

*Visual intelligence... for
humans and machines*

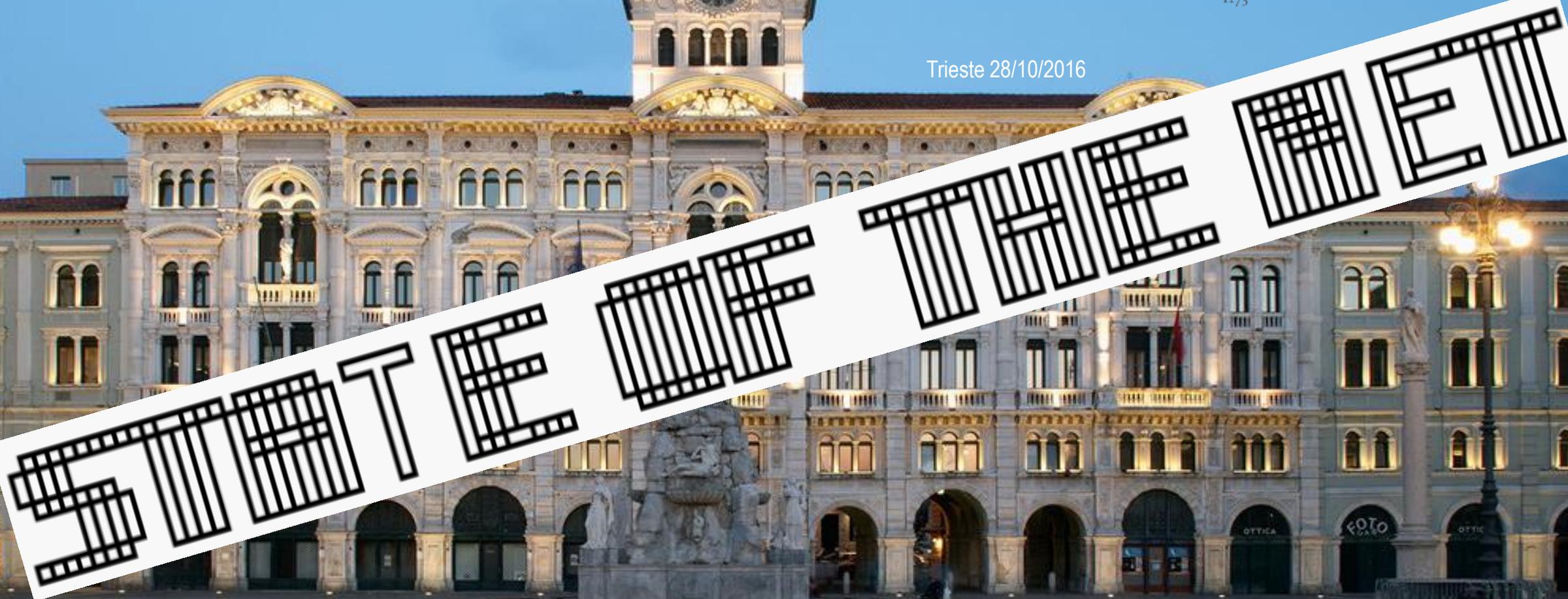
Rita Cucchiara



UNIMORE

UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

Trieste 28/10/2016

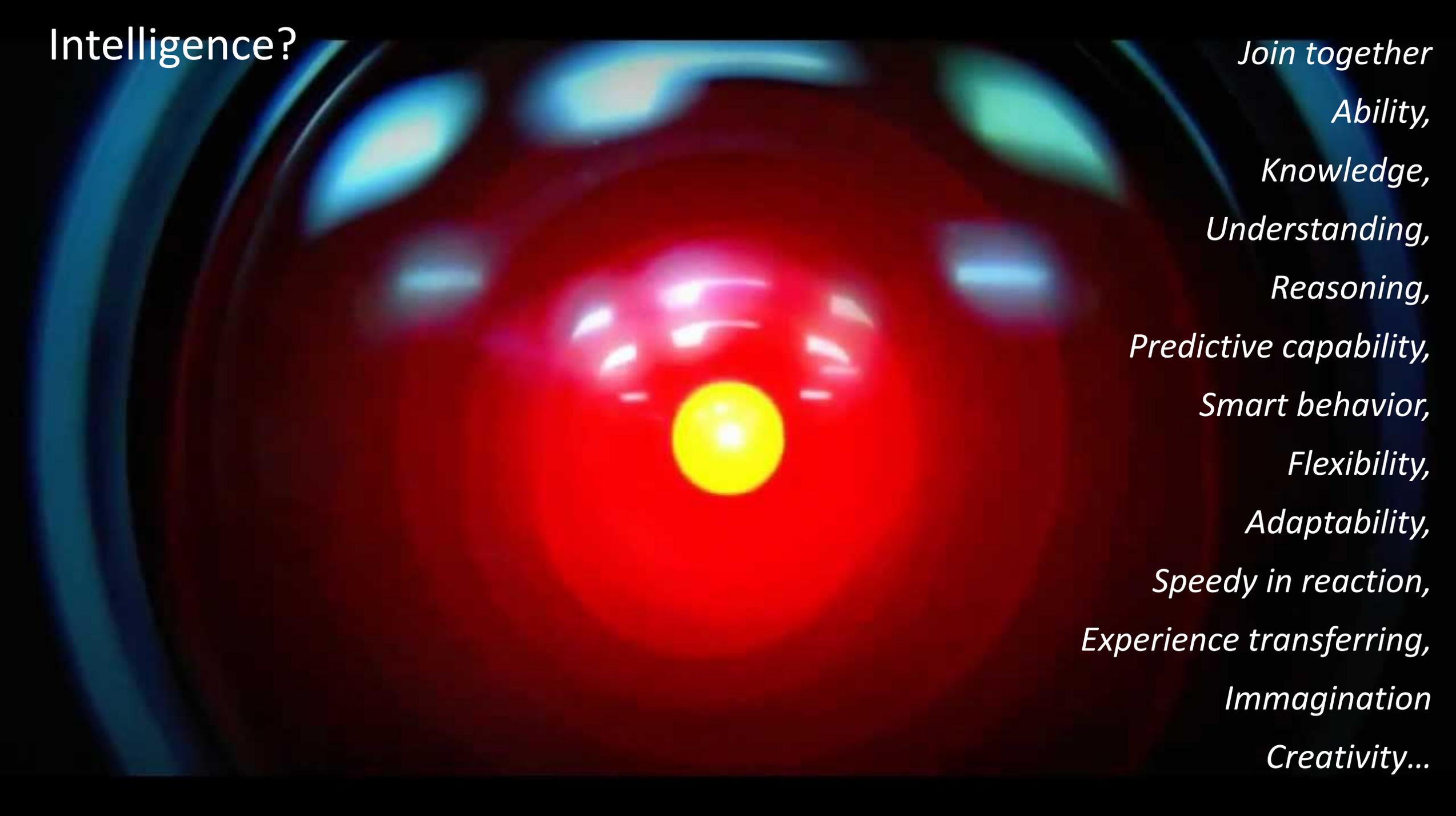


Intelligence ?

*Many forms of natural intelligence due to evolution.
AI is in the game now, as a new form of (quickly) evolving intelligence.*

But actually, IT IS ONLY SOFTWARE

Intelligence?



*Join together
Ability,
Knowledge,
Understanding,
Reasoning,
Predictive capability,
Smart behavior,
Flexibility,
Adaptability,
Speedy in reaction,
Experience transferring,
Imagination
Creativity...*

Visual Intelligence ?

