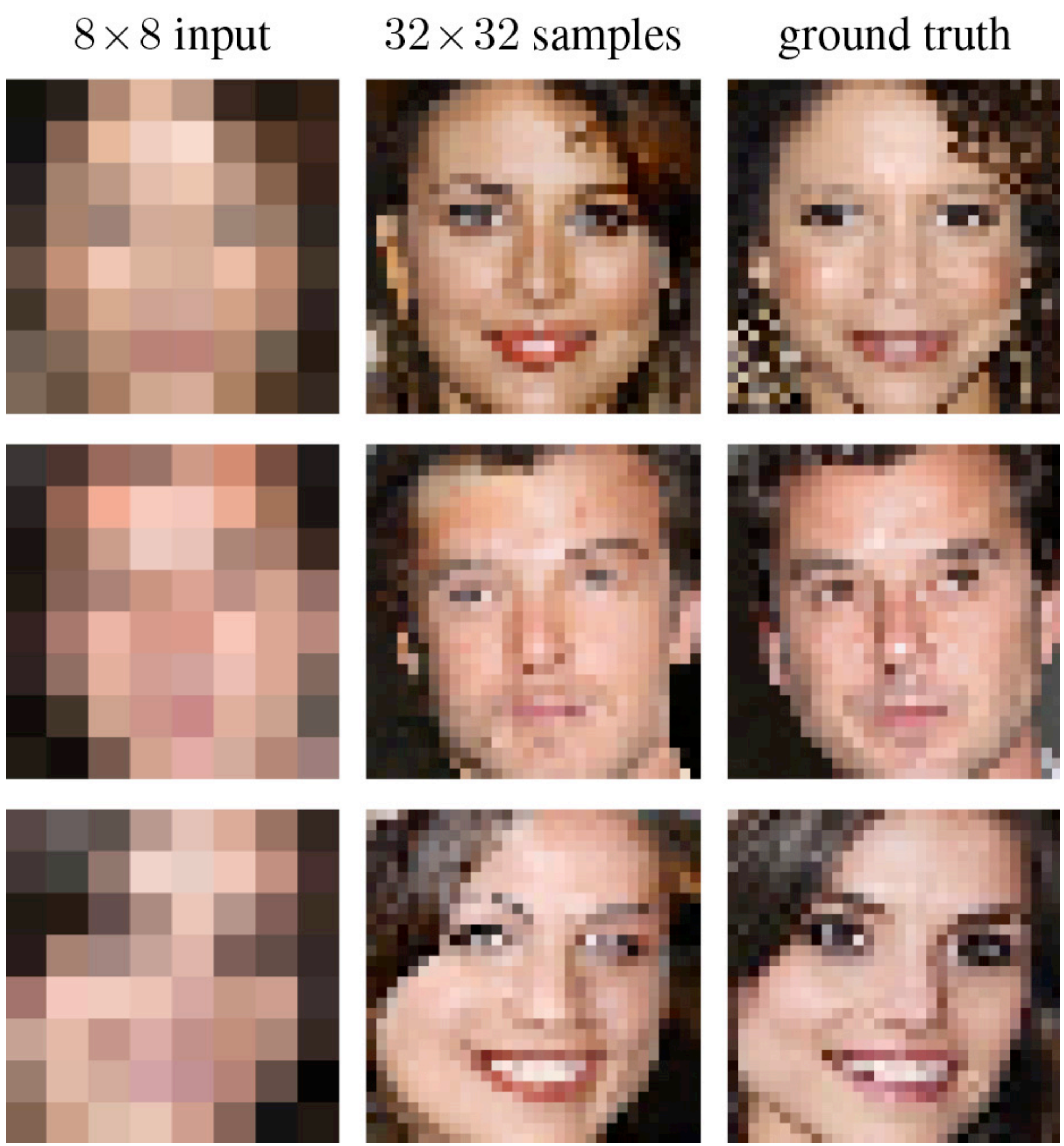
Analogique

Intelligence artificielle forte

Indépendante

Singularité

Coexister avec nos artifices



Pixel Recursive Super Resolution - Ryan Dahl & al. / Google Brain – Février 2017

Coexister avec nos artifices

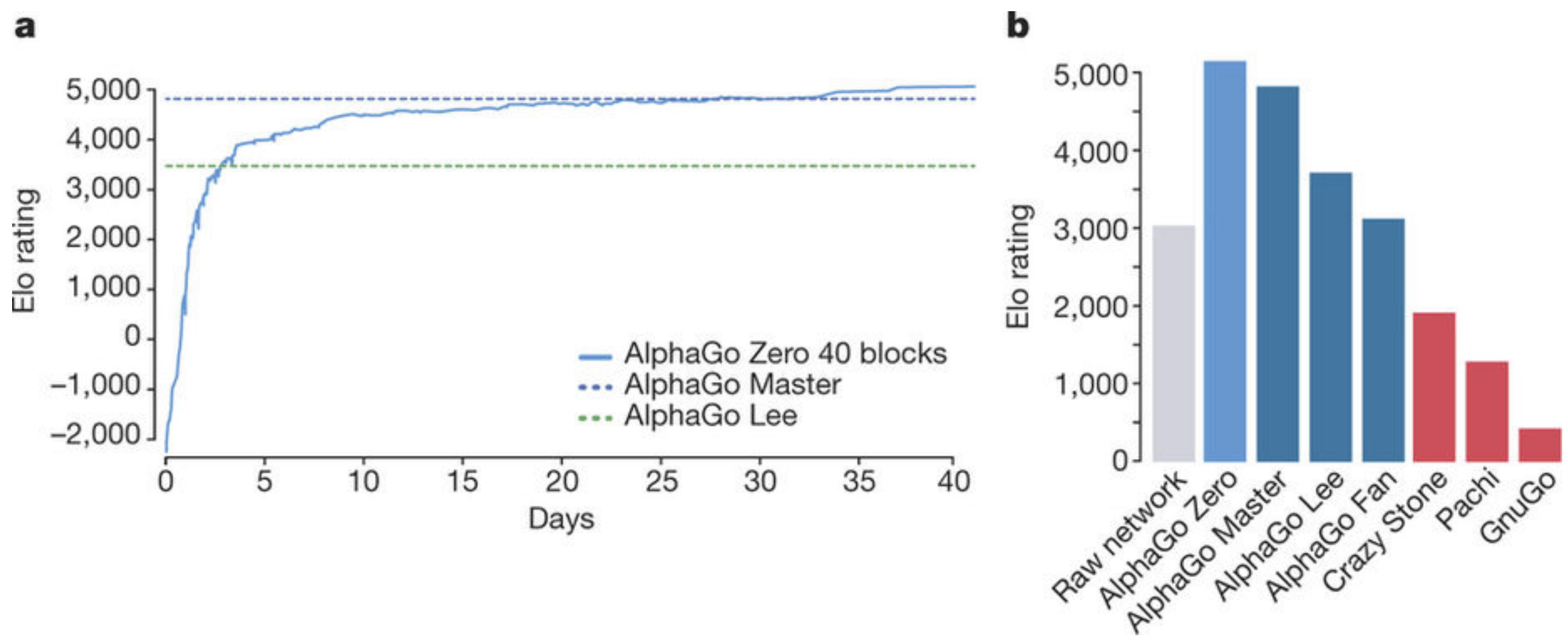
UK National Grid control room and DeepMind - Mars 2017

Coexister avec nos artifices

LipNet - University of Oxford and DeepMind - Novembre 2016

Wei Xiaoyong (Sichuan University) - Students being checked for boredom levels in class - West China Metropolis Daily - Septembre 2016

Mastering the game of Go without human knowledge



(1) **AlphaGo** Fan is the previously published program that played against Fan Hui in **October 2015**. This program was distributed over many machines using 176 GPUs.

(2) **AlphaGo Lee** is the program that defeated Lee Sedol 4–1 in **March 2016**. It was previously unpublished, but is similar in most regards to AlphaGo Fan - 48 TPUs, rather than GPUs

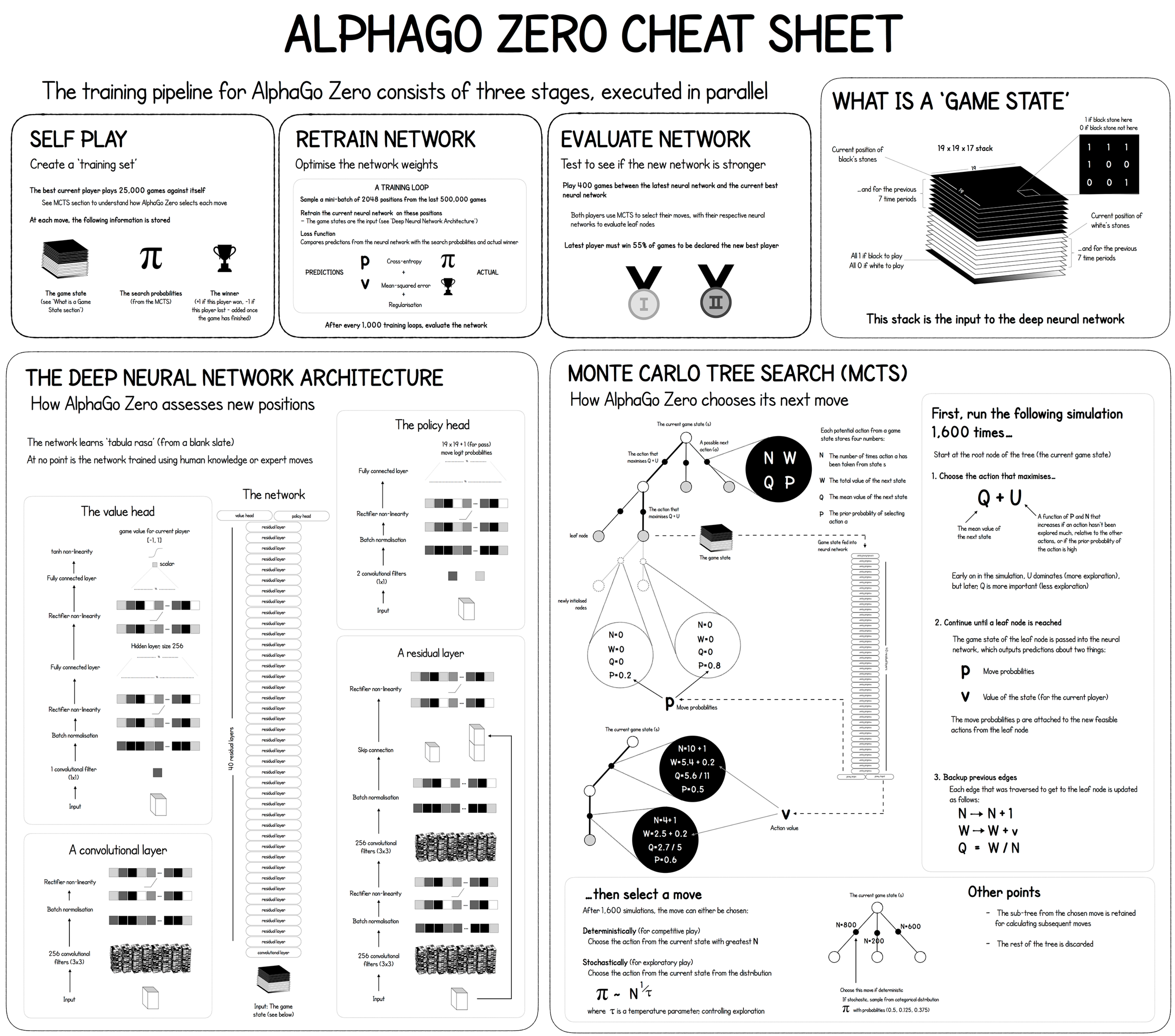
(3) **AlphaGo Master** is the program that defeated top human players by 60–0 in **January 2017** - initialized by supervised learning from human data + reinforcement

(4) **AlphaGo Zero** learns from self-play reinforcement learning, with no human supervision in **April 2017**. It uses just a single machine in the Google Cloud with 4 TPUs

(x) **AlphaZero** is a more generalized variant of the AlphaGo Zero (AGZ) algorithm, and is able to play **shogi** and **chess** as well as **Go**.
- Defeated Stockfish 8 after **9 hours** of training.
- Defeated AlphaGo Zero after **34 hours** of training.

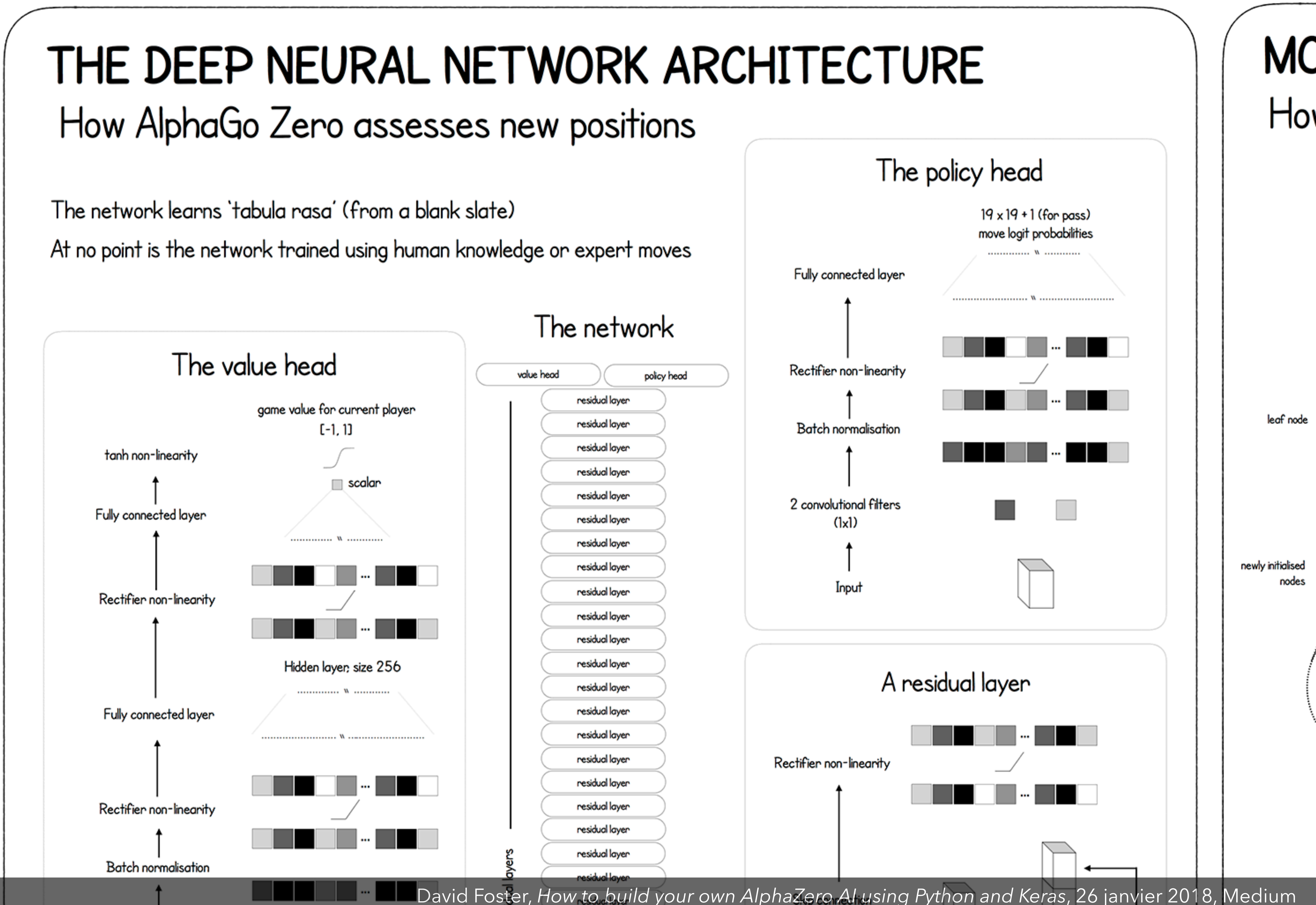
David Silver et al., *Mastering the game of Go without human knowledge*, 7 avril / 18 octobre 2017, Nature

