



# Critical Review of AI Regarding Localization

Andrzej Zydrón CTO XTM Intl

LT-Innovate Summit, Brussels 2019

**Better** Translation **Technology**





ELSEVIER

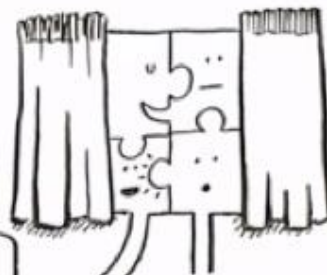


# LANGUAGE

**STEVEN  
PINKER**  
THE STUFF OF THOUGHT:  
LANGUAGE AS A WINDOW  
INTO **HUMANNATURE**



LANGUAGE  
IS A WINDOW  
INTO SOCIAL  
RELATIONS

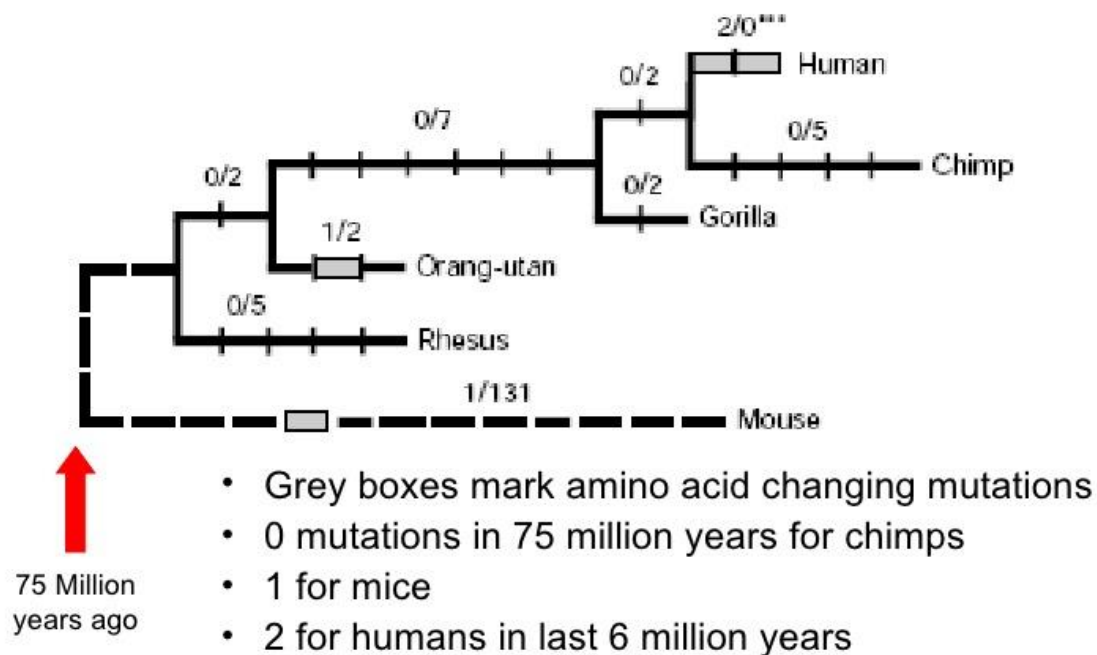


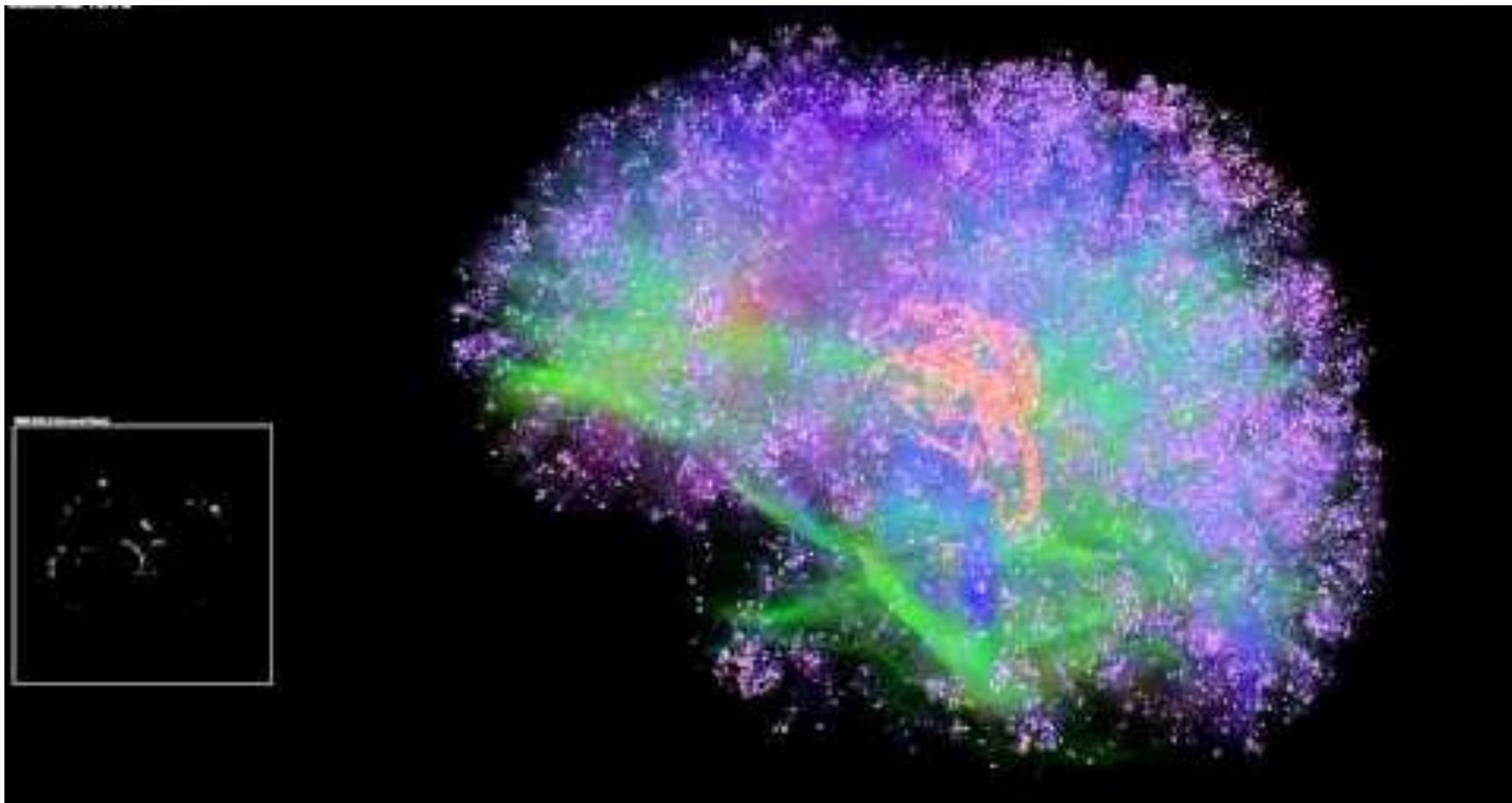
I'M BEGIN  
WITH A