*Artificial
Intelligence*

*Pattern
Recognition*

*Machine (Deep)
Learning*

*Computer
Vision*

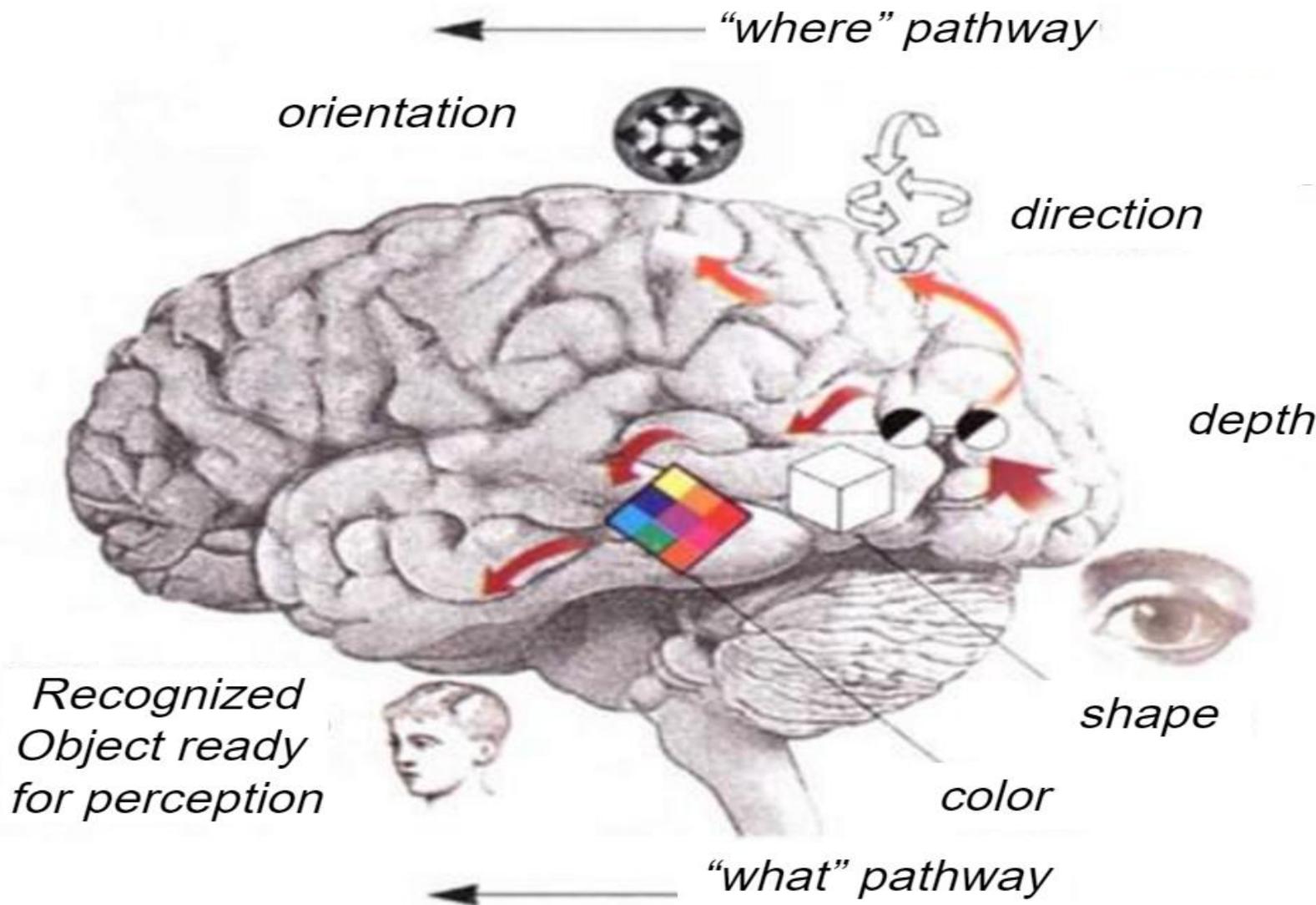


*Intelligence uses **senses** as input data,
memory for learned facts, and **brain** connections for learned lessons..
Intelligence makes data association and recognition of unknown patterns
Intelligence uses learning and reasoning to **predict facts** and **eventually to act**.*



Visual intelligence starts with eyes

Intelligence ?



Cat

Neurons : 0,8 billion
Synapses: 10 billion

Human

Neurons: 80 billion
In the cortex: 20-28 billion*
Synapses: 100 billion

Retinae cones hodd: 150 million

Cortex for vision: 30%

Cortex for touch: 8%

Cortex for hearing: 3%

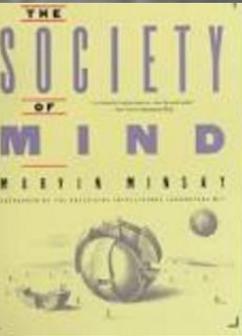
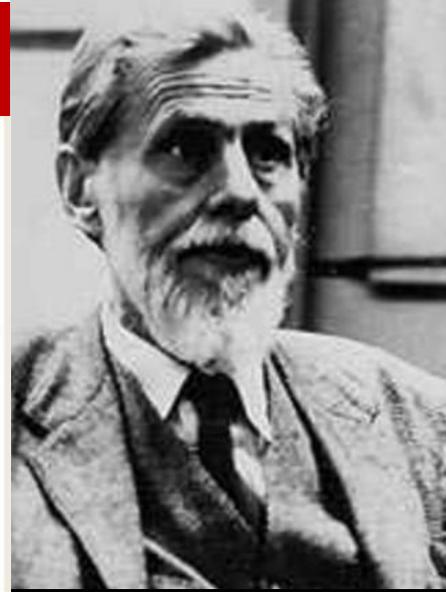
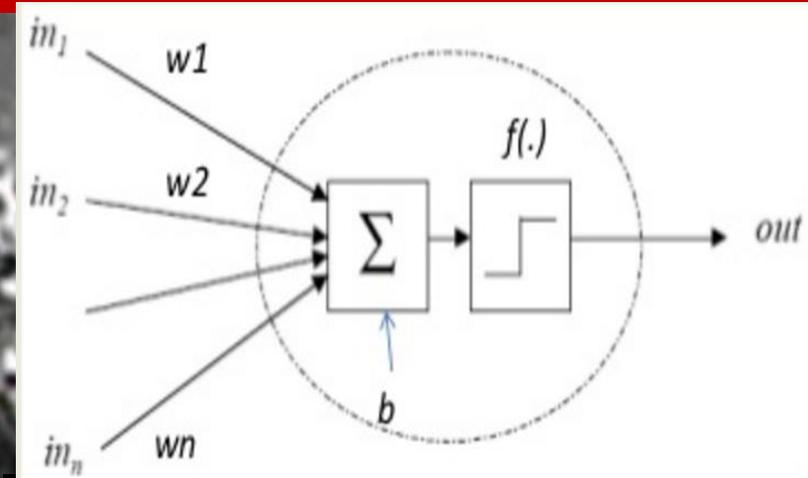
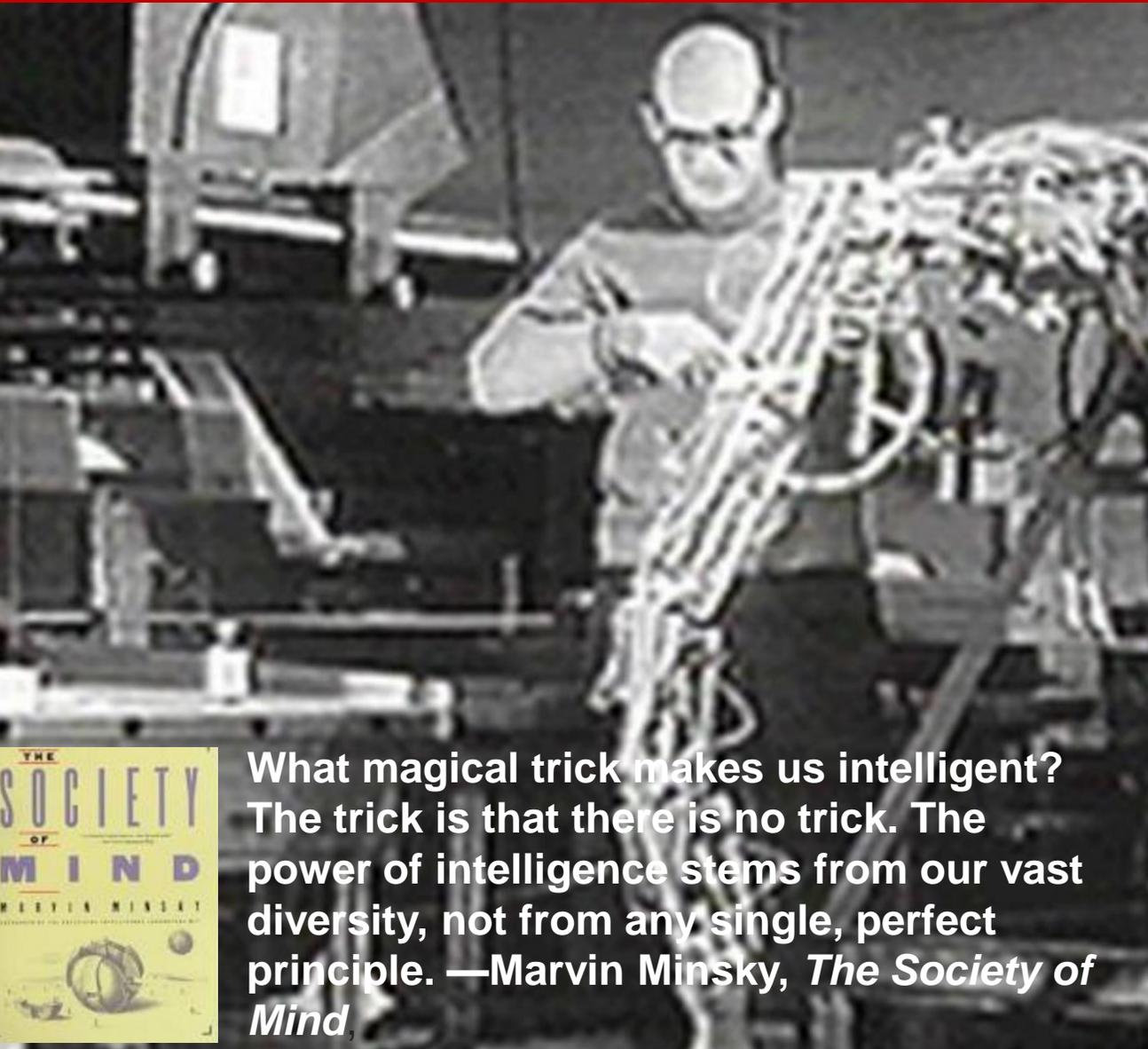
Each optical nerves: 1 million