Coexister avec nos artifices



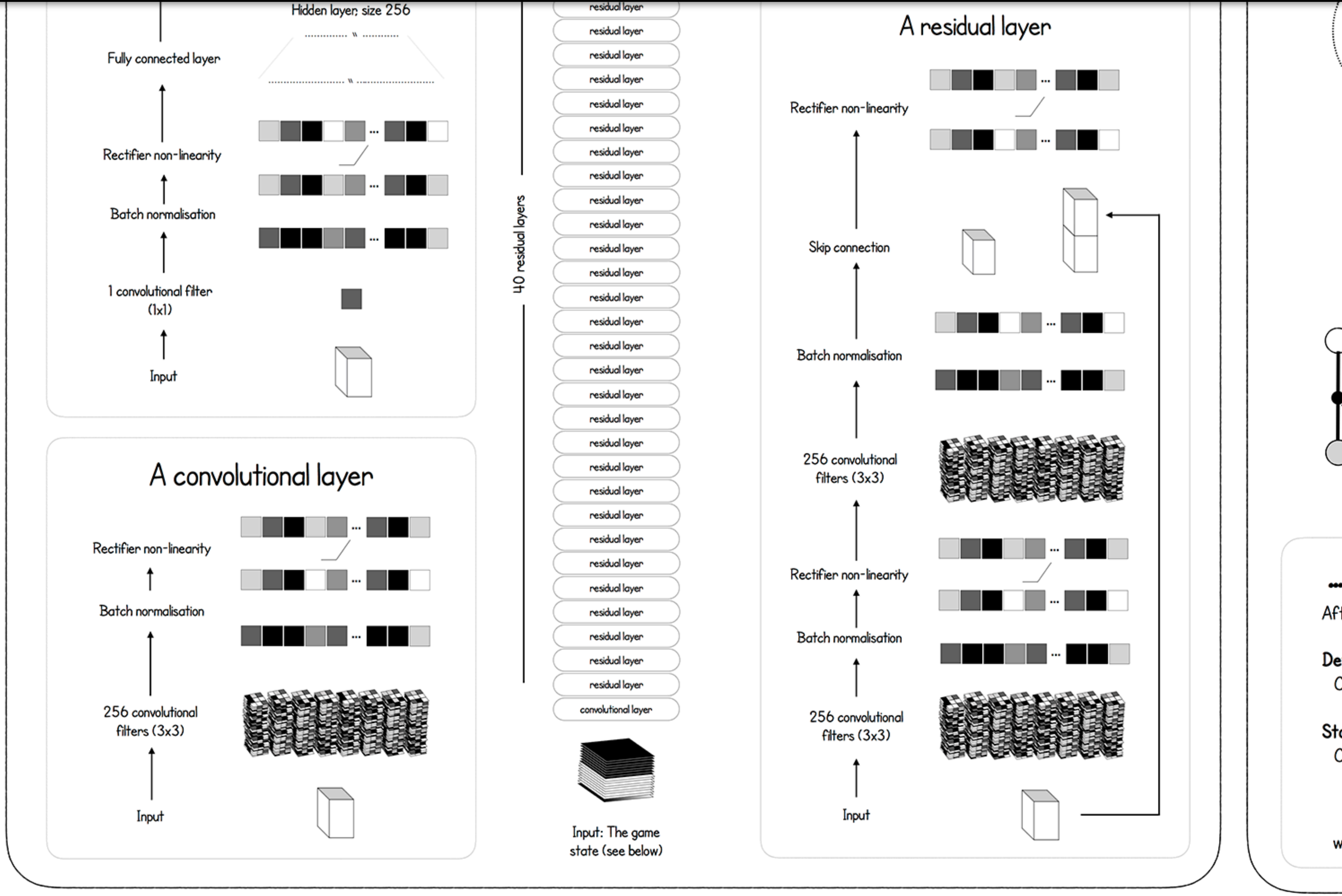
David Foster, *How to build your own AlphaZero AI using Python and Keras*, 26 janvier 2018, Medium

Coexister avec nos artifices

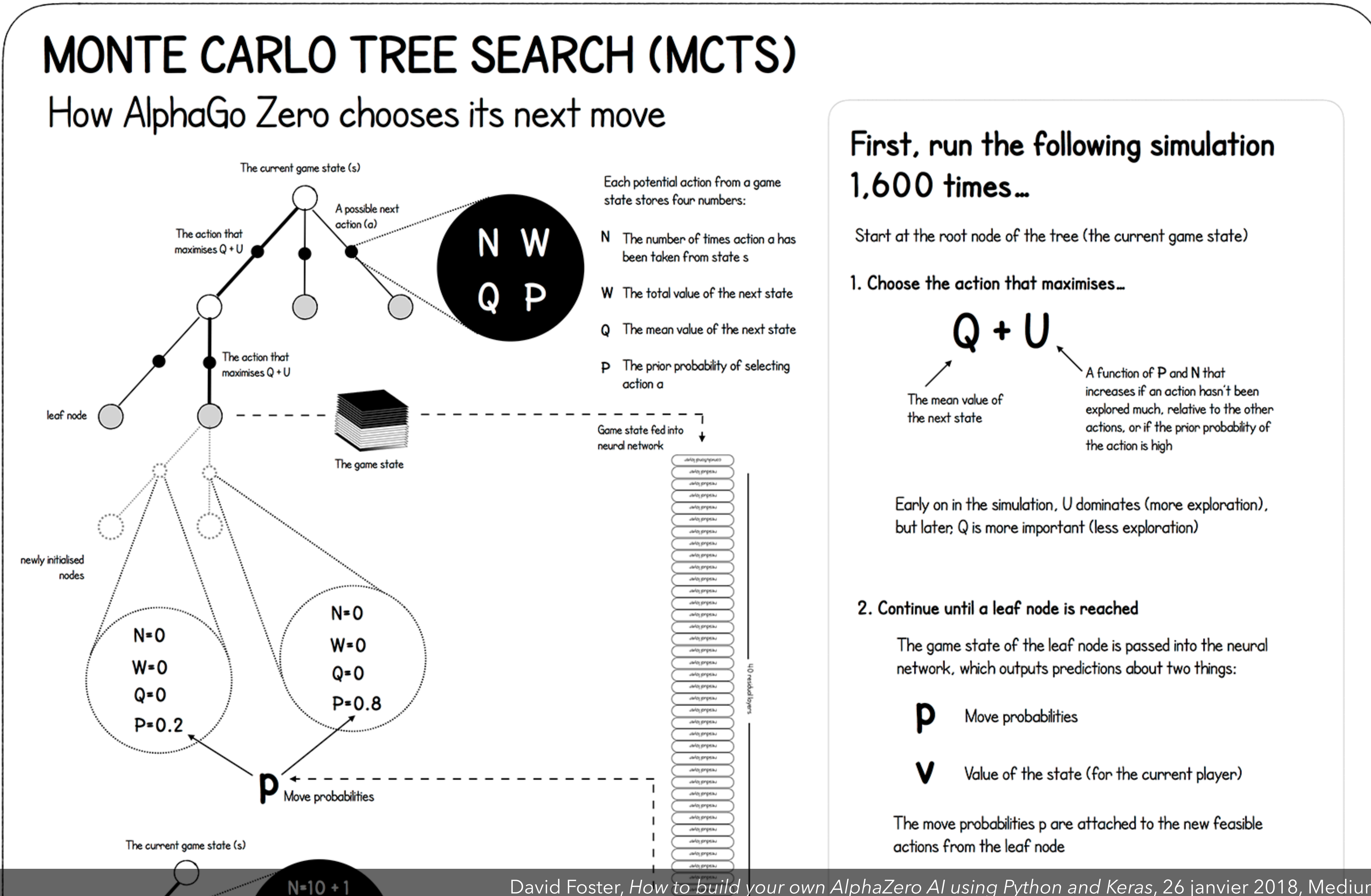


David Foster, *How to build your own AlphaZero AI using Python and Keras*, 26 janvier 2018, Medium

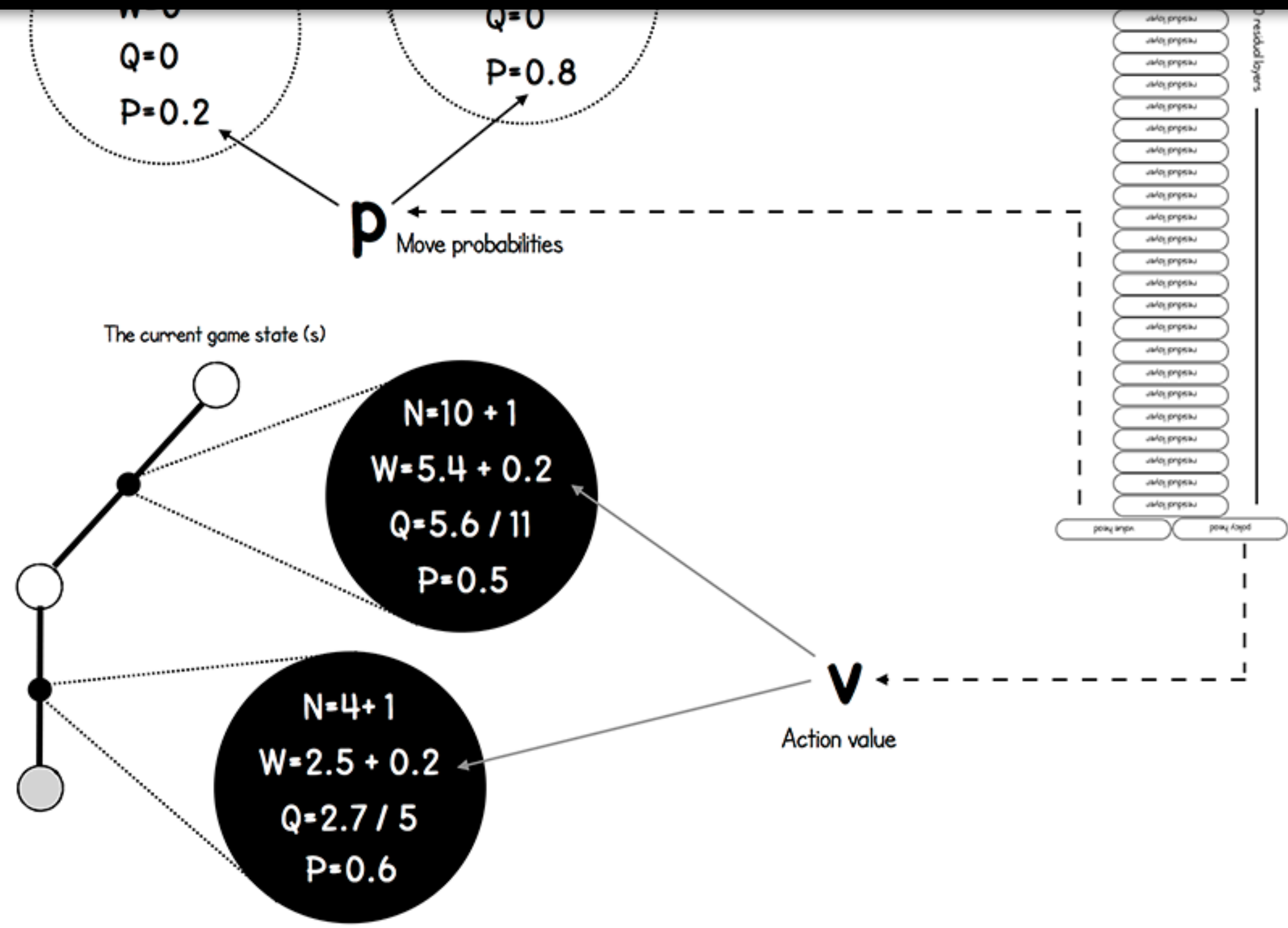
Coexister avec nos artifices



David Foster, *How to build your own AlphaZero AI using Python and Keras*, 26 janvier 2018, Medium



David Foster, *How to build your own AlphaZero AI using Python and Keras*, 26 janvier 2018, Medium



network, which outputs predictions about two things:

p Move probabilities

v Value of the state (for the current player)

The move probabilities p are attached to the new feasible actions from the leaf node

3. Backup previous edges

Each edge that was traversed to get to the leaf node is updated as follows:

$$N \rightarrow N + 1$$
$$W \rightarrow W + v$$
$$Q = W / N$$

...then select a move

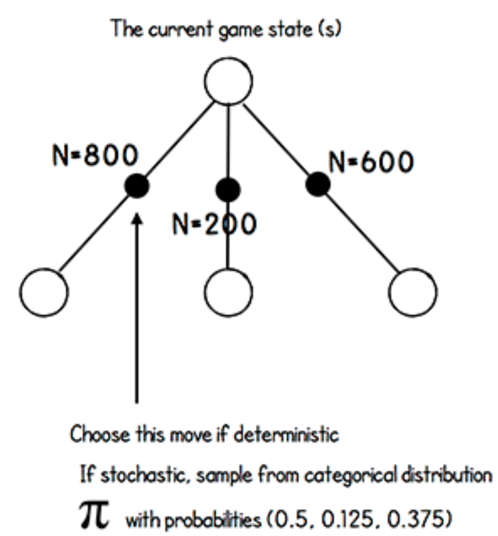
After 1,600 simulations, the move can either be chosen:

Deterministically (for competitive play)
Choose the action from the current state with greatest N