# Evolution of FOXP2





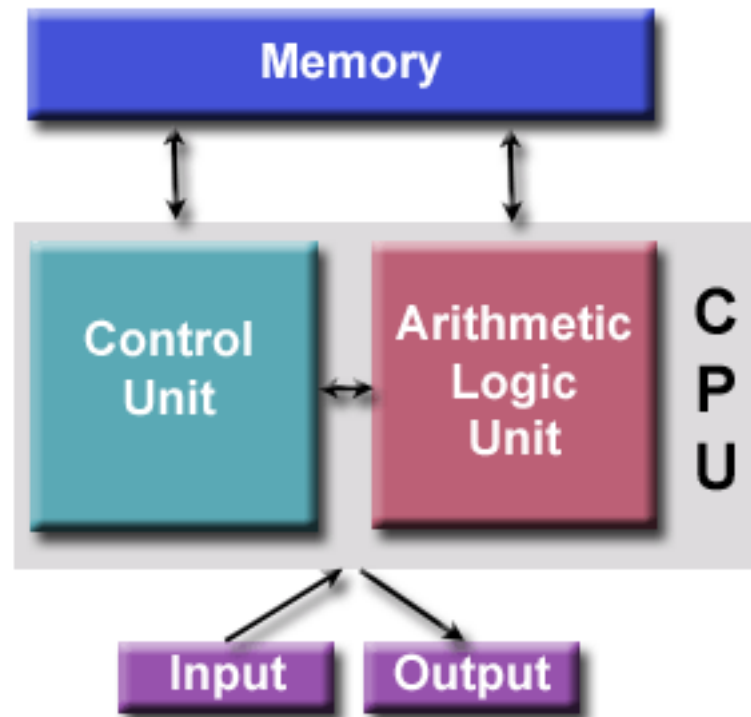
30 billion cells, 100 trillion synapses



- Trajectory
- Velocity
- Angle
- Wind speed
- Direction



# von Neumann architecture





=





# Efficiency

## Jeopardy!



= 80 Watts



= 90,000 Watts

# von Neumann architecture limitations



Very Good



Not so good



# Person!

Try the API

Objects

Labels

Logos

Web

Properties

Safe Search



xxx.jpg

Person

87%

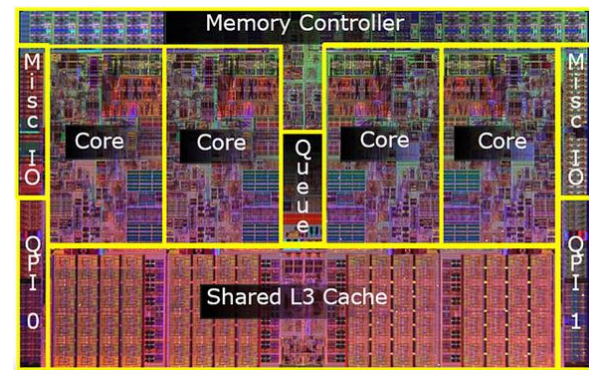


# Human vs Computer

- Human 200 OPS