Semir Zeki «*Inner vision: an exploration of Art and Brain*» 1999

*Suzana Herculano-Houzel, *The Human Brain in Numbers: A Linearly Scaled-up Primate Brain*, *Frontier in Human Neuroscience* 2009

Visual intelligence needs a brain (and a cortex)

John McCulloch, Marvin Minsky @MIT 1960



What magical trick makes us intelligent? The trick is that there is no trick. The power of intelligence stems from our vast diversity, not from any single, perfect principle. —Marvin Minsky, *The Society of Mind*

- Pattern Recognition
- Cognitive computing
- Media Lab
- Social physics
- Global Intelligence
-



Face Recognition using Eigenfaces CVPR 1991



The Global Artificial Intelligence (GAI) has already been born. Its eyes and ears are the digital devices all around us: credit cards, land use satellites, cell phones, and the pecking of billions of people using the Web. Its central brain is rather like a worm at the moment: nodes that combine some sensors and some effectors, but the whole is far from what you would call a coordinated intelligence.

It is not the Global Artificial Intelligence itself that is worrisome; it is how it is controlled.



"A fascinating look at a new field by one of its principal geeks." —*The Economist*



SOCIAL PHYSICS



2014
strategy+business
Best Business
Book

HOW SOCIAL NETWORKS CAN
MAKE US SMARTER

ALEX PENTLAND

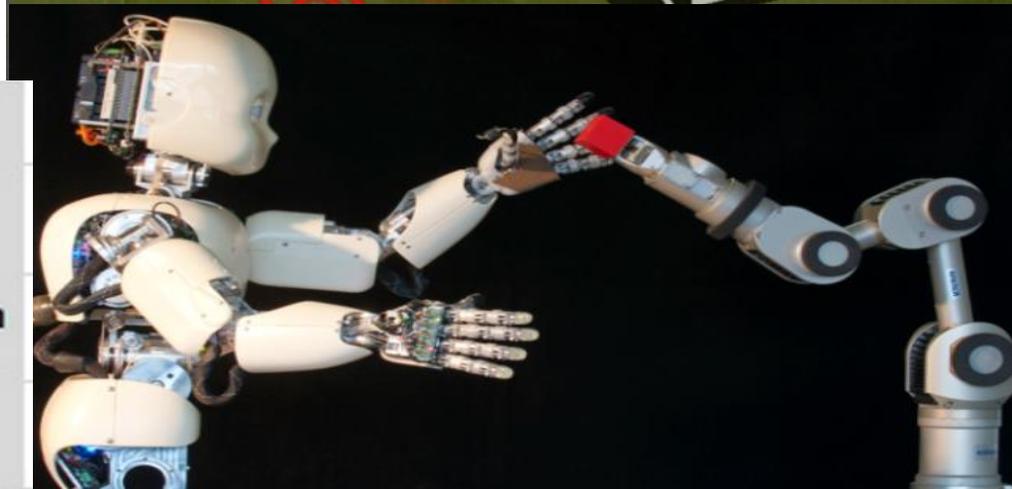
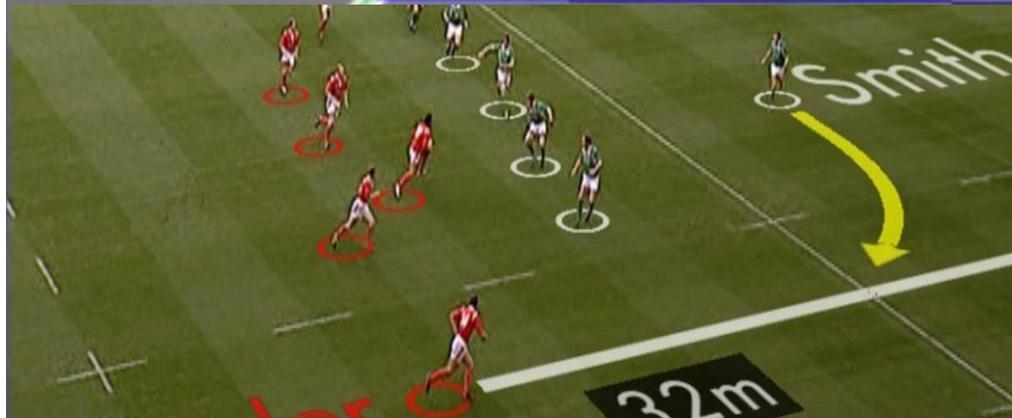
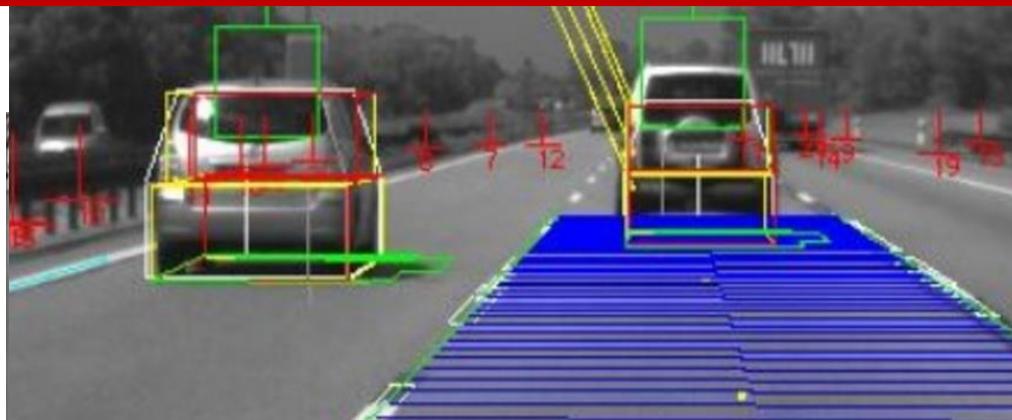
<https://www.edge.org/response-detail/26113>

AI and Computer Vision from '90s up to now

AI from '90s ...