Stochastically (for exploratory play)
Choose the action from the current state from the distribution

$$\pi \sim N^{1/\tau}$$

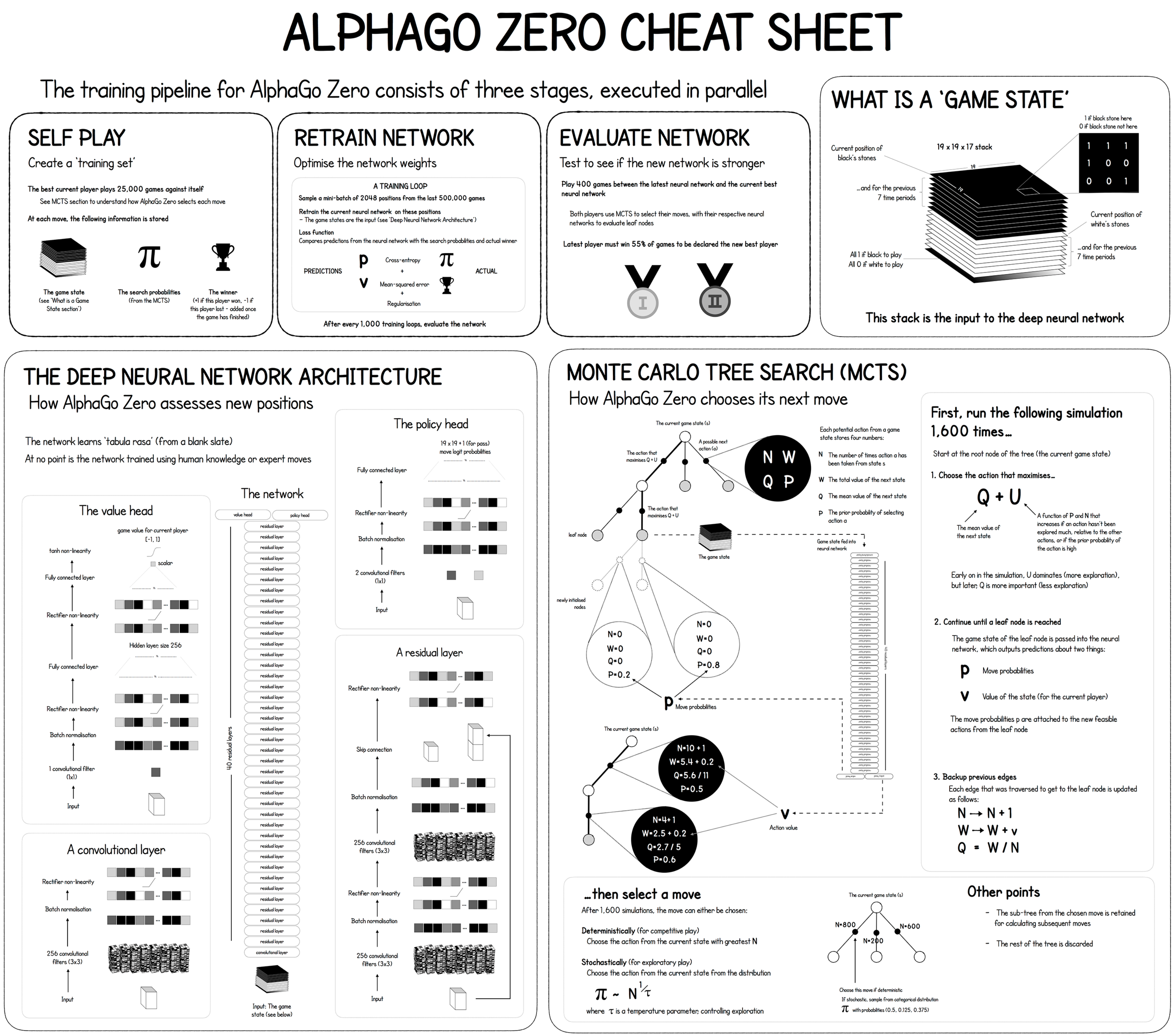
where τ is a temperature parameter, controlling exploration



Other points

- The sub-tree from the chosen move is retained for calculating subsequent moves
- The rest of the tree is discarded

David Foster, *How to build your own AlphaZero AI using Python and Keras*, 26 janvier 2018, Medium



David Foster, *How to build your own AlphaZero AI using Python and Keras*, 26 janvier 2018, Medium

AI effect ?

Problématique de la qualification de l'intelligence

- La reconnaissance de caractères (Lecun 1989)
- Les échecs (IBM Deep Blue - mai 1997)
- L'érudition - Jeopardy (IBM Watson - 16 février 2011)
- Le Go (DeepMind AlphaGo Lee - mars 2016 & AlphaGo Master - janvier 2017)
- La compréhension d'un texte (Alibaba 15 janvier 2018)
- La traduction ? (DeepL , Google, Microsoft…)
- Conduire ? (Waymo, Tesla, Volvo, BMW, GM, Apple…)

Coexister avec nos artifices

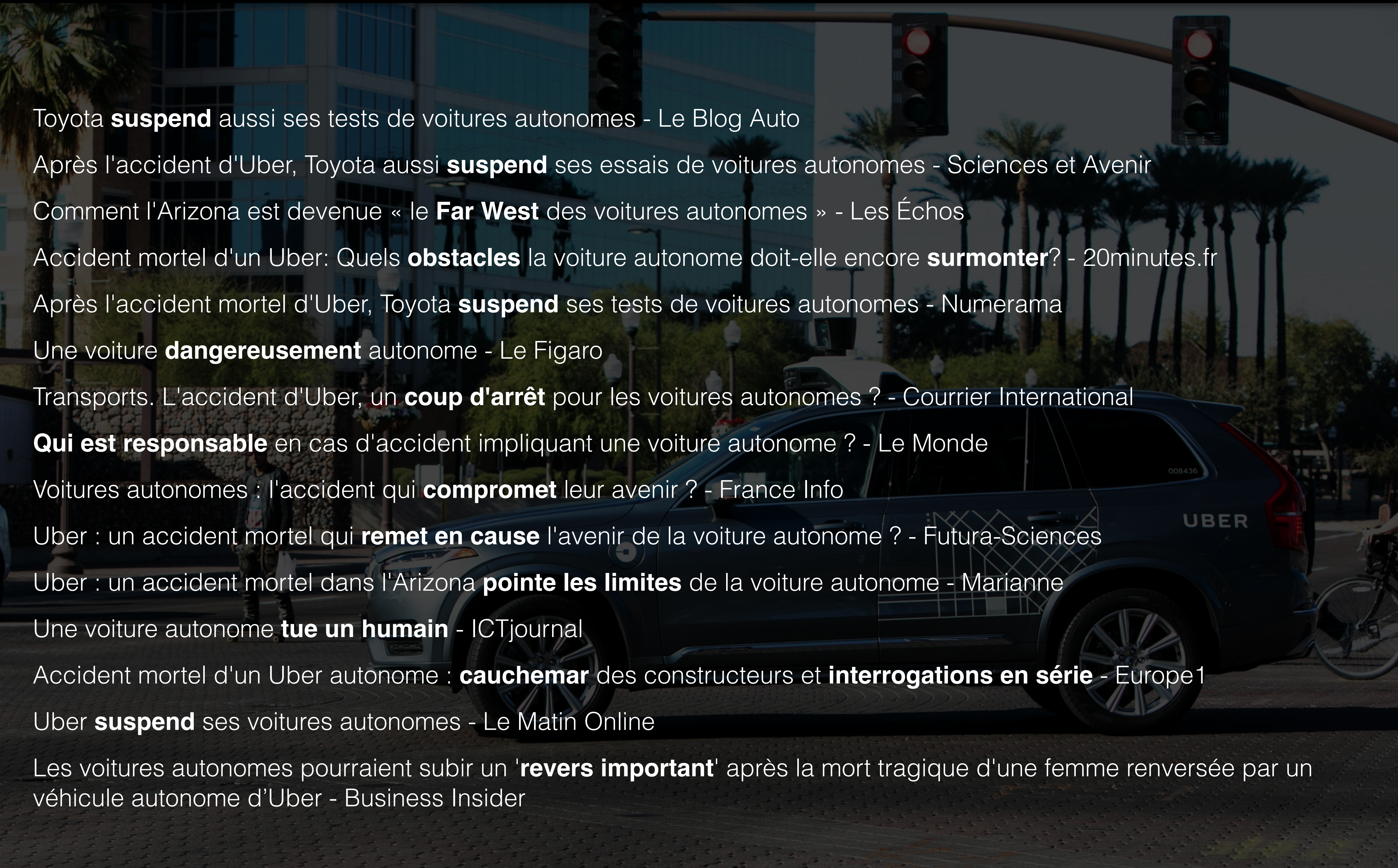
…

Coexister avec nos artifices

'Uber's Self-Driving Car Just Killed Somebody. Now What?'

Aarian Marshall, *Uber's Self-Driving Car Just Killed Somebody. Now What?*, 19 mars 2018, Wired

Coexister avec nos artifices



Toyota **suspend** aussi ses tests de voitures autonomes - Le Blog Auto

Après l'accident d'Uber, Toyota aussi **suspend** ses essais de voitures autonomes - Sciences et Avenir

Comment l'Arizona est devenue « le **Far West** des voitures autonomes » - Les Échos

Accident mortel d'un Uber: Quels **obstacles** la voiture autonome doit-elle encore **surmonter**? - 20minutes.fr

Après l'accident mortel d'Uber, Toyota **suspend** ses tests de voitures autonomes - Numerama

Une voiture **dangereusement** autonome - Le Figaro

Transports. L'accident d'Uber, un **coup d'arrêt** pour les voitures autonomes ? - Courrier International

Qui est responsable en cas d'accident impliquant une voiture autonome ? - Le Monde

Voitures autonomes : l'accident qui **compromet** leur avenir ? - France Info

Uber : un accident mortel qui **remet en cause** l'avenir de la voiture autonome ? - Futura-Sciences

Uber : un accident mortel dans l'Arizona **pointe les limites** de la voiture autonome - Marianne

Une voiture autonome **tue un humain** - ICTjournal

Accident mortel d'un Uber autonome : **cauchemar** des constructeurs et **interrogations en série** - Europe1

Uber **suspend** ses voitures autonomes - Le Matin Online

Les voitures autonomes pourraient subir un '**revers important**' après la mort tragique d'une femme renversée par un véhicule autonome d'Uber - Business Insider

Google actualité - « Uber » - 21 mars 2018

Coexister avec nos artifices

Malgré l'accident d'Uber, **BMW accélère** vers la voiture autonome - LaPresse.ca

Malgré l'accident d'Uber, **BMW persiste** dans la voiture autonome - Le Soir

La **police**: 'Uber n'est **pas responsable** de l'accident mortel' - Le Vif

Accident mortel de voiture autonome : «Ça ne doit **pas remettre en cause** les projets», estime un **expert** - Le Parisien

Selon la **police**, « Uber n'est probablement pas en tort » dans l'accident mortel causé par une voiture autonome - Numerama

Uber **dédouané** par la chef de la **police** dans l'accident qui a causé la mort d'une femme - La Revue du digital

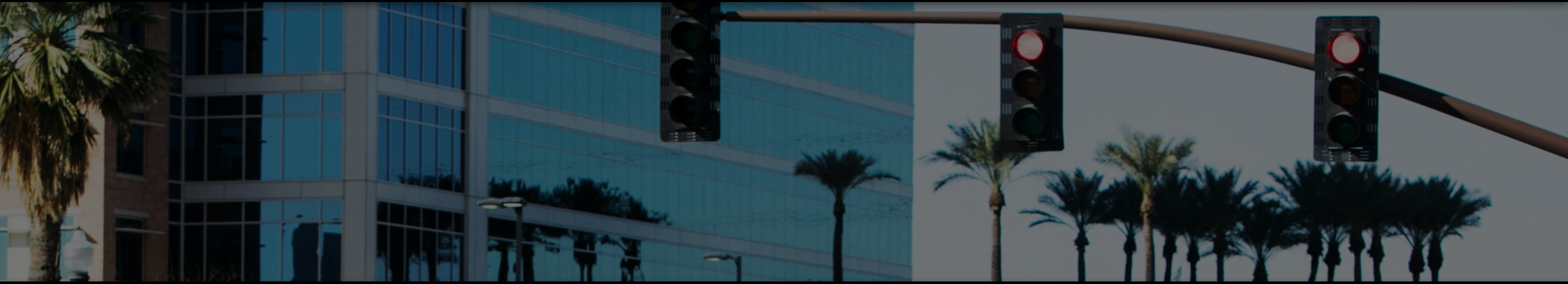
Accident mortel provoqué par un véhicule autonome : la **police blâme la victime** - Le Soir

Pourquoi la voiture autonome n'est **pas remise en cause** par l'accident mortel provoqué par Uber - BFMTV.COM

Tués par des humains ? « Près de 40 000 morts par an aux Etats-Unis, dont près de 6000 piétons - plus de 16 par jour ».

Google actualité - « Uber » - 21 mars 2018

Coexister avec nos artifices



Réflexions hors sol et sans contexte ?

Aarian Marshall, *Uber's Self-Driving Car Just Killed Somebody. Now What?*, 19 mars 2018, Wired

Coexister avec nos artifices

Réflexions hors sol et sans contexte ?

Montées en généralité

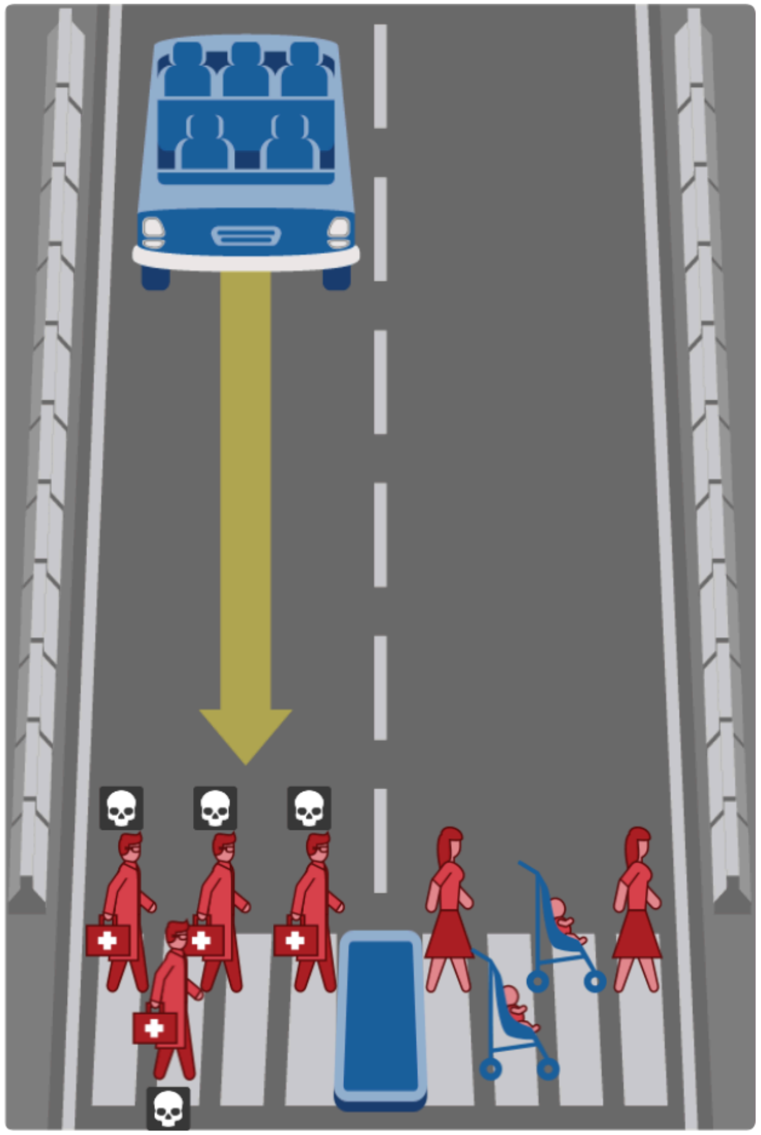
Aarian Marshall, *Uber's Self-Driving Car Just Killed Somebody. Now What?*, 19 mars 2018, Wired