- Computer 82,300,000,000 OPS



VS

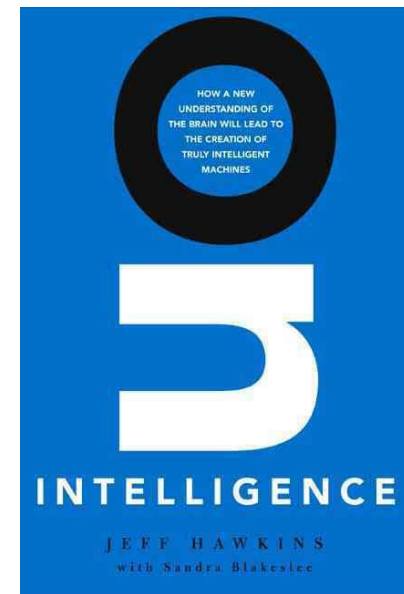


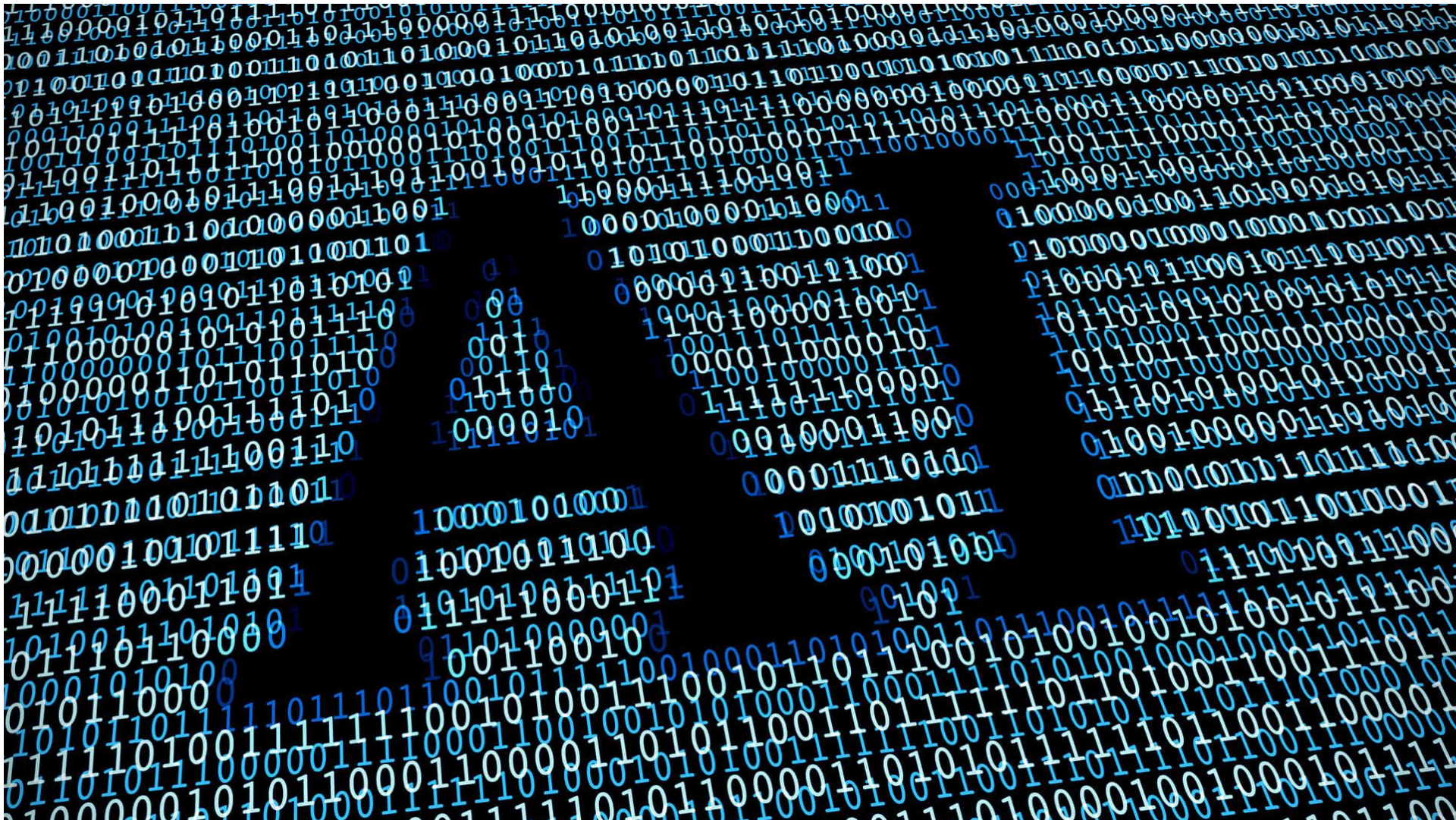


# What is intelligence?

1. Recognize patterns
2. Remember those patterns
3. Use the recognized patterns to predict future events

Jeff Hawkins, and Sandra Blakeslee, On Intelligence 2005





# What is Artificial Intelligence?

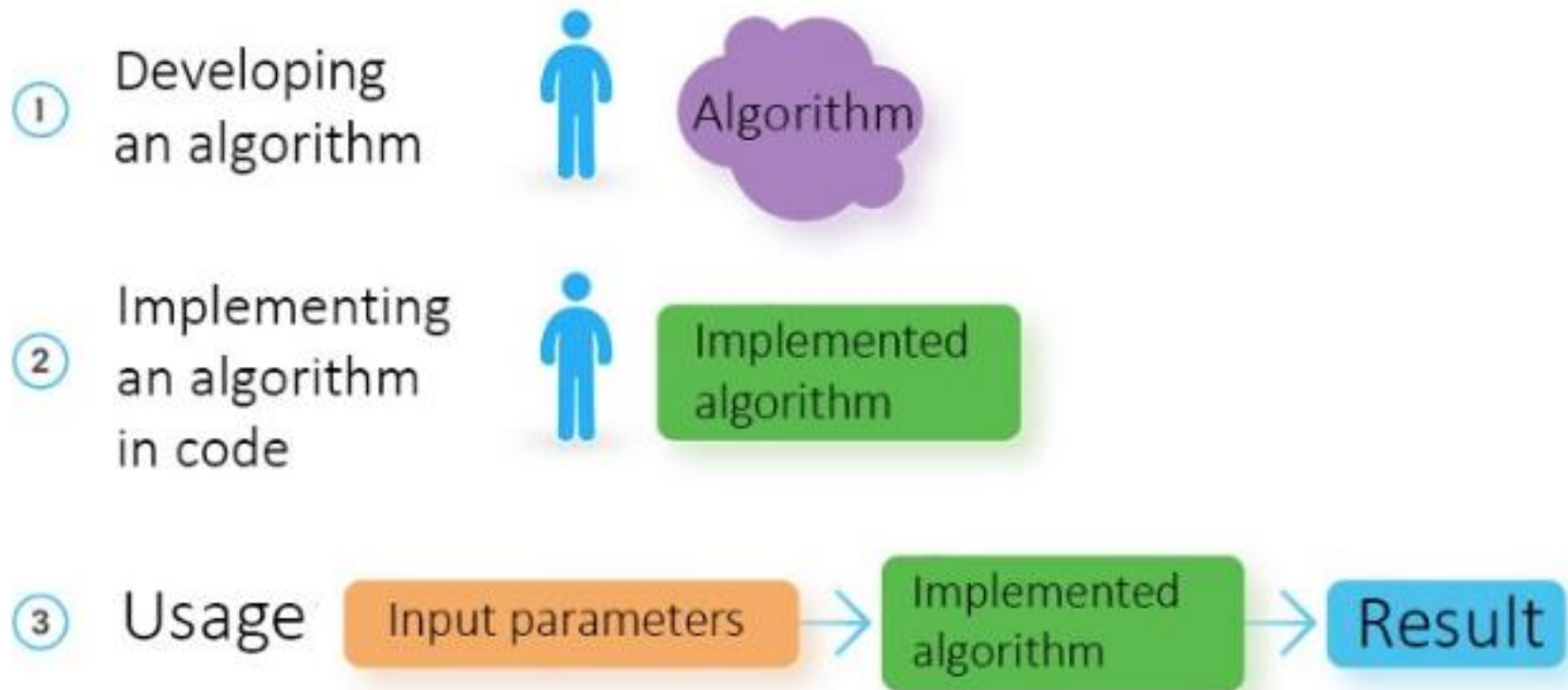
## 1. Expert Systems

- a. Rule based
- b. Knowledge based

## 2. Machine Learning

- a. Bayesian probability