- Logics
- Planning
- NLP
- Robotics
-

Pattern Recognition
Computer Vision
Machine Learning



- Automotive
- Sports & Entertainment
- Consumer
- Robotics and Machine Vision
- Medical
- Security & Surveillance

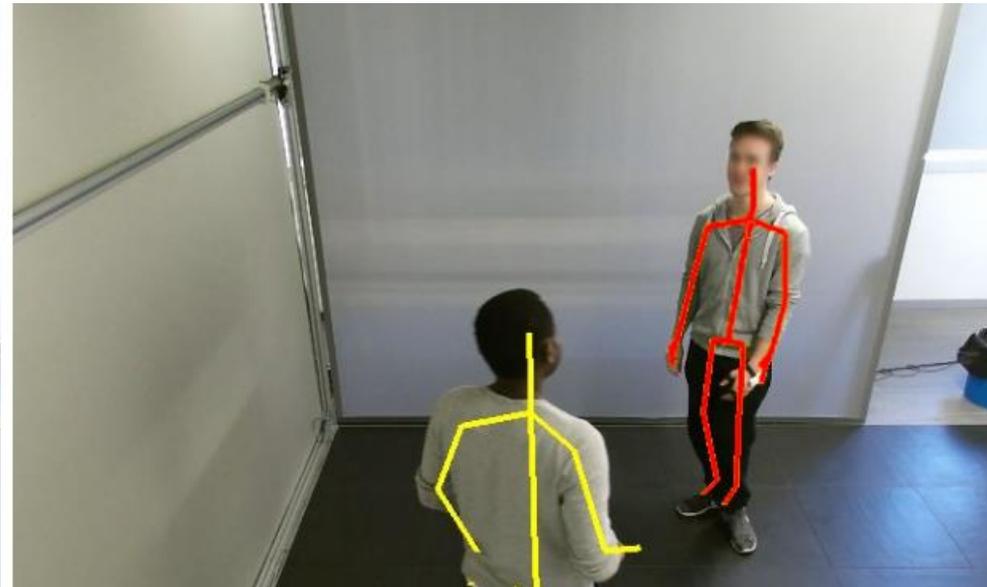
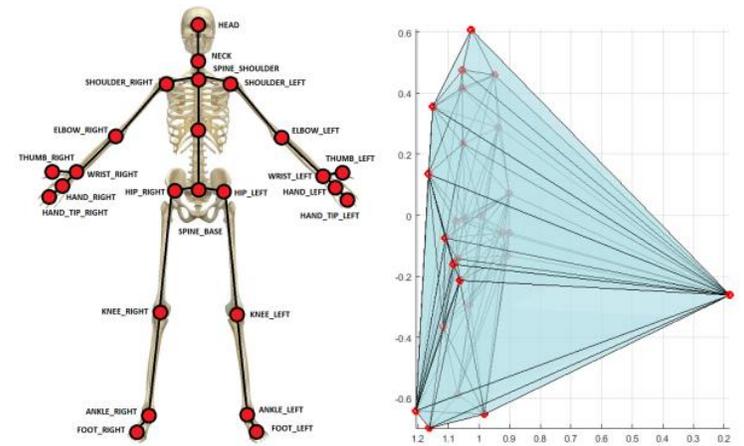
What can we do now?

Now: Computer vision can move robots (Students @Imagelab &Lapis)

Is it intelligence?

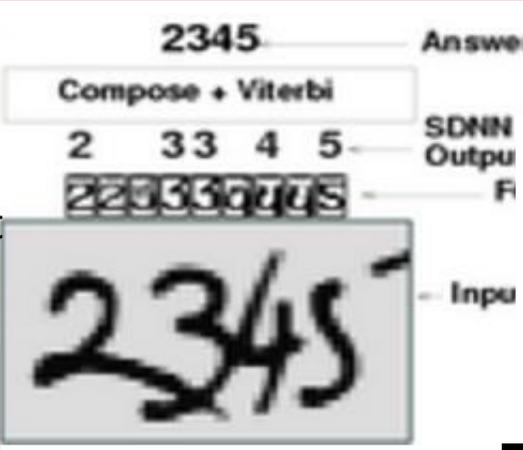


Camera spots your hidden prejudices from your body language

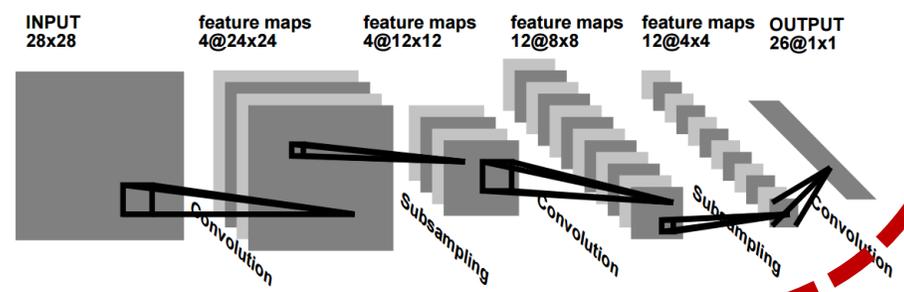


Palazzi,A.; Calderara,S.; Biccocchi,N; Vezzali,L.; di Bernardo,G.; Zambonelli,F; Cucchiara, R. "[Spotting prejudice with nonverbal behaviours](#)" *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Heidelberg, Settembre 2016,

From Computer Vision to Visual Intelligence

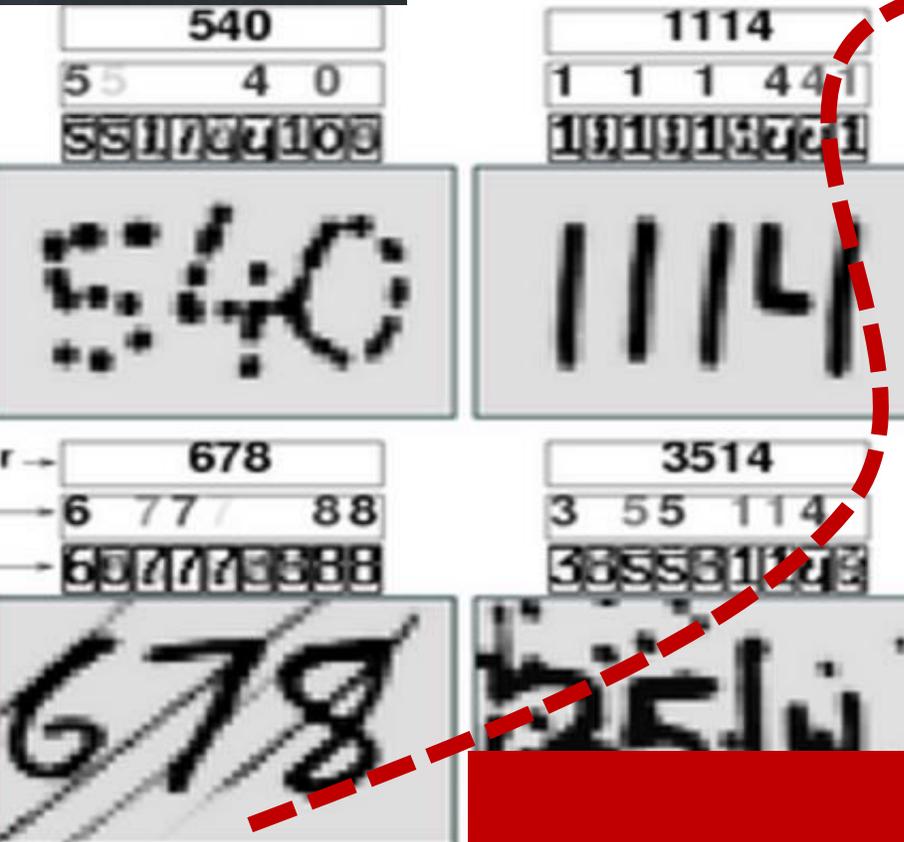


Yan LE Cun, Yoshua Benjo "Convolutional Networks for Images, Speech, and Time-Series, The Handbook of brain theory and neural networks 1995



IMAGENET

From 2010 1.2 billion images
1000 categories



Fei-Fei Li:

How we're teaching computers to understand pictures

TED2015 · 17:58 · Filmed Mar 2015

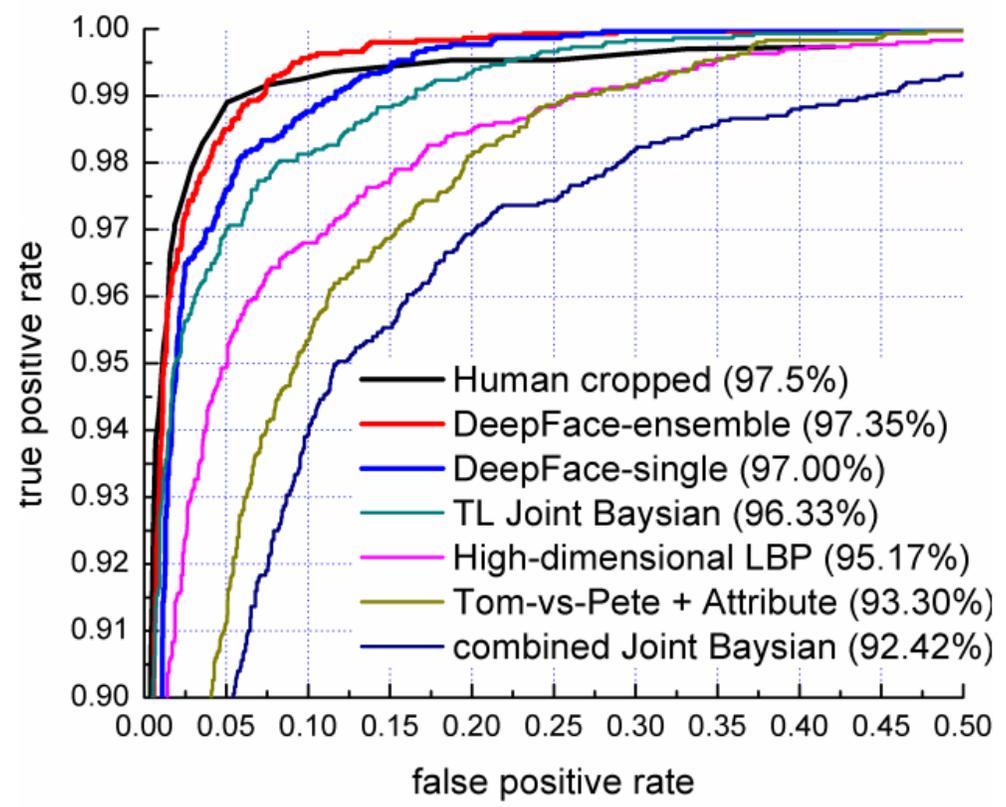
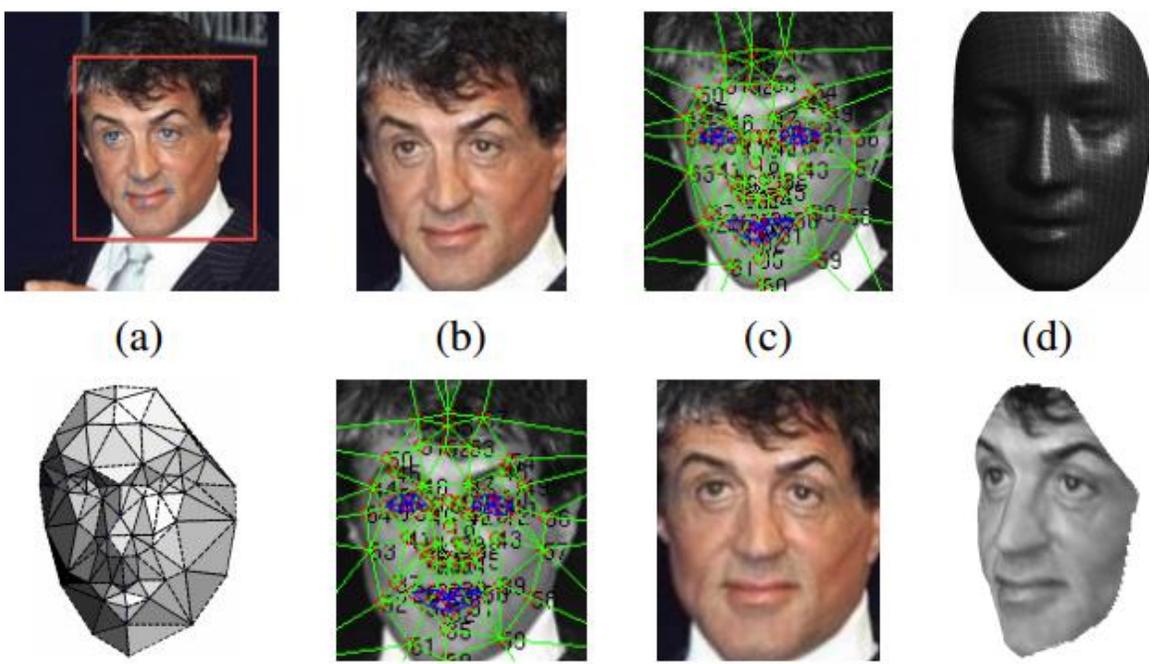
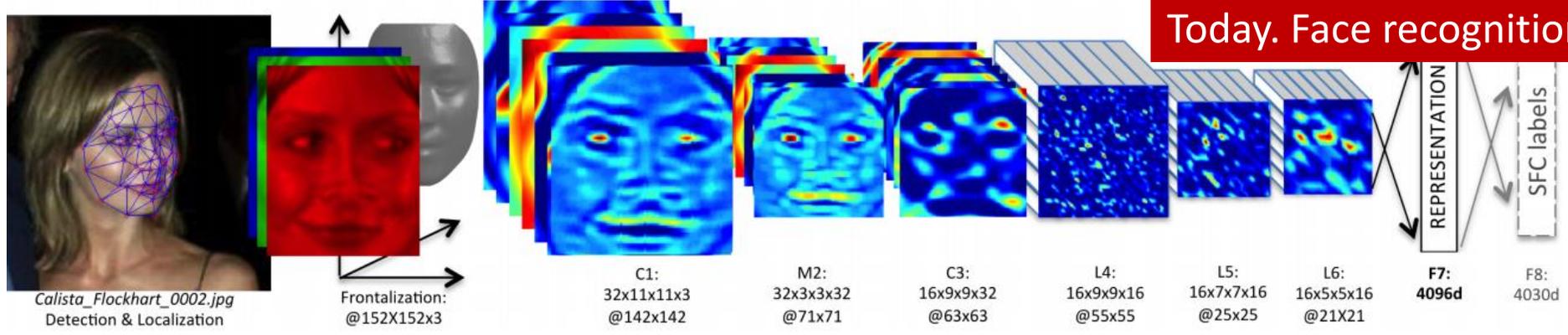
25 subtitle languages