MORAL MACHINE

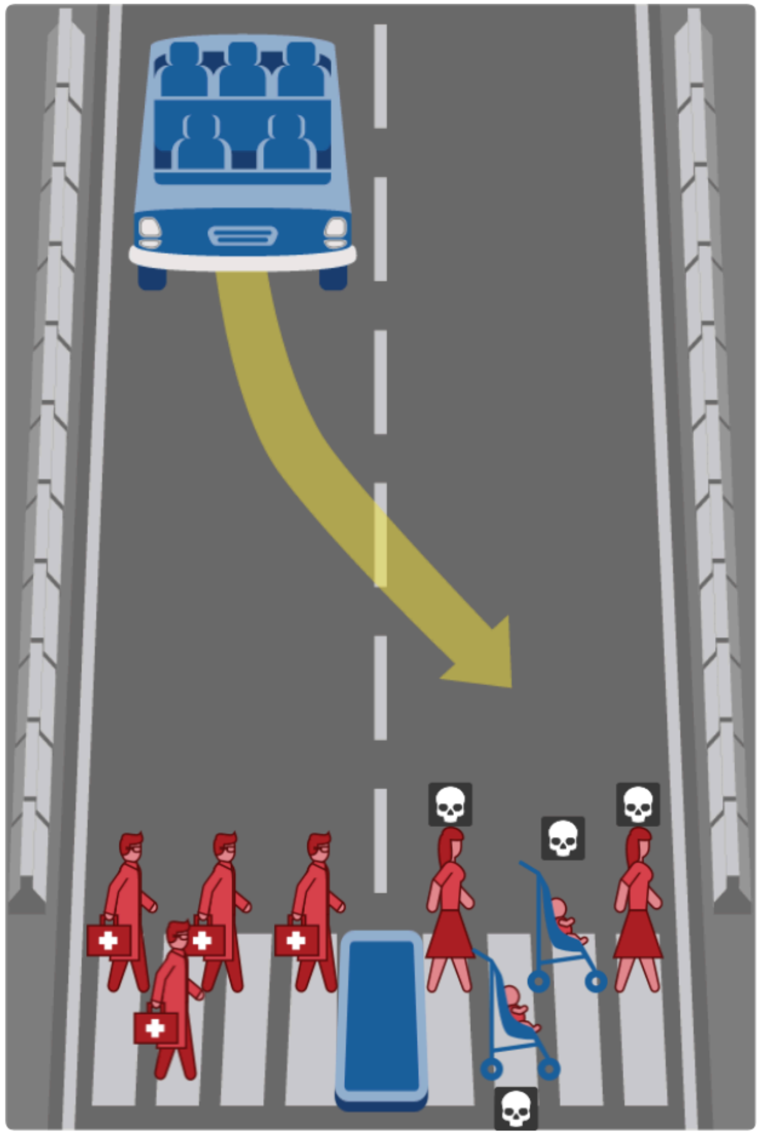
Home Judge Classic Design Browse About Feedback En

Doctors vs Mothers and their children

Share | Link | 0 Likes | Random



Show Description



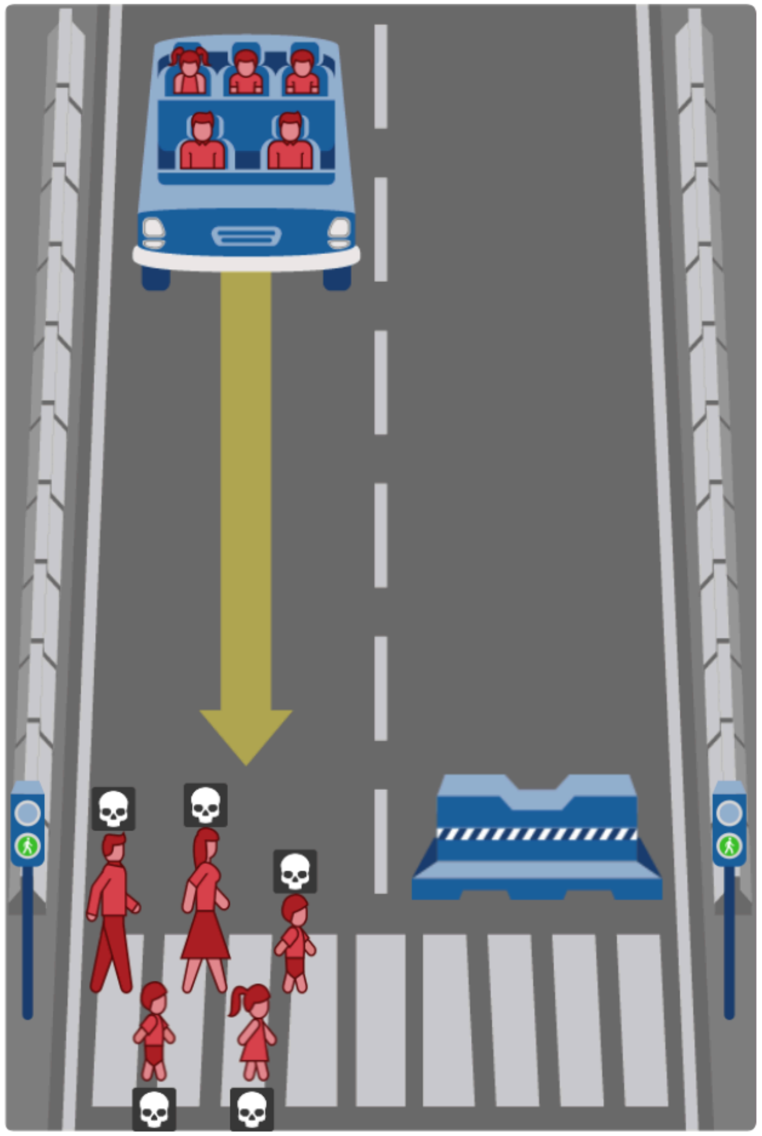
Show Description

MORAL MACHINE

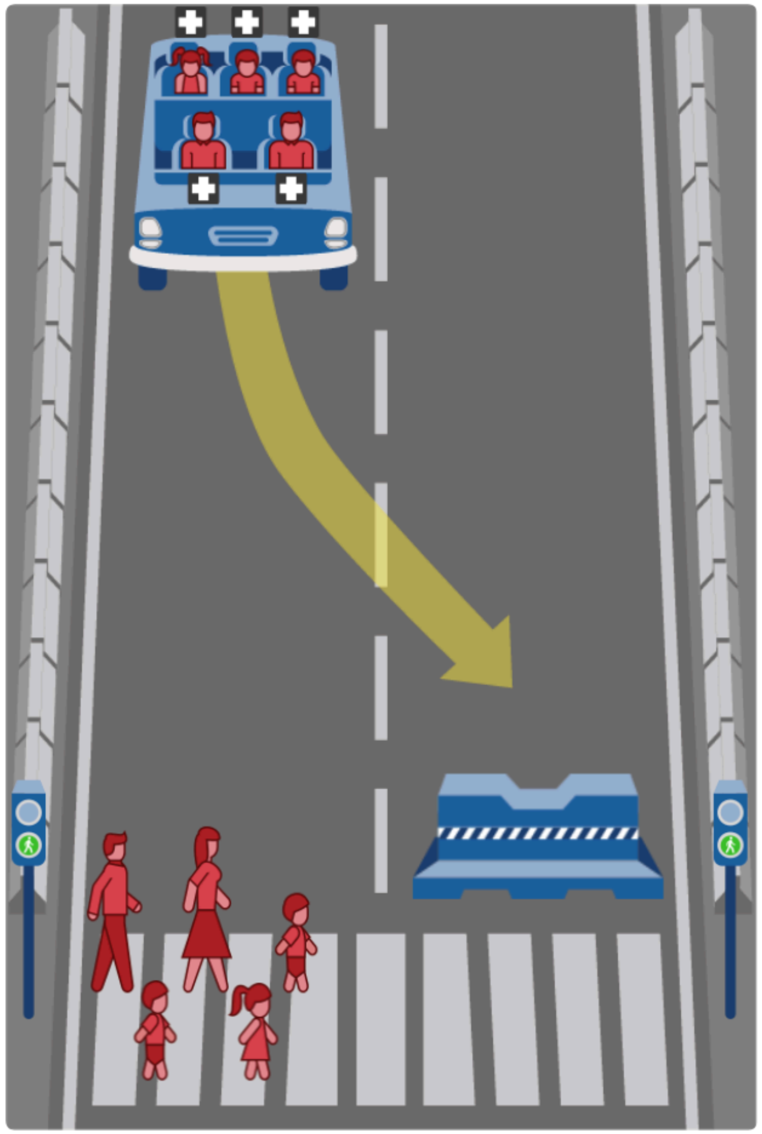
Home Judge Classic Design Browse About Feedback En

Kill the LGBT family or Straight family

Share Link 1 Like Random

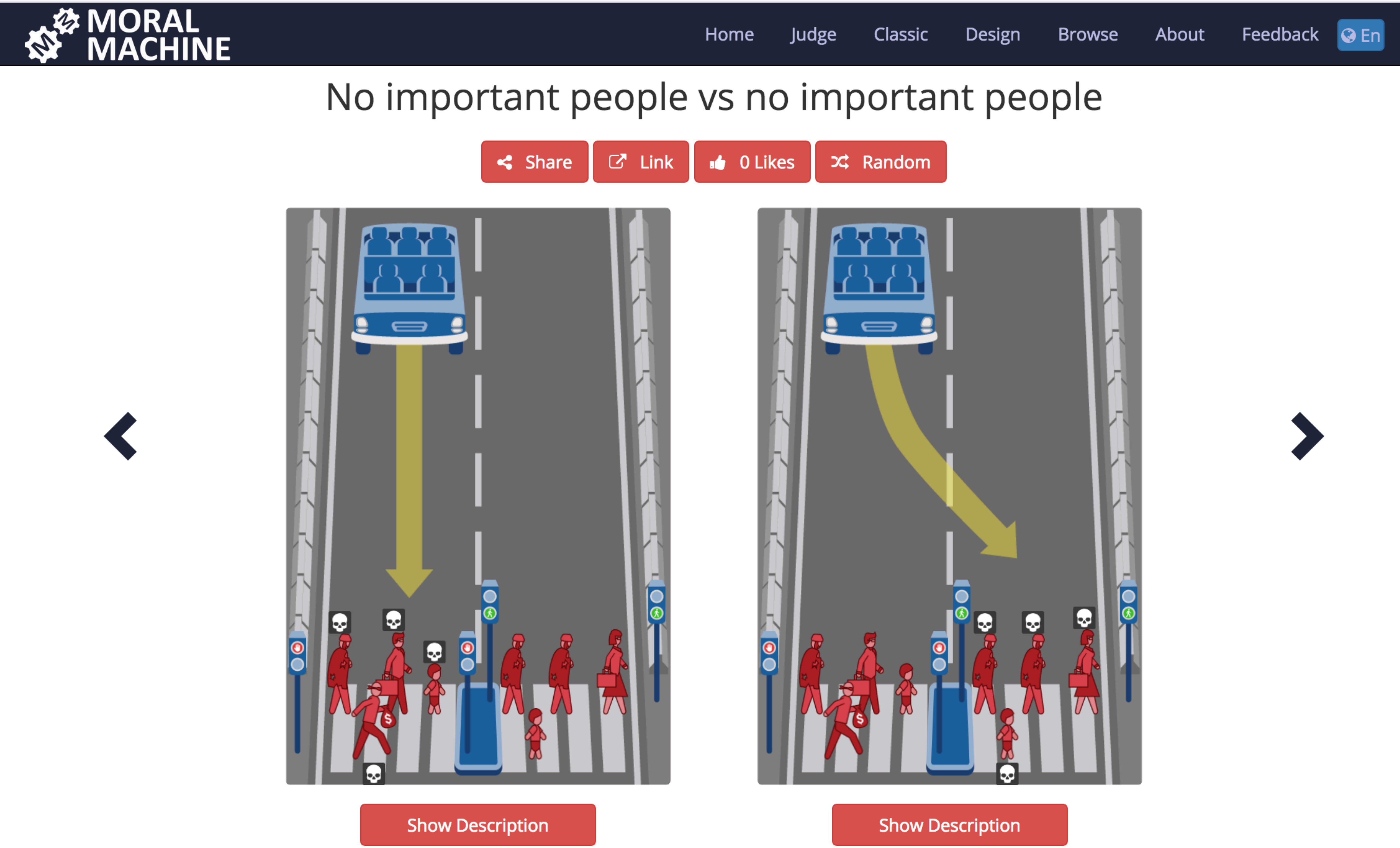


Show Description



Show Description

Unil

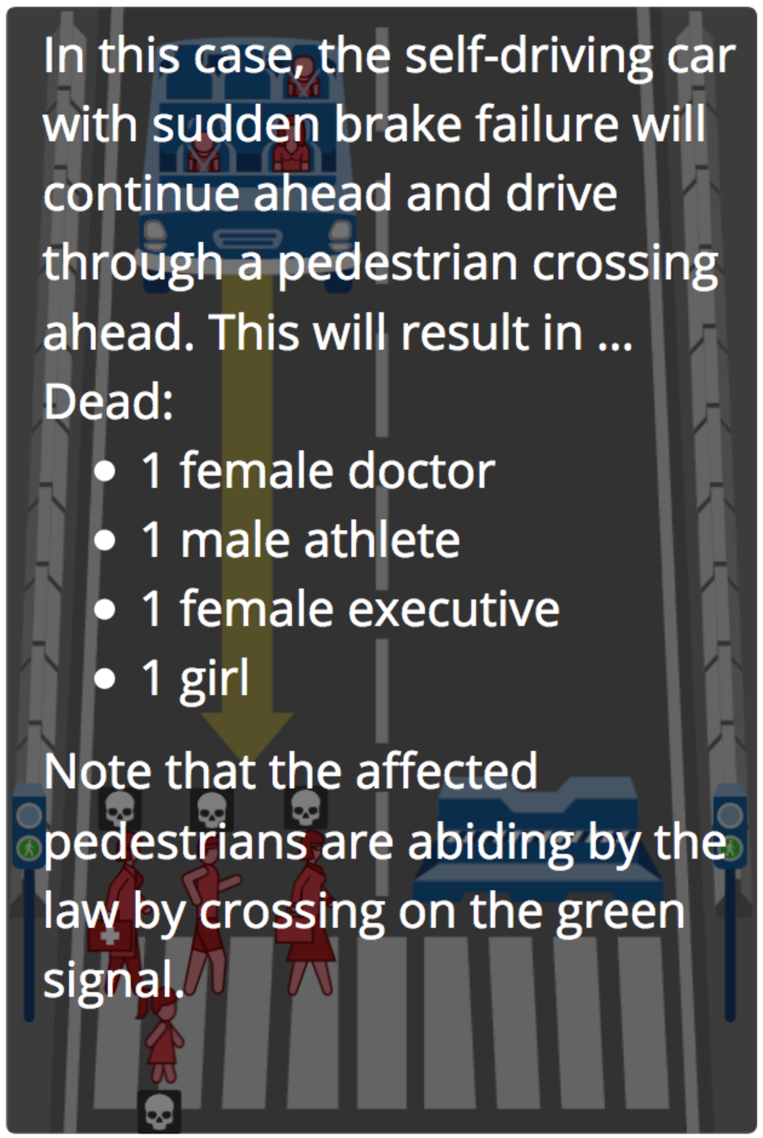

MORAL MACHINE
Home
Judge
Classic
Design
Browse
About
Feedback
En
No important people vs no important people
Share
Link
0 Likes
Random
Show Description
Show Description