- b. Statistical probability
- c. Neural Networks

# Expert Systems: rule based



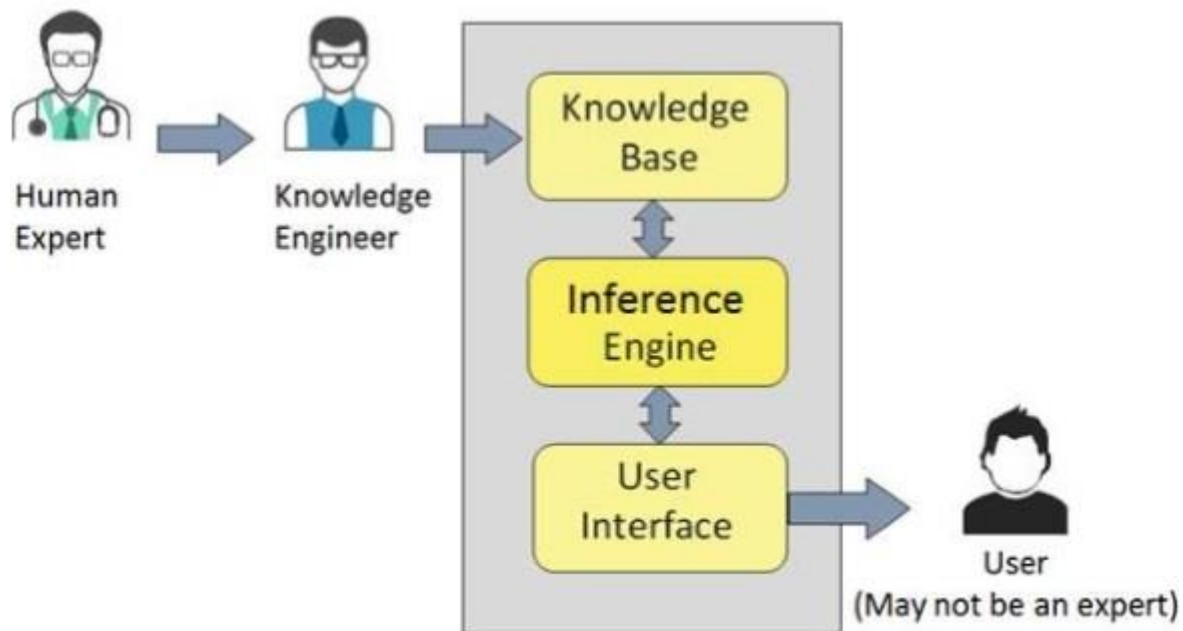
# Expert Systems: rule based

Table 1-1

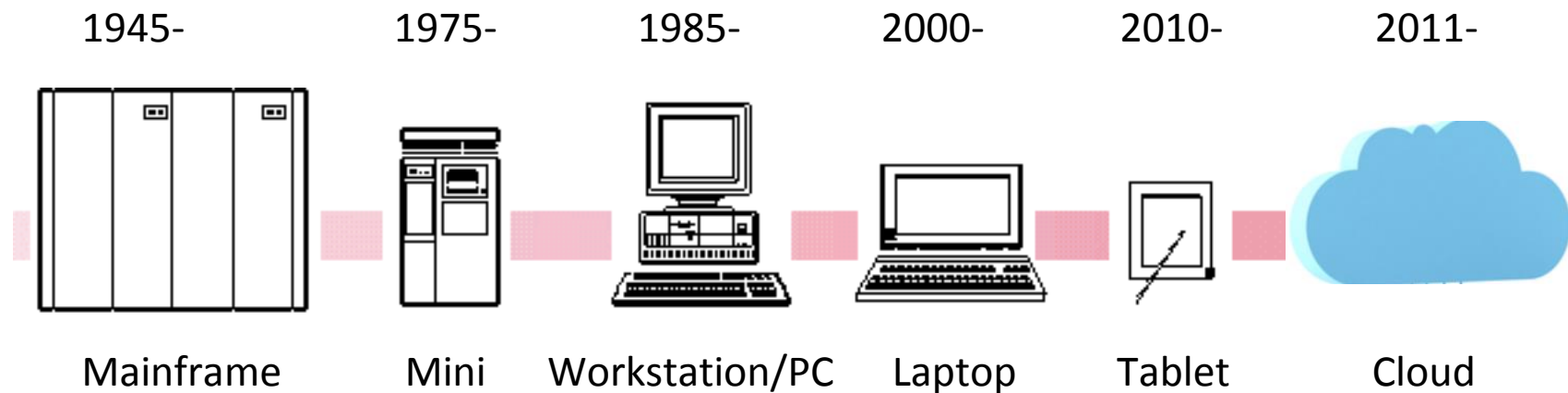
Run No.	Composition (wt %)													Tensile Extension At a Temperature	Existance of Cracks	Corrosion Rate (mm/year)		
	Cr	Mo	W	Cu	Ca	Zr	Nb	Ta	Hf	Y	Rare Earth Elements	Impurities				Ni and the other impurities	Mixed Acid	Inverse Aqua Regia
												C	N					