View interactive transcript



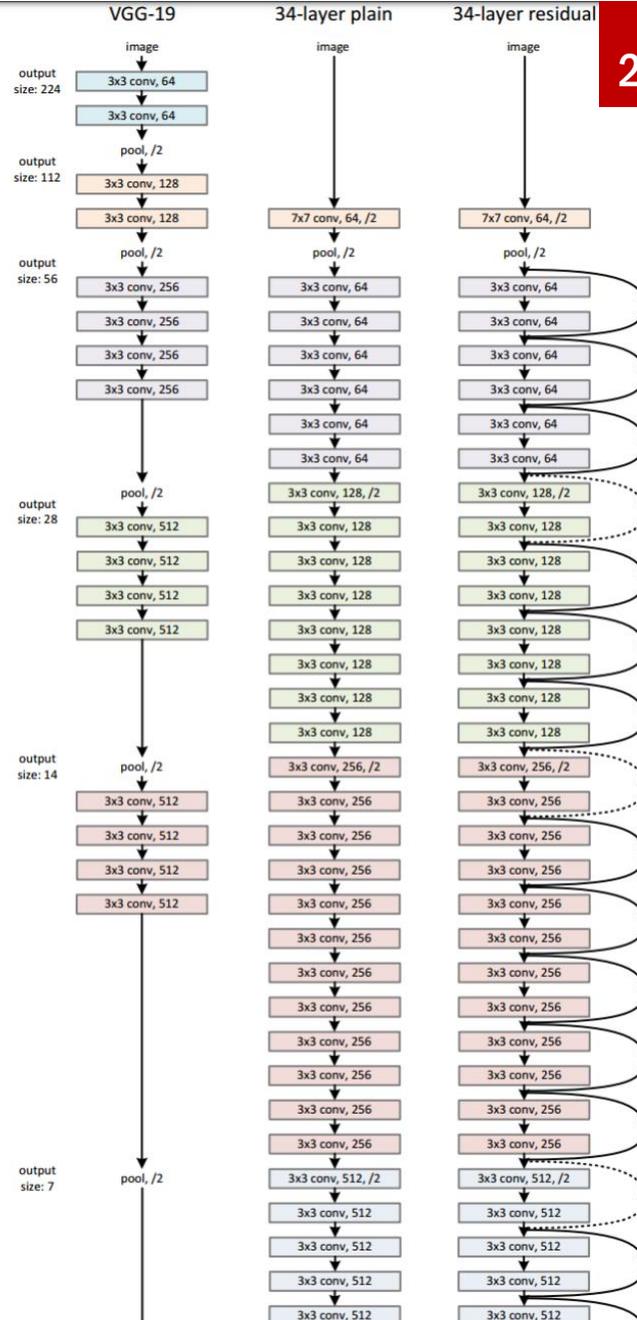
Deep Learning & Big (annotated) Data

Today. Face recognition with DL as humans



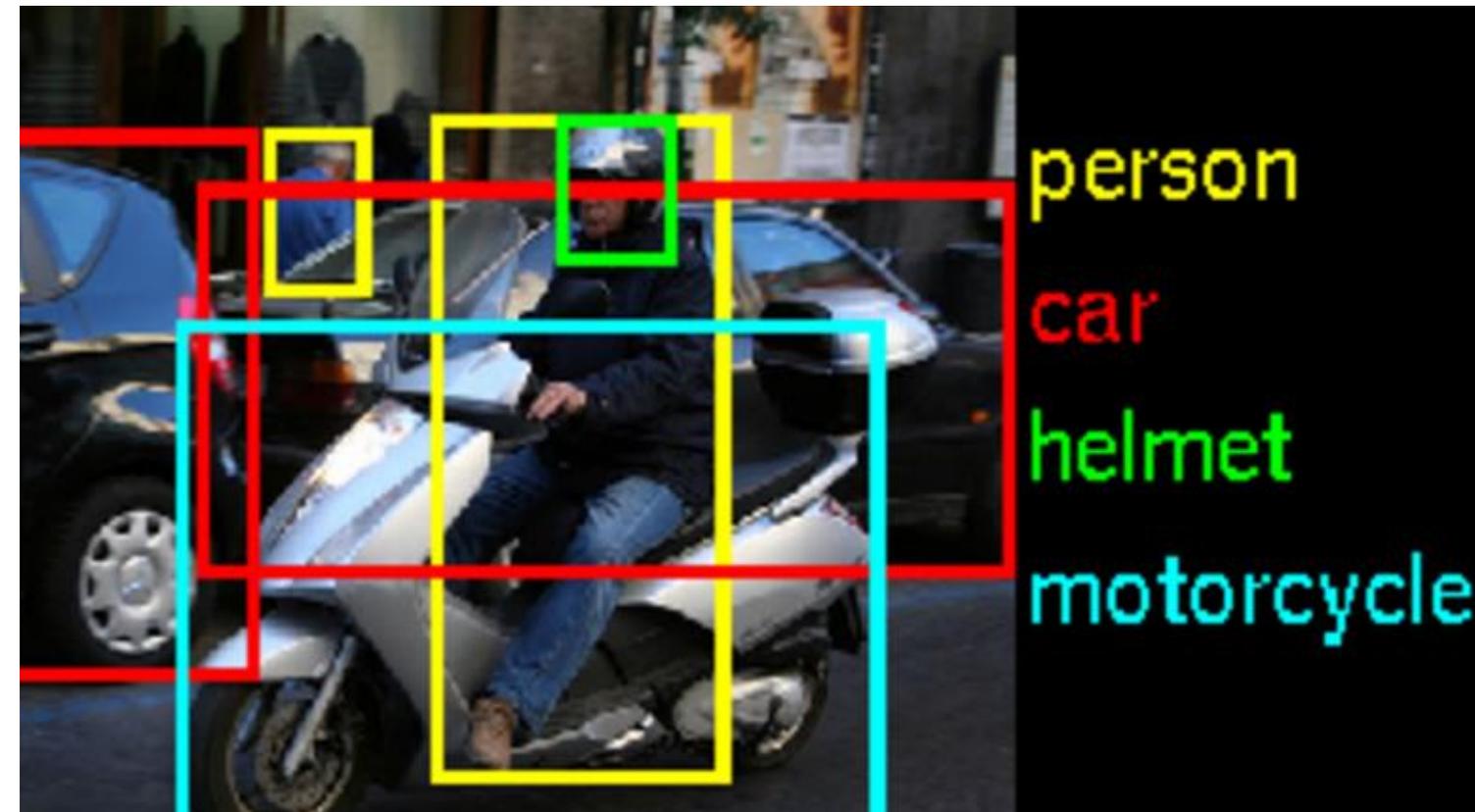
Trained on 4 million images, 4000 identities

2015 object recognition in the wild



over 1.2 billion images and 1000 concepts

2014 GoogleNet 7.8% top-5 err
2015 Microsoft ResNet-152 4.5% top-5 err
+3.6 billions FLOPS





Visual concepts

camel · snow · slope · land · side

Speaker transcript

Winter is the time for breeding. This extraordinary performance is a male Bactrian camel's way of attracting the attention of a passing female. In summer, the camels can't stray far from waterholes, but now, with mouthfuls of snow lying everywhere, they can travel widely in search of mates. Today less than a thousand of these desert specialists remain in the wild. The Gobi, hostile though it is, is their last stronghold.

Scene 6

Shots



Visual concepts

sand · place · Dunes · mouthful · snow

Speaker transcript

There's no other desert quite like the Gobi. But why is this place a desert? There is one simple and massive cause the Himalayas.

Scene 7

Shots



Deserts

This instalment features the harsh environment that covers one-third of the land on Earth: the deserts. Due to Siberian winds, Mongolia's Gobi Desert reaches extremes of temperature like no other, ranging from -40 °C to +50 °C (-40 °F to 122 °F). It is home to the rare Bactrian camel, which eats snow to maintain its fluid level and must limit itself to 10 liters (2.6 U.S. gal; 2.2 imp gal) a day if it is not to prove fatal. Africa's Sahara is the size of the USA, and just one of its severe dust storms could cover the whole of Great Britain. While some creatures, such as the dromedary, take them in their stride, for others the only escape from such



Drag your shots here (double click to remove)



Deserts

Probability: 0.77%

Winter is the time for breeding. This extraordinary performance is a male Bactrian camel's way of attracting the attention of a passing female. In summer, the camels can't stray far from waterholes, but now, with mouthfuls of snow lying everywhere, they can travel widely in search of mates. To...



Deserts

Probability: 0.69%

Not all deserts are hot. Fifty mile an hour winds, blowing in from Siberia, bring snow to the Gobi Desert in Mongolia. From a summer high of 50 degree centigrade, the temperature in mid-winter can drop to minus 40 making this one of the harshest deserts of all. Few animals can survive these ex...



Deserts

Probability: 0.52%

The ferocious wind, armed with grains of sand, is the agent that shapes all deserts...



Deserts

Probability: 0.51%

A third of the land on our planet is desert. These great scars on the face of the earth appear to be lifeless. But surprisingly none are. In all of them, life manages somehow to keep a precarious hold...



- GT1: the woman is riding a horse
- GT2: the horse and rider trotted down the field
- GT3: a person is riding a horse
- GT4: a woman is riding a horse
- GT5: a girl is riding a horse

Pr: a woman is riding a horse



- GT1: a woman is slicing potatoes
- GT2: a woman is cutting a potato into small pieces
- GT3: a person is slicing a potato into pieces
- GT4: a woman is slicing potatoes
- GT5: a woman is cutting a potato