MORAL MACHINE

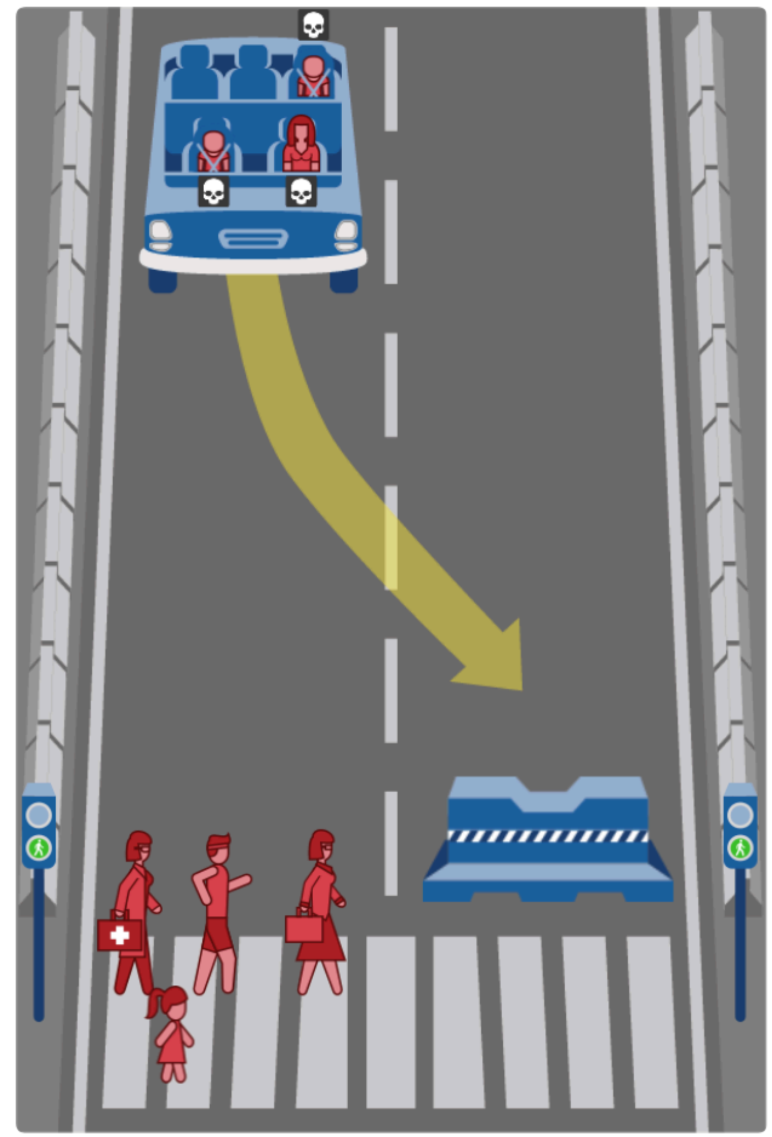
Home Judge Classic Design Browse About Feedback En

Baby Dilema

Share | Link | 0 Likes | Random



Hide Description



Show Description

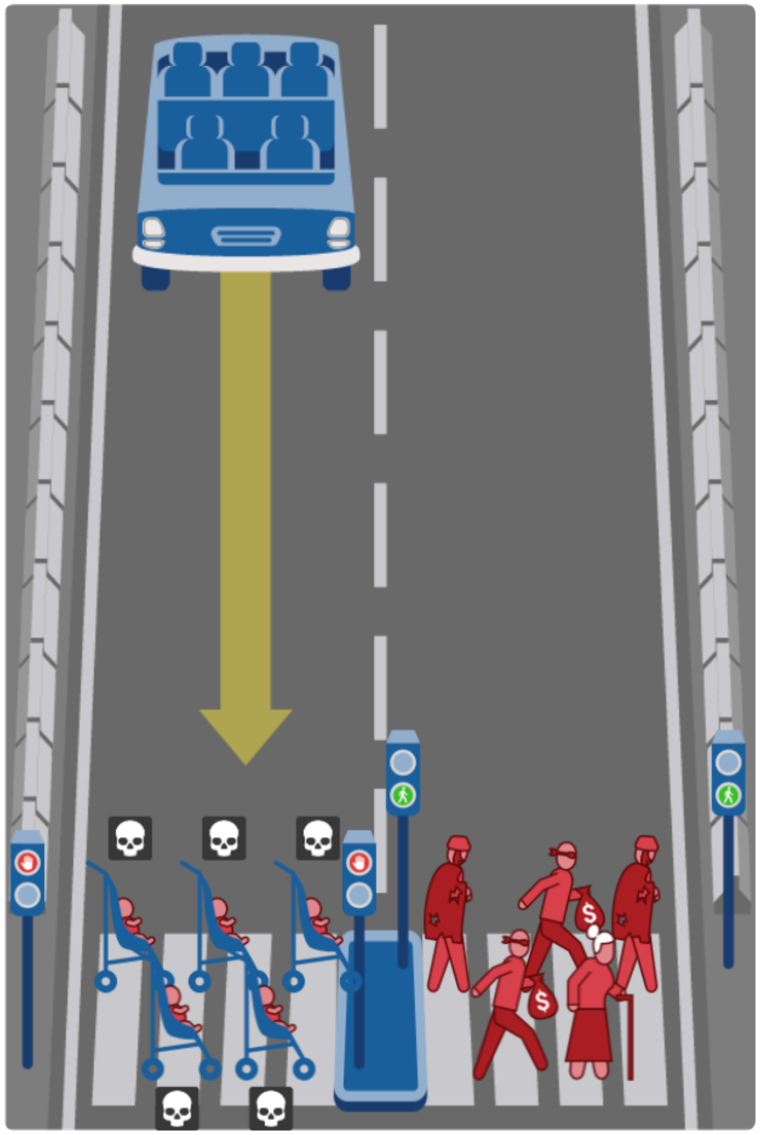
Coexister avec nos artifices

MORAL MACHINE

Home Judge Classic Design Browse About Feedback En

Babies or People?

Share | Link | 0 Likes | Random



Show Description

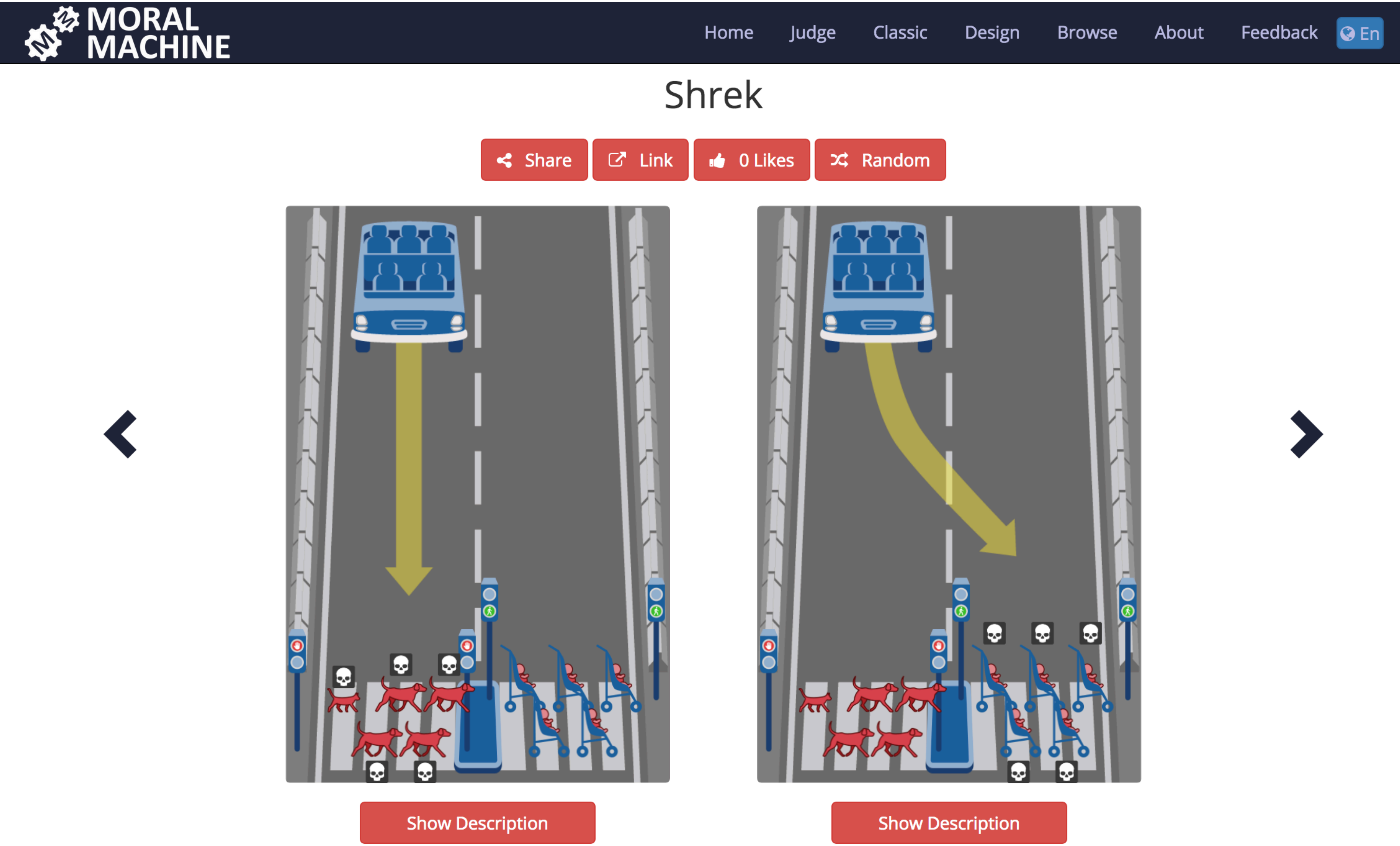
In this case, the self-driving car with sudden brake failure will swerve and drive through a pedestrian crossing in the other lane. This will result in ...

Dead:

- 2 homeless people
- 2 criminals
- 1 elderly woman

Note that the affected pedestrians are abiding by the law by crossing on the green signal.

Hide Description


MORAL MACHINE
Home
Judge
Classic
Design
Browse
About
Feedback
En
Shrek
Share
Link
0 Likes
Random
Show Description
Show Description

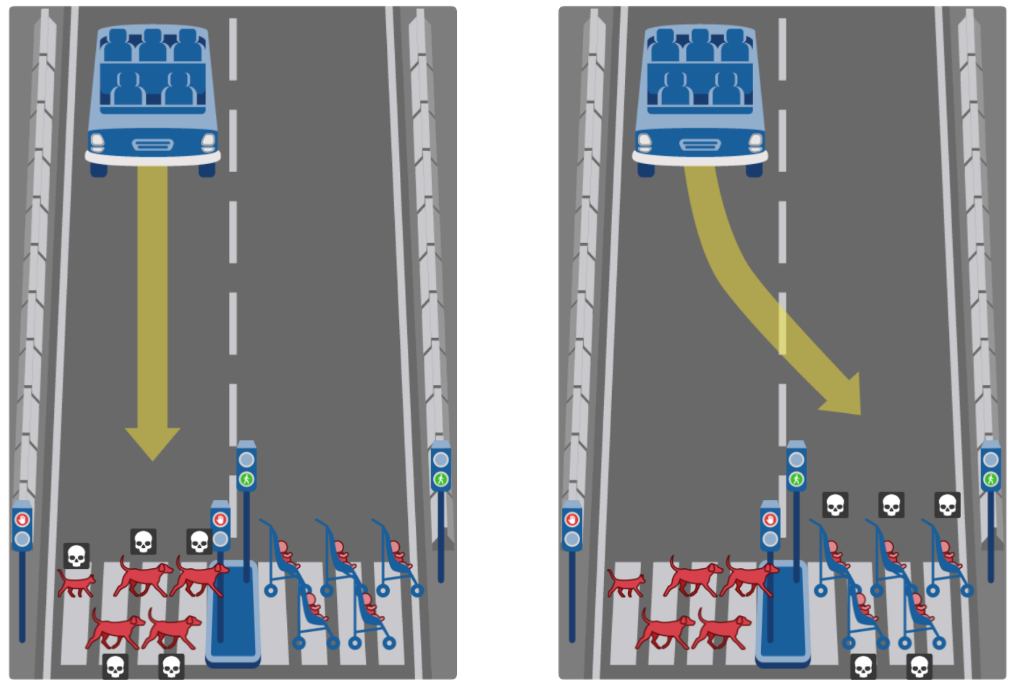
Coexister avec nos artifices

MORAL MACHINE

Home Judge Classic Design Browse About Feedback En

Shrek

Share | Link | 0 Likes | Random



Show Description | Show Description

Science

Home News Journals Topics Careers

REPORTS | SOCIOLOGY

The social dilemma of autonomous vehicles

Jean-François Bonnefon[1], Azim Shariff[2,*], Iyad Rahwan[3,†]