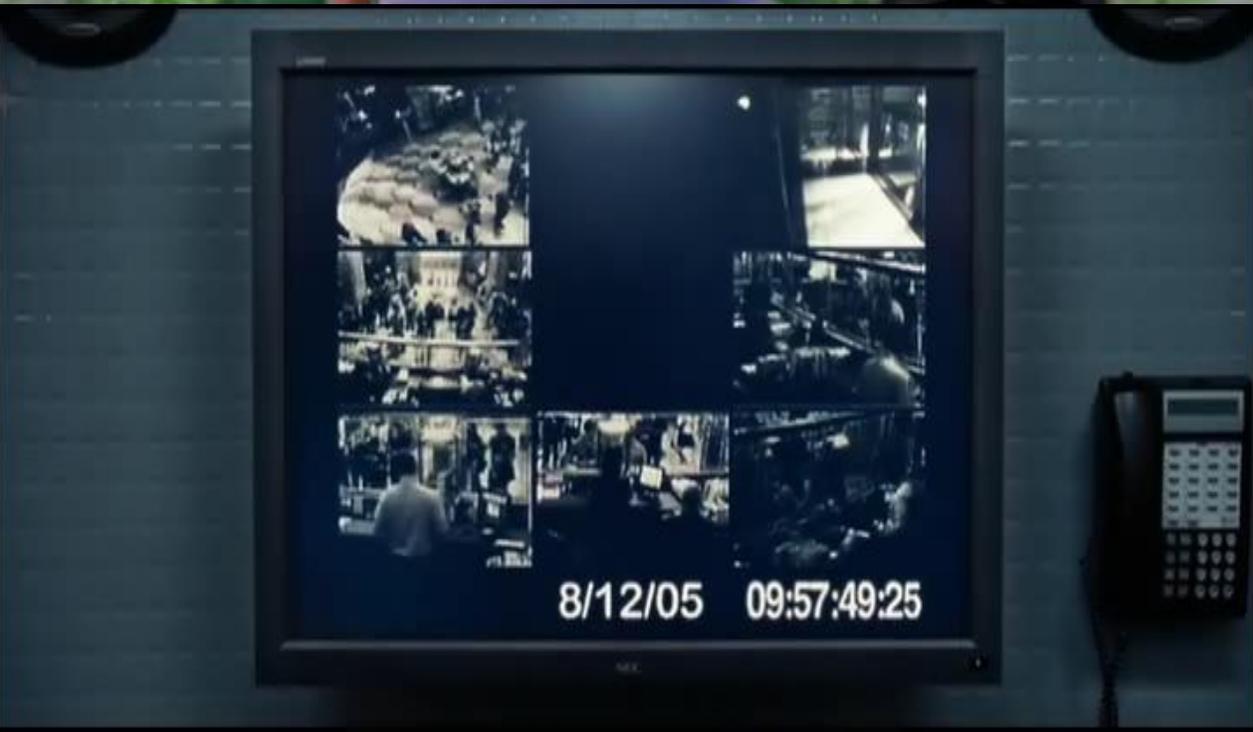
Pr: a person is cutting a potato



Video captioning..
Understanding facts and emotions...
A still un-completely solved problem

GT: he stands and offers her the small bouquet

Pr: **someone looks up at someone**



GT: an image appears on a screen

Pr: **someone and someone watch the screen**

Visual Intelligence for humans and machines in (autonomous) car

- Understanding the human pose/gaze
- Understanding the human attention/distraction
- Learn the human driving behavior

...use both intelligences together.



Look outside

Form INSIDE TO OUTSIDE: Dr(Eye)Ve project @Imagelab



Dr(Eye)Ve project @Imagelab 2016

*Dr(eye)ve learned where the drivers see,
and what the drivers pay attention on...*

it is learning an intelligent visual behavior!

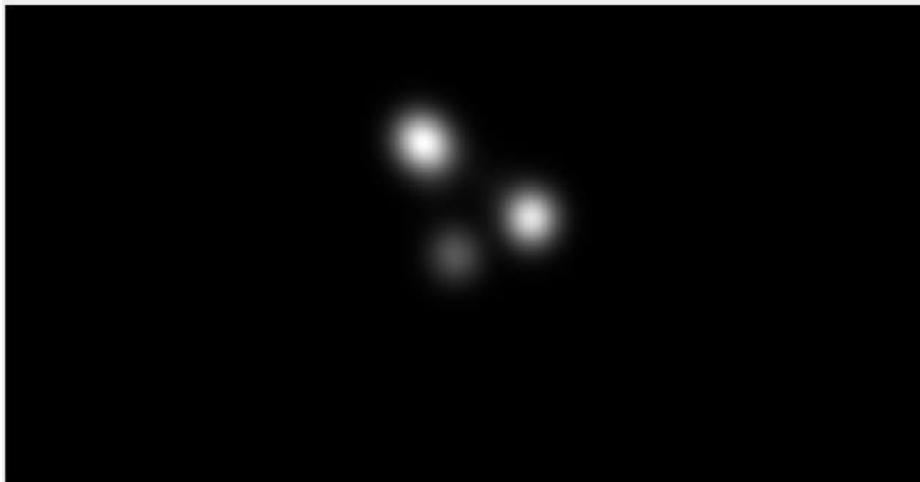
roof mounted camera



semantic segmentation



attention model prediction



overlay

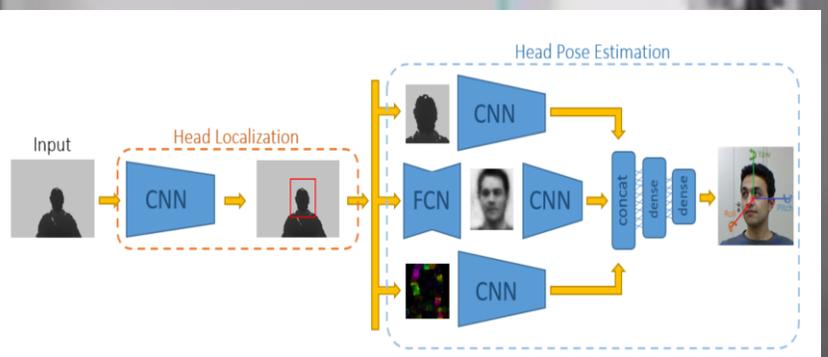


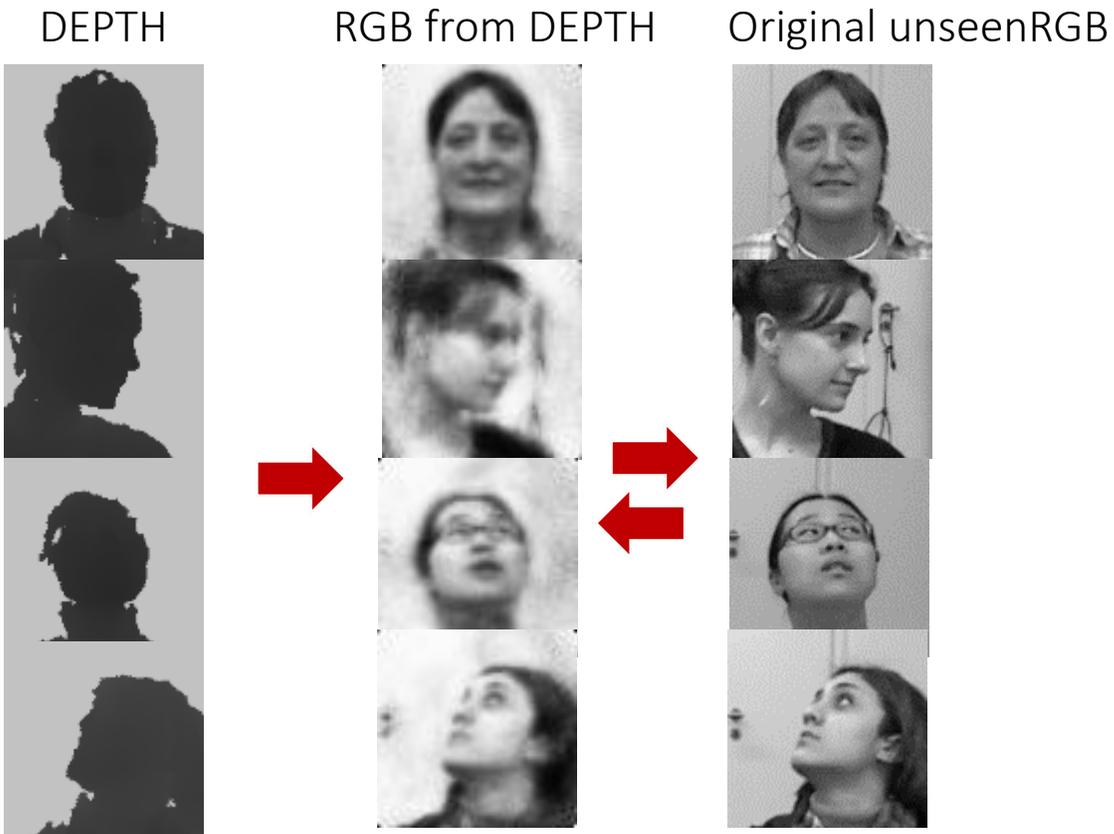
Visual Intelligence for humans and machines in (autonomous) car



Look outside

Understanding the human pose by depth only (with DL)

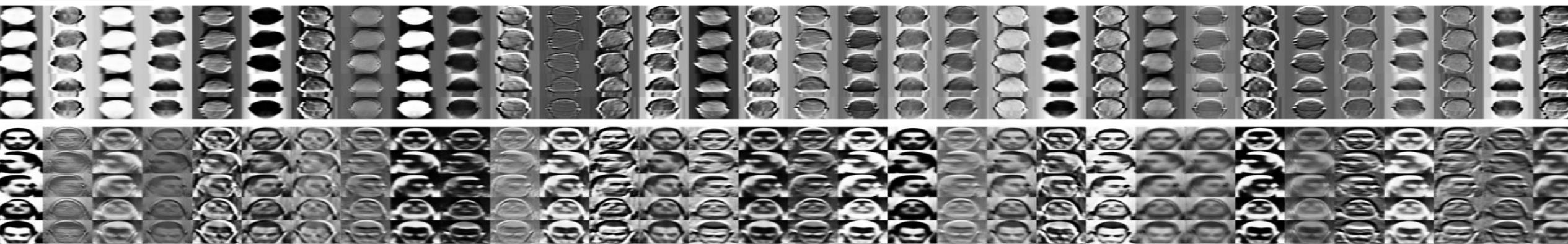




A (spectacular) side effect.



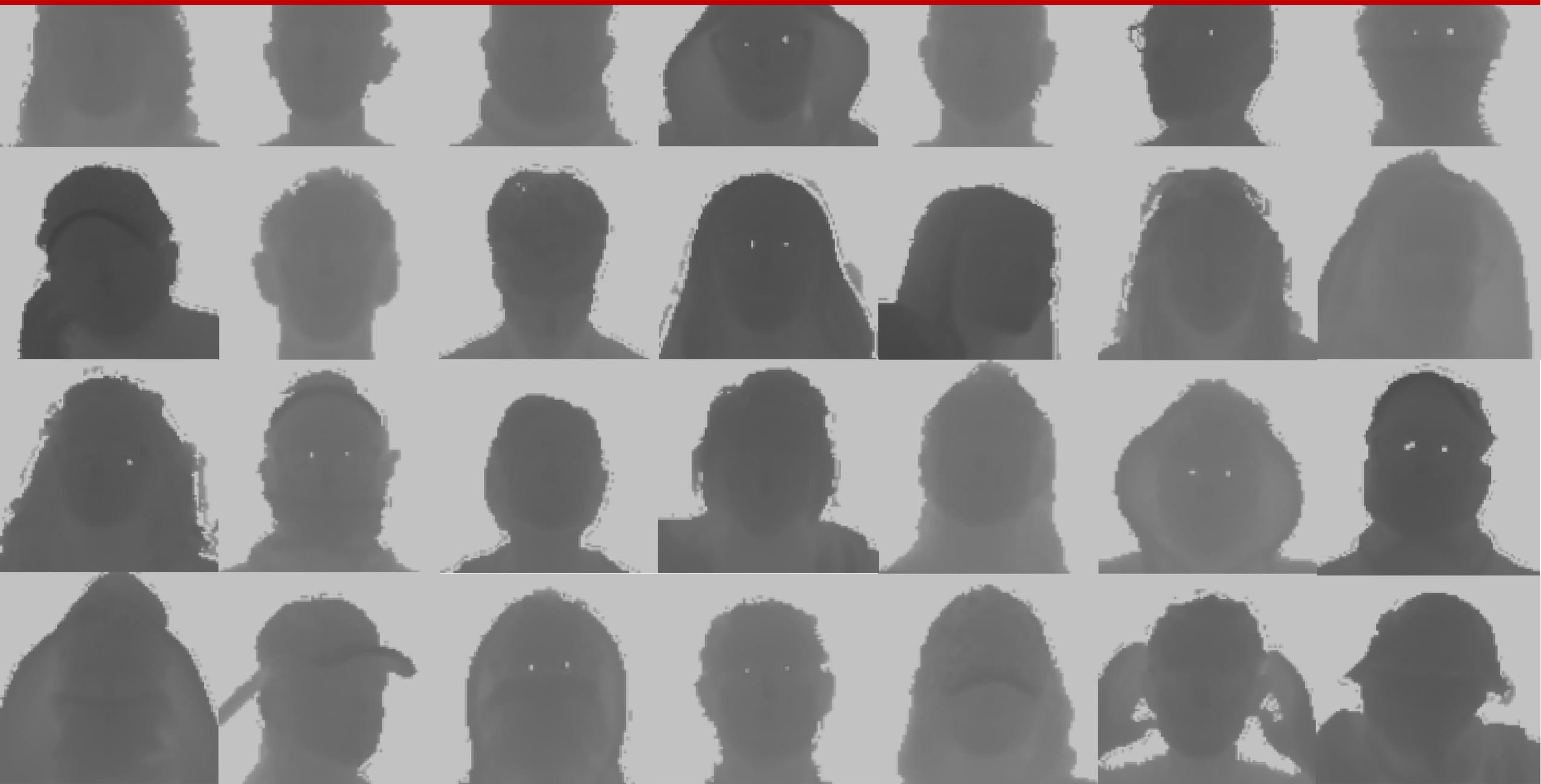
POSEidon learned something more...



A TRANSFER LEARNED EXPERIENCE.

*To have a mental image on what it did't see
To imagine face by depth!*

3D Pandora dataset @Imagelab



2D Pandora dataset @Imagelab



Learned by the Poseydon Net @Imagelab

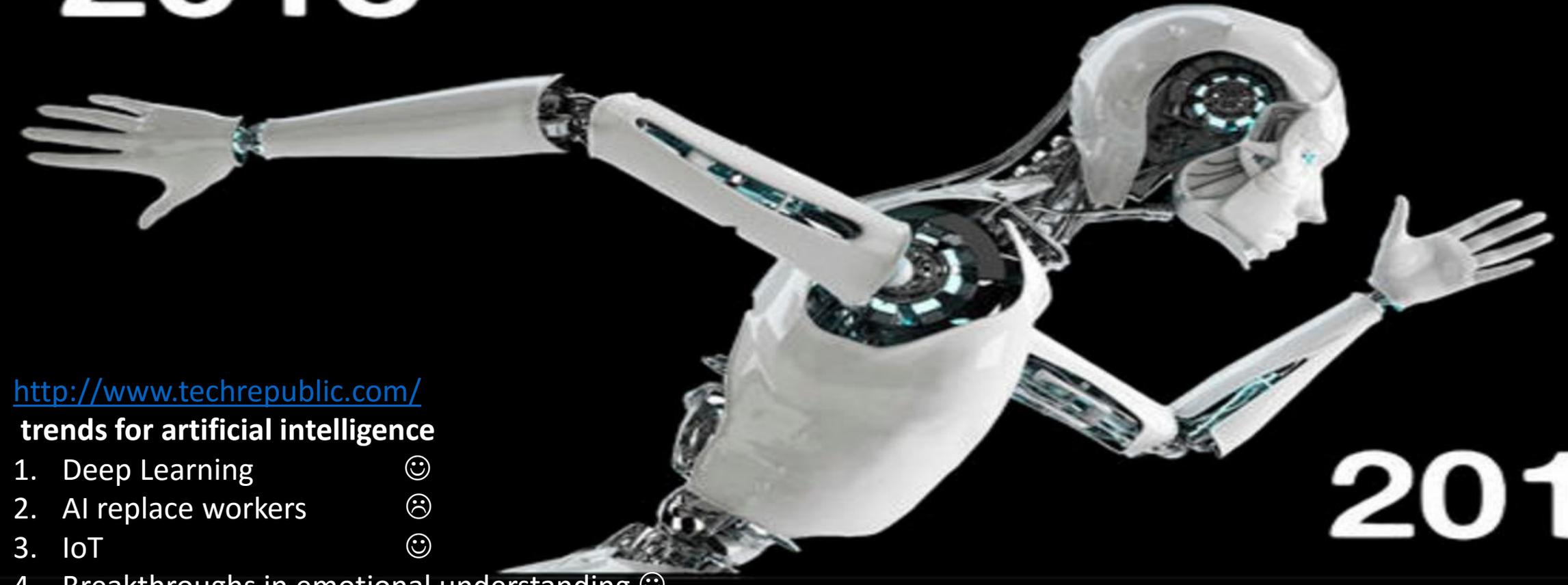


2015

2016

[http://www.techrepublic.com/trends for artificial intelligence](http://www.techrepublic.com/trends-for-artificial-intelligence)

1. Deep Learning ☺
2. AI replace workers ☹
3. IoT ☺
4. Breakthroughs in emotional understanding ☺
5. AI in shopping and customer service ☹
6. Ethical questions ☺
7. A problem with (gender) representation ☹



We have artificial Intelligence Now.

And now.. Ride the bike!

Thanks .

Thanks to Imagelab
www.imagelab.unimore.it
Rita Cucchiara
rita.cucchiara@unimore.it