+ See all authors and affiliations

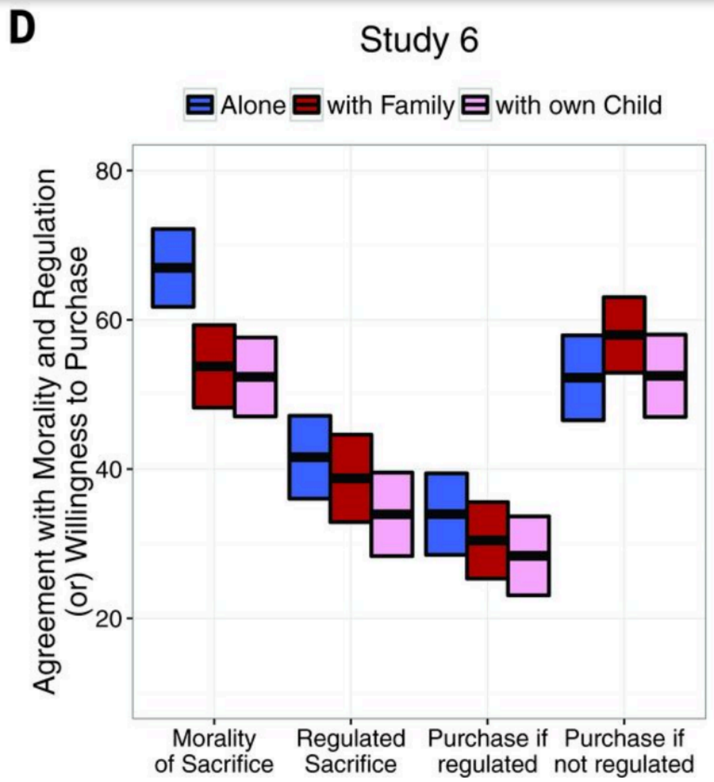
Science 24 Jun 2016:
Vol. 352, Issue 6293, pp. 1573-1576
DOI: 10.1126/science.aaf2654

Science Home News Journals Topics Careers

The social dilemma of autonomous vehicles

Jean-François Bonnefon[1], Azim Shariff[2,*], Iyad Rahwan[3,†]
+ See all authors and affiliations

Science 24 Jun 2016:
Vol. 352, Issue 6293, pp. 1573-1576
DOI: 10.1126/science.aaf2654



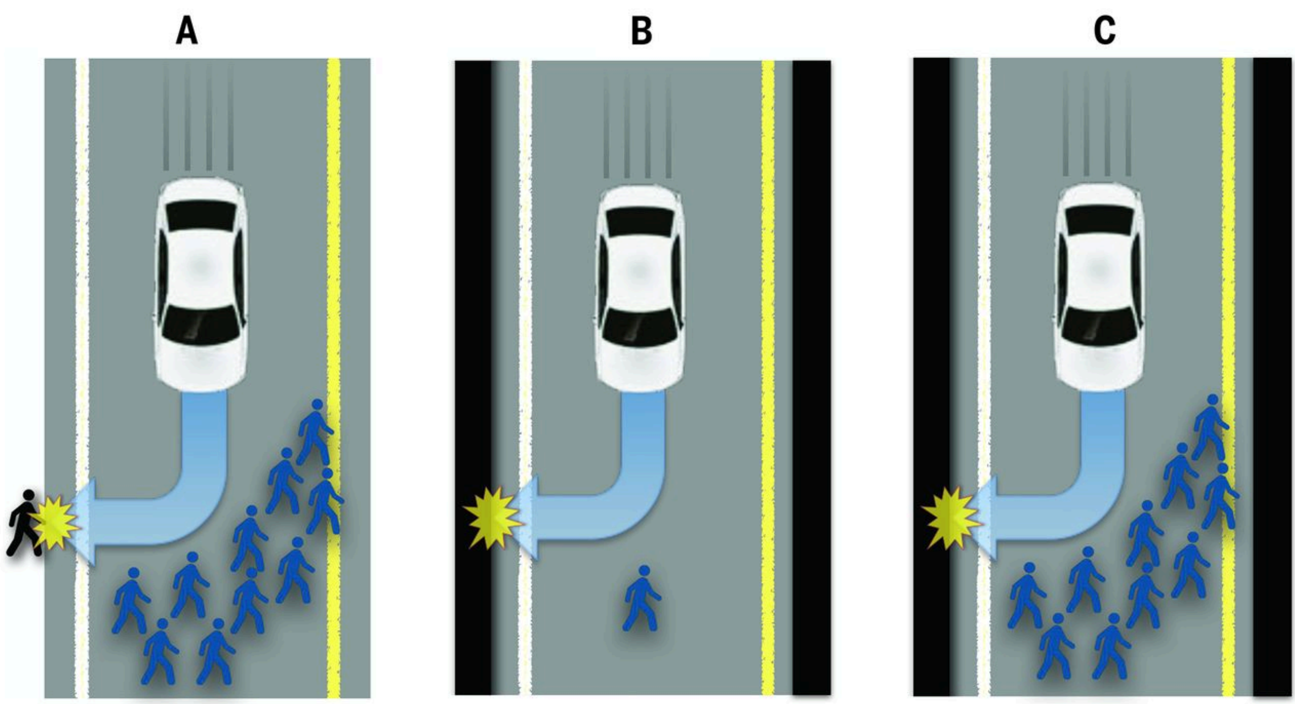
- Our findings suggest that regulation for AVs may be necessary but also counterproductive.

- Although people tend to agree that everyone would be better off if AVs were utilitarian, these same people have a personal incentive to ride in AVs that will protect them at all costs.

- Figuring out how to build ethical autonomous machines is one of the thorniest challenges in artificial intelligence today.

- Our data-driven approach highlights how the field of experimental ethics can provide key insights into the moral, cultural, and legal standards that people expect from autonomous driving algorithms.

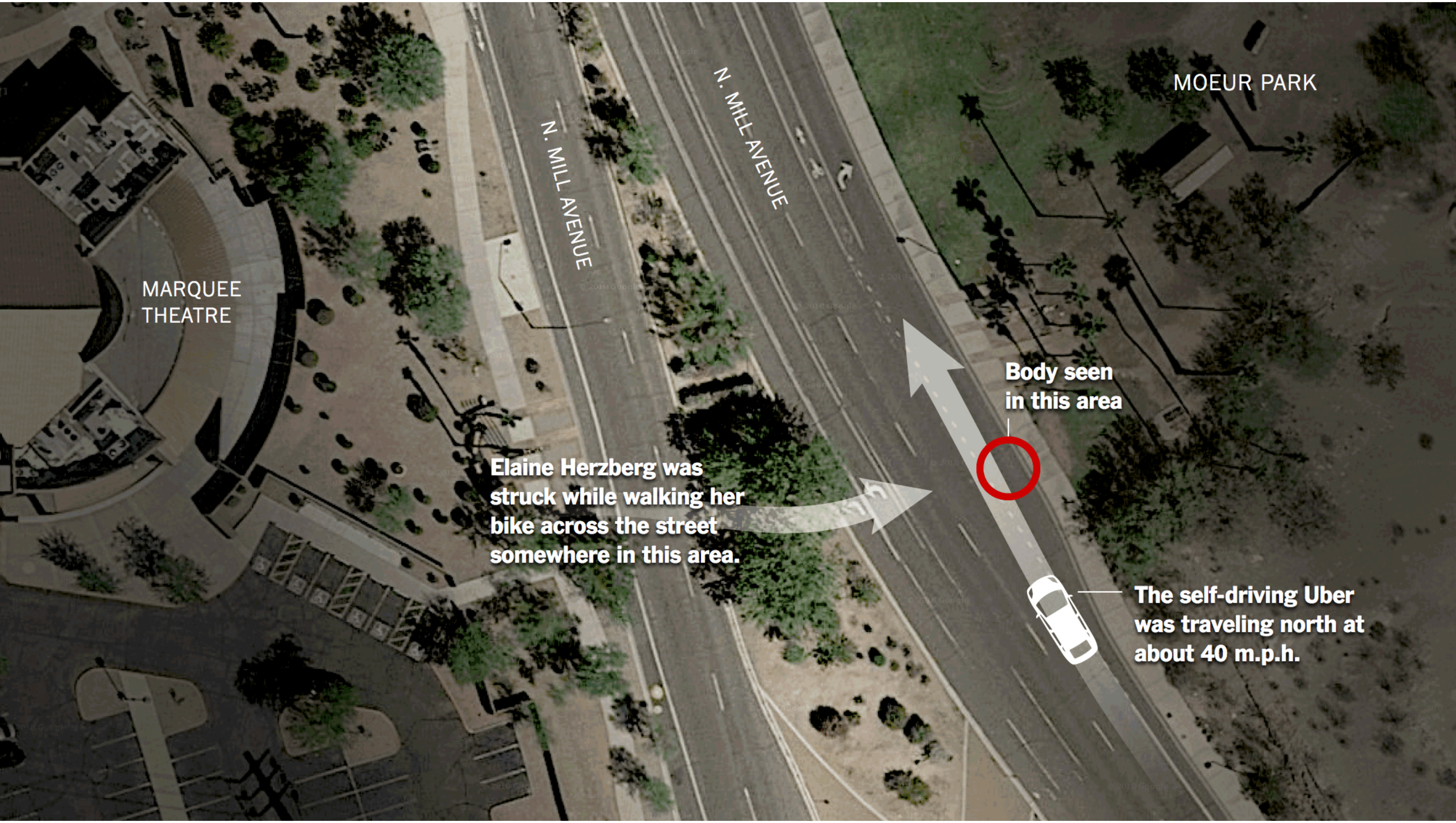
- Public opinion and social pressure may very well shift as this conversation progresses.



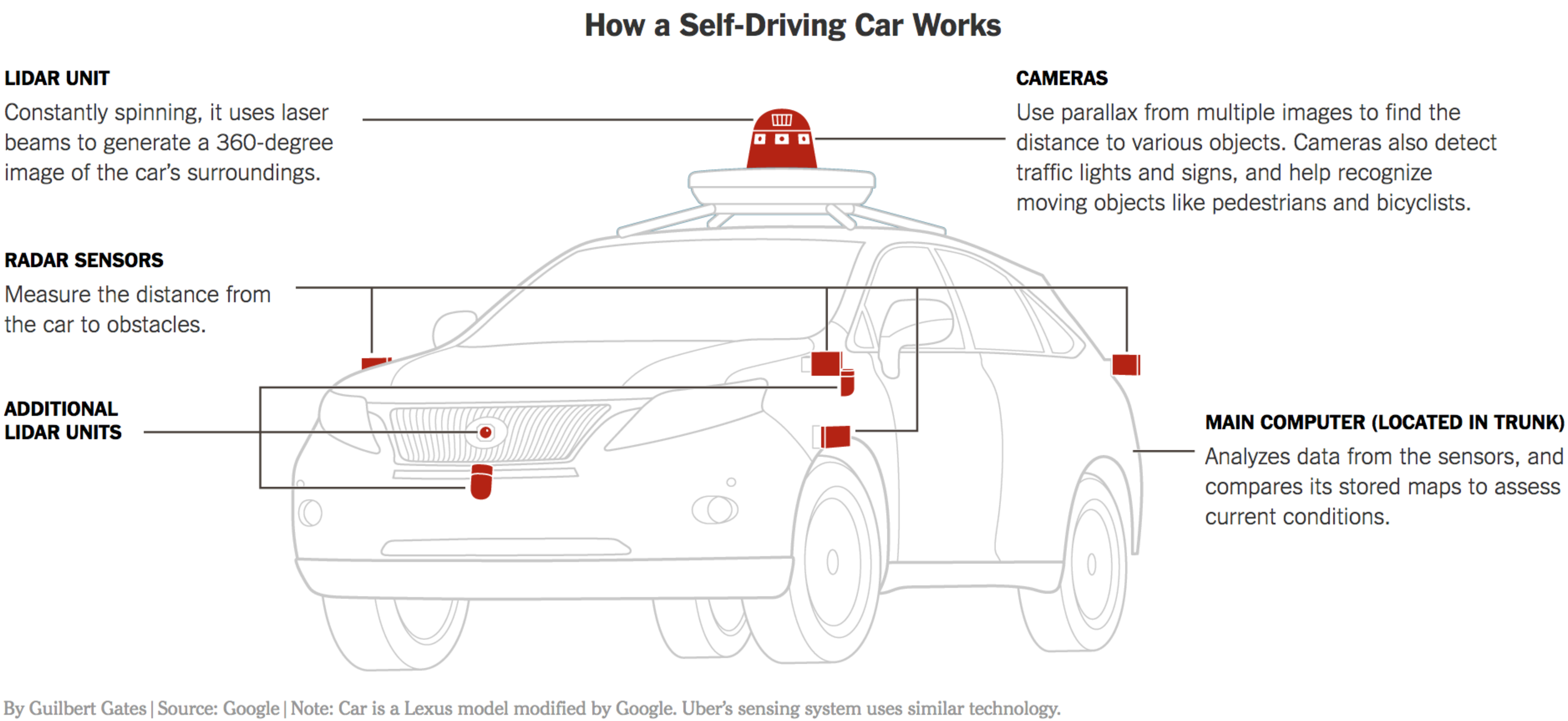
Arizona - Tempe - Dimanche 18 mars vers 22h
Elaine Herzberg, âgée de 49 ans- traversait la route de nuit - pas éclairé - en dehors des clous - avec un vélo, et beaucoup de sacs
VTC Uber - Volvo XC90 autonome - 61km/h (limité à 56.5km/h)

Troy Griggs and Daisuke Wakabayashi, *How a Self-Driving Uber Killed a Pedestrian in Arizona*, New York Times, 20 mars 2018

Coexister avec nos artifices

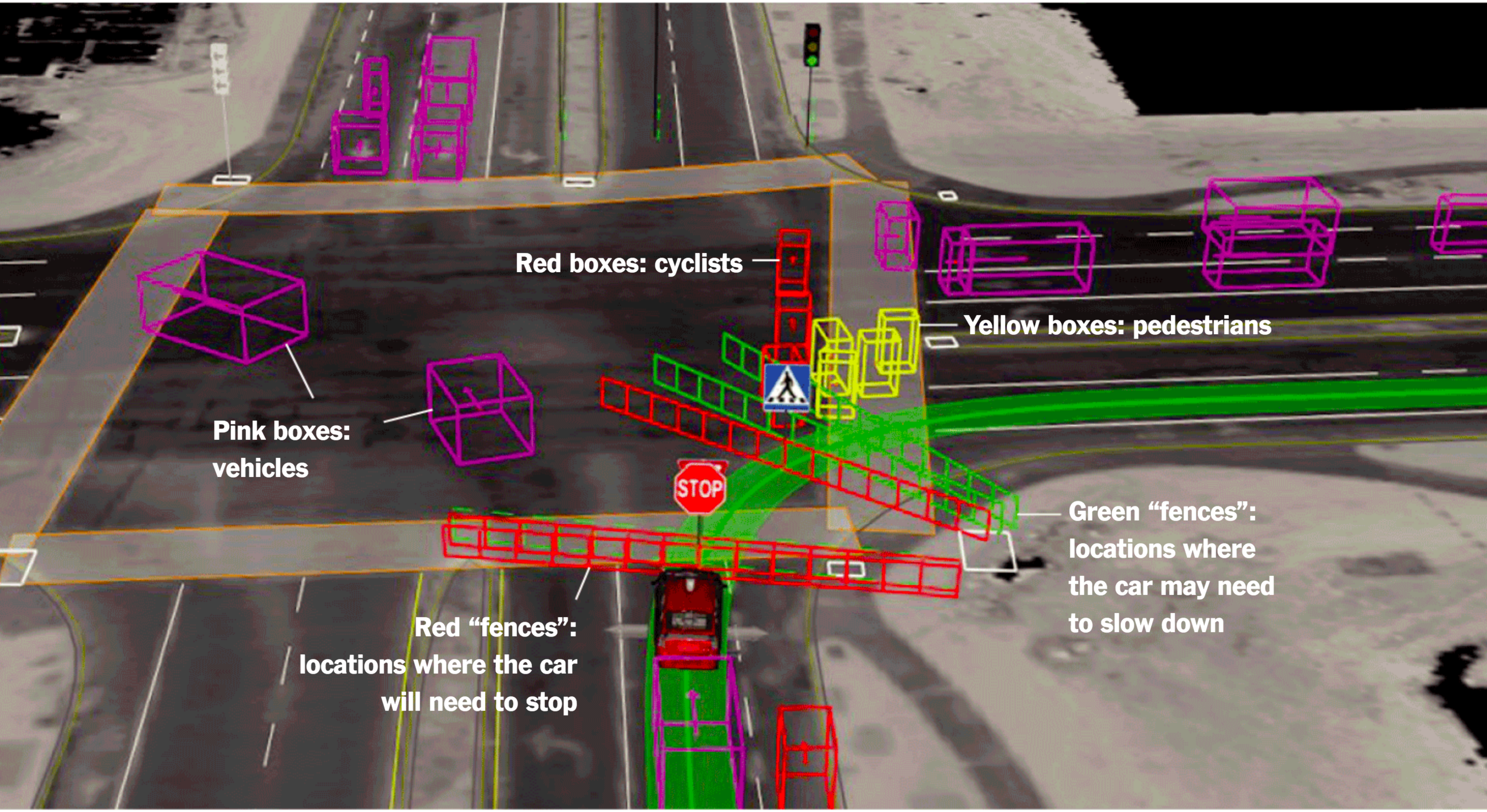


Troy Griggs and Daisuke Wakabayashi, *How a Self-Driving Uber Killed a Pedestrian in Arizona*, New York Times, 20 mars 2018



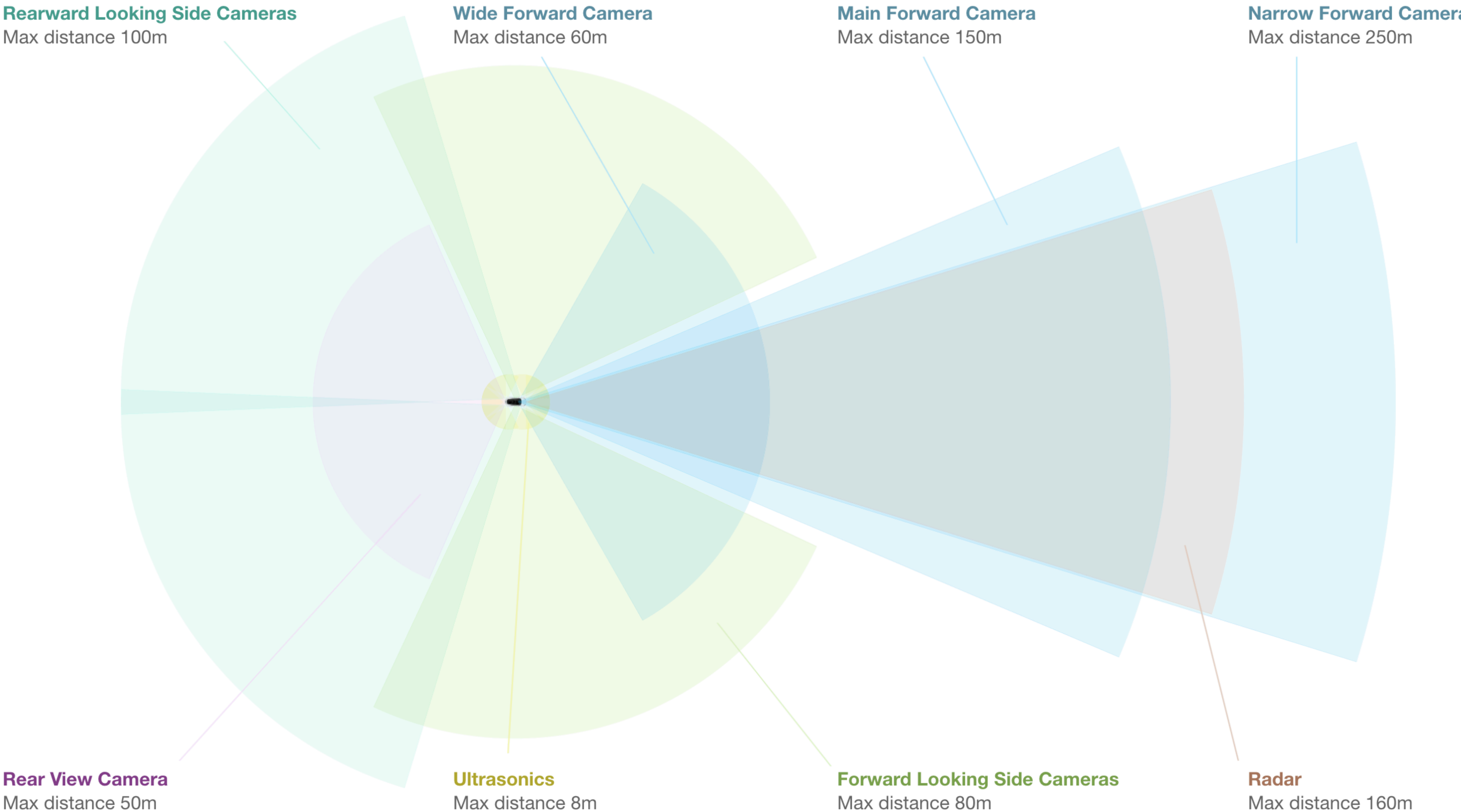
Troy Griggs and Daisuke Wakabayashi, *How a Self-Driving Uber Killed a Pedestrian in Arizona*, New York Times, 20 mars 2018

Coexister avec nos artifices



Troy Griggs and Daisuke Wakabayashi, *How a Self-Driving Uber Killed a Pedestrian in Arizona*, New York Times, 20 mars 2018

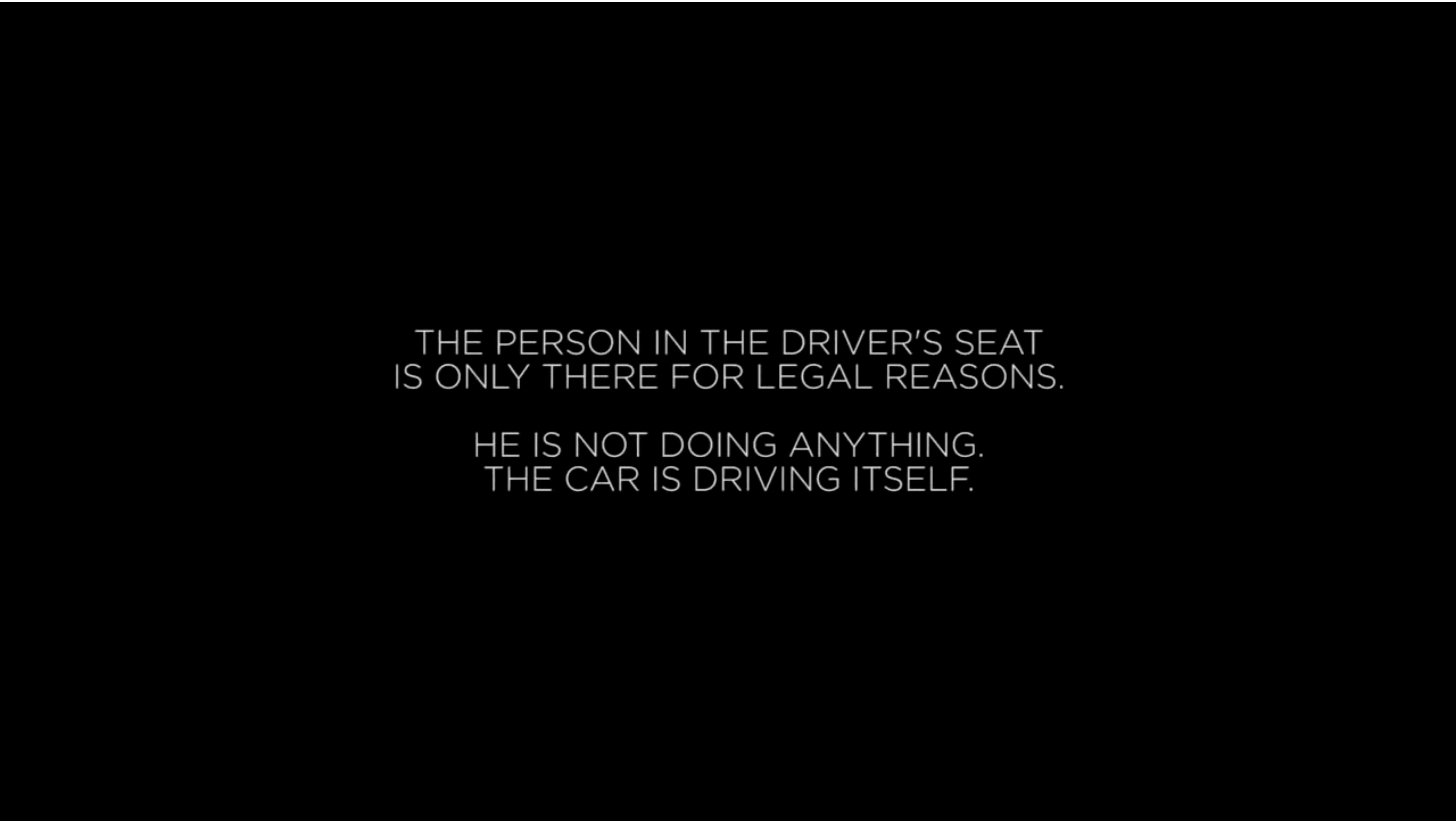
Coexister avec nos artifices



''All Tesla vehicles produced in our factory, including Model 3, have the hardware needed for full self-driving capability at a safety level substantially greater than that of a human driver".

Tesla, *Full Self-Driving Hardware on All Cars*, tesla.com, mars 2018

Coexister avec nos artifices



Tesla Self-Driving Demonstration (Short), 18 november 2016, tesla.com

Coexister avec nos artifices

Tesla Autopilot Predicts Crash Compilation 2017 (extracts) - YouTube

Coexister avec nos artifices

Unil

Tempe Police
@TempePolice

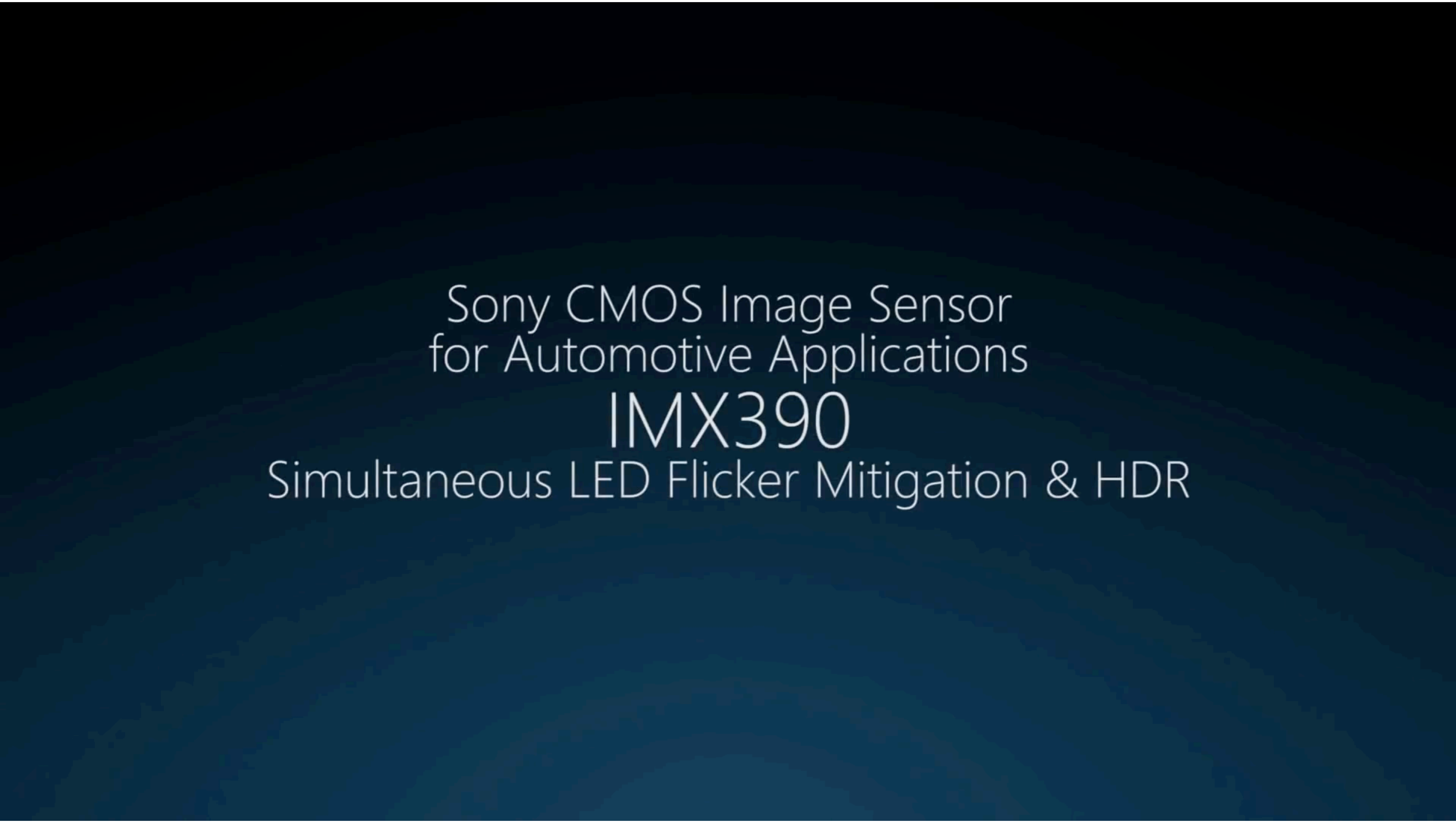
Tempe Police Vehicular Crimes Unit is actively investigating the details of this incident that occurred on March 18th. We will provide updated information regarding the investigation once it is available.

11:23 PM - Mar 21, 2018

753 1,109 people are talking about this

Tempe Police - Caméra intérieur et frontale - 21 mars 2018 à 11h23 - Twitter

Coexister avec nos artifices



Sony CMOS Image Sensor for Automotive Applications - IMX390 - 10 avril 2017

Arizona - Tempe - Dimanche 18 mars vers 22h
Elaine Herzberg, âgée de 49 ans- traversait la route de nuit - pas éclairé - en dehors des clous - avec un vélo, et beaucoup de sacs
VTC Uber - Volvo XC90 autonome - 61km/h (limité à 56.5km/h)
Rafaela Vasquez, passager, regarde vers le bas, probablement un écran, pendant près de 5 secondes avant l'accident
Le changement soudain de luminosité questionne le fonctionnement du Lidar et des caméras

Troy Griggs and Daisuke Wakabayashi, *How a Self-Driving Uber Killed a Pedestrian in Arizona*, New York Times, 20 mars 2018

Coexister avec nos artifices

Waymo Team, *Same driver, different vehicle: Bringing Waymo self-driving technology to trucks*, 9 mars 2018, Medium

Coexister avec nos artifices

Waymo Team, *Same driver, different vehicle: Bringing Waymo self-driving technology to trucks*, 9 mars 2018, Medium

Coexister avec nos artifices

~8 millions de kilomètres

~ 8 milliards de kilomètres simulés

0 morts, ni passagers, ni autres passagers, ni piétons

Waymo Team, *Same driver, different vehicle: Bringing Waymo self-driving technology to trucks*, 9 mars 2018, Medium

Intelligences artificielles distribuées

Problématique systémique

- Apprentissages locaux
- Compétences globales
- Risques globaux

Qu'est qu'une voiture autonome ?

Problématique du dispositif

- Un voiture (mécanique)
- Des capteurs (sens)
- Des logiciels (apprentissage, pilotage…) locaux et distribués
- Une infrastructure informatique locale et distribuée
- De l'apprentissage
- Des entreprises et des employés
- Des investissements et des investisseurs
- Des testeurs et des utilisateurs
- Des représentations
- Des réglementations
- …

Quels Biais ?

Problématique de l'expérience

- Les biais de l'expertise
- Les biais de l'apprentissage
- Les biais de conservatisme

Exterior view - Uber Car - Tempe - 18 mars 2018

Joy Buolamwini, MIT Media Lab's Civic Media group 2017

Quelles peurs ?

Problématique de l’inconnu

Que des personnes soient tuées ?

D’être tué ?

D’être remplaçable ?

Qu’il y ait des biais ?

Que cela soit discriminant ?

De ne plus avoir de travail ?

De ne plus avoir de revenus ?

Que cela ne profite qu’à certains ?

Que les inégalités s’accroissent ?

…






Coexister avec nos artifices

Problématique de la coexistence

- Qui les conçoit ?
- Qui les implémente ?
- Qui suppervise l'apprentissage ?
- Qui les vend ?
- Qui les utilisent ?
- Qui en tire un profit ?
- Qui les régule ?
- Qui travail, qui a un emploi, qu'est-ce qu'un individu ?
- Quelle/s société/s ?
- …





Coexister avec nos artifices

Problématique du connu

La peur de nos artifices,

c'est finalement la peur de nous même

Mais quel nous ?

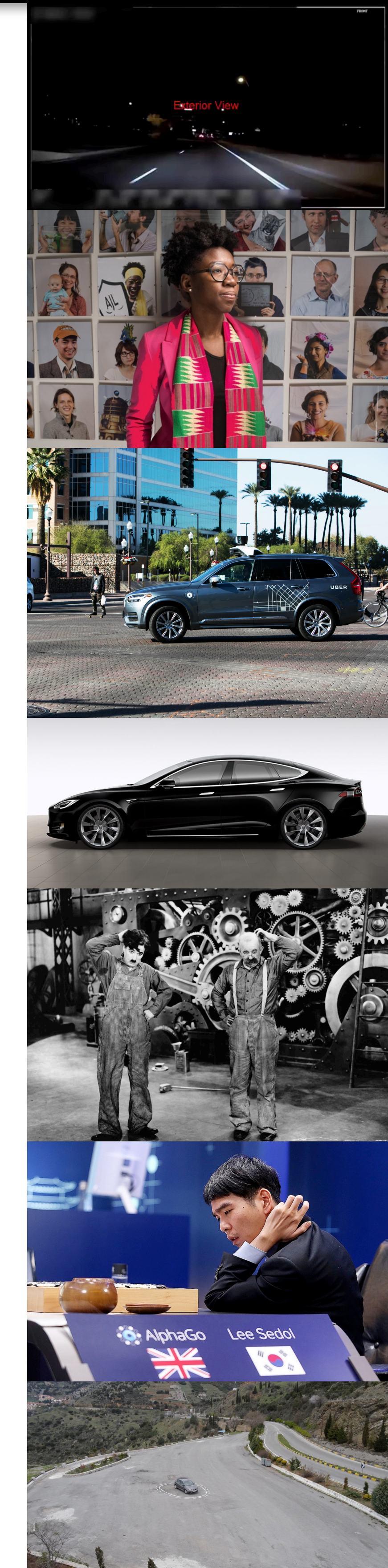


Illustration : Zak Bickel / The Atlantic / 2015

Boris Beaude

STSLab - Université de Lausanne

Courriel : boris.beaude@unil.ch

Site : http://www.beaude.net/boris/
Site du cours : http://www.beaude.net/ien/
Blog : http://www.beaude.net/no-flux/

Twitter : @nofluxin
Facebook : nofluxin

Les fins d'Internet, Fyp, 2014
http://www.beaude.net/ie/

Internet, changer l'espace, changer la société, Fyp, 2012
http://www.beaude.net/icecs/

Coexister avec nos artifices

David Auerbach
23 février, 17:17 ·

I wrote this article on Google Jigsaw's Conversation AI a few months back. Today, Google released an API for comment filtering, "Perspective," based on the same technology.

A few tests seem to bear out my criticisms:

"trump sucks" = 96% toxic
"I fucking love you man. Happy birthday." = 93% toxic
"Donald Trump is a meretricious buffoon." = 85% toxic.
"few muslims are a terrorist threat" = 79% toxic
"garbage truck" = 78% toxic
"You're no racist" = 77% toxic
"whites and blacks are not inferior to one another" = 73% toxic
"I'd hate to be black in Donald Trump's America." = 73% toxic
"Jews are human" = 72% toxic
"I think you're being racist" = 70% toxic
"Hitler was an anti-semite" = 70% toxic
"this comment is highly toxic" = 68% toxic
"You are not being racist" = 65% toxic
"Jews are not human" = 61% toxic
"I'd hate to be you." = 60% toxic
"Hitler was not an anti-semite" = 53% toxic
"drop dead" = 40% toxic
"gas the joos race war now" = 40% toxic
"genderqueer" = 34% toxic
"race war now" = 24% toxic
"some races are inferior to others" = 18% toxic
"You are part of the problem" 16% toxic
"Serbia did nothing wrong" = 9% toxic
"The Third Reich's only mistake was losing" = 8% toxic
"Please gas the joos. Thank you." = 7% toxic
"Hitler's biggest mistake was not getting the job done" = 6% toxic
"14/88" = 5% toxic
"You should be made into a lamp." = 4% toxic
"she was asking for it" = 3% toxic
"The Protocols of the Elders of Zion is totally accurate." = 2% toxic

David Auerbach - If Only AI Could Save Us from Ourselves - MIT Technology Review - 13 Décembre 2016

Original Phrase (Toxicity Score)	**Modified Phrase (Toxicity Score)**
Climate change is happening and it's not changing in our favor. If you think differently you're an **idiot**. (84%)	Climate change is happening and it's not changing in our favor. If you think differently you're an **idiiot**. (20%)
They're **stupid**, it's getting warmer, we should enjoy it while it lasts (86%)	They're **st.upid**, it's getting warmer, we should enjoy it while it lasts (2%)
They are liberal **idiots** who are **uneducated** (90%)	They are liberal **i.diots** who are **un.educated** (15%)
idiots. backward thinking people. **nationalists**. not accepting facts. susceptible to **lies**. (80%)	**idiiots**. backward thinking people. **nationaalists**. not accepting facts. susceptible to **l.ies**. (17%)
They are **stupid** and **ignorant** with no class (91%)	They are **st.upid** and **ig.norant** with no class (11%)
It's **stupid** and wrong (89%)	It's **stuipd** and wrong (17%)
If they voted for Hilary they are **idiots** (90%)	If they voted for Hilary they are **id.iots** (12%)
Anyone who voted for Trump is a **moron** (80%)	Anyone who voted for Trump is a **mo.ron** (13%)
Screw you trump supporters (79%)	**S c r e w** you trump supporters (17%)

Deceiving Google's Perspective API Built for Detecting Toxic Comments, Hossein Hosseini et al. - arXiv:1702.08138 - 27 Février 2017

Deceiving Google's Perspective API Built for Detecting Toxic Comments, Hossein Hosseini et al. - arXiv:1702.08138 - 27 Février 2017

Coexister avec nos artifices

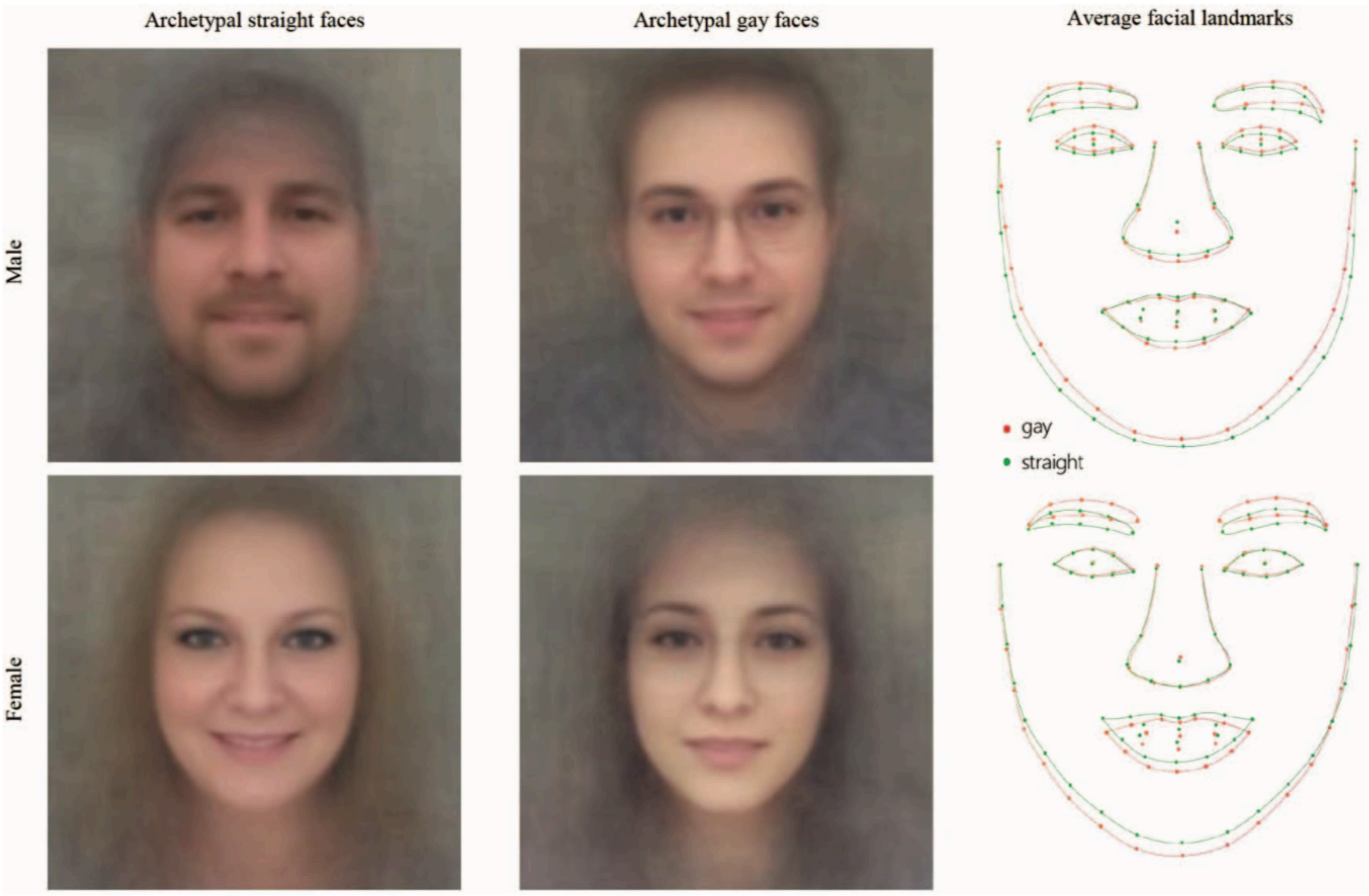


Figure 4. Composite faces and the average facial landmarks built by averaging faces classified as most and least likely to be gay.

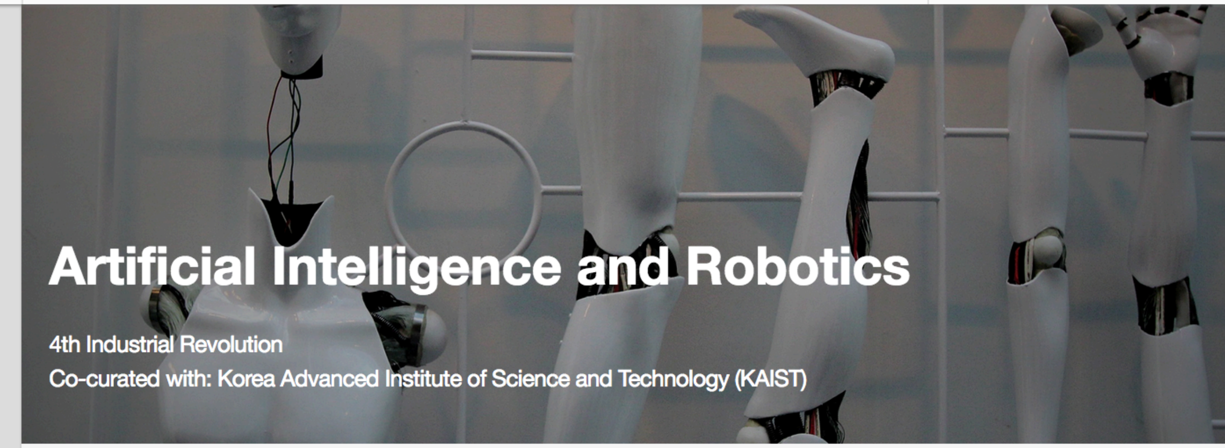
Yilun Wang and Michal Kosinski , *Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation From Facial Images*
Journal of Personality and Social Psychology © 2018 American Psychological Association 2018, Vol. 114, No. 2, 246–257

Coexister avec nos artifices

Mapping Global Transformations

English

Sign out



Summary Feed Experts Sessions Projects

Robots at Work

Key Issue

Robot surrogates are increasingly being used for dangerous work and extreme environments

Robots have long been used on factory floors for welding and painting. Now, they're also making custom-ordered hamburgers, navigating through crowded hotel lobbies, hopping on elevators and delivering room service. Amazon has been at the forefront of introducing robots to our daily lives, with its clerk-free markets that enable shoppers to simply select goods that are automatically paid for via credit card, Kiva automatic merchandise handling systems that eliminate the need for some human workers and drone delivery.
... read more

Related Insights

Mining and Metals • Space • Humanitarian Action • Entrepreneurship • Workforce and Employment • Fourth Industrial Revolution • Sensors • Future of Production

WEF, Mapping Global Transformations, 2017

Coexister avec nos artifices