1	38.6	0.96	-	-	-	-	-	-	-	-	-	0.045	0.004	rem.	65.5	⊙	0.106	0.050
2	43.0	-	1.83	-	-	-	-	-	-	-	-	0.028	0.003	rem.	60.3	⊙	0.088	0.049
3	44.2	0.97	0.27	-	-	-	-	-	-	-	-	0.014	0.021	rem.	62.6	⊙	0.074	0.046
4	40.5	0.68	-	-	-	-	-	-	-	-	-	0.029	0.004	rem.	61.6	⊙	0.098	0.040
5	45.5	0.12	-	-	-	-	-	-	-	-	-	0.022	0.006	rem.	58.5	⊙	0.071	0.052
6	42.1	0.96	-	-	-	-	-	-	-	-	-	0.018	0.003	rem.	50.3	⊙	0.087	0.053
7	44.6	1.52	-	-	-	-	-	-	-	-	-	0.031	0.003	rem.	55.6	⊙	0.078	0.035
8	45.6	-	0.11	-	-	-	-	-	-	-	-	0.032	0.004	rem.	59.2	⊙	0.075	0.054
9	44.6	-	0.96	-	-	-	-	-	-	-	-	0.037	0.004	rem.	66.1	⊙	0.078	0.042
10	46.5	1.21	-	-	-	-	-	-	-	-	-	0.029	0.006	rem.	60.4	⊙	0.077	0.049
11	44.2	0.97	0.27	1.0	-	-	-	-	-	-	-	0.014	0.021	rem.	62.0	⊙	0.074	0.030
12	44.6	1.52	-	1.5	-	-	-	-	-	-	-	0.031	0.003	rem.	52.3	⊙	0.058	0.020
13	44.6	-	0.96	0.5	-	-	-	-	-	-	-	0.037	0.004	rem.	53.4	⊙	0.068	0.037
14	44.2	0.97	0.27	-	0.002	-	-	-	-	-	-	0.014	0.004	rem.	63.6	⊙	0.072	0.048
15	44.6	1.52	-	-	0.005	-	-	-	-	-	-	0.031	0.003	rem.	52.7	⊙	0.075	0.038
16	44.6	-	0.96	-	0.008	-	-	-	-	-	-	0.014	0.004	rem.	69.3	⊙	0.081	0.041
17	44.2	0.97	0.27	-	-	3.0	-	-	-	-	-	0.014	0.004	rem.	64.7	⊙	0.083	0.051
18	44.6	1.52	-	-	-	-	0.5	-	-	-	-	0.031	0.003	rem.	57.8	⊙	0.063	0.021
19	44.6	-	0.96	-	-	-	-	0.8	-	-	-	0.014	0.004	rem.	69.3	⊙	0.060	0.027
20	45.6	-	0.11	-	-	-	-	-	1.2	-	-	0.032	0.004	rem.	63.2	⊙	0.075	0.041

EP 0 255 854 A2

# Expert Systems: knowledge based



# Information Technology Evolution



# Unimaginable Scales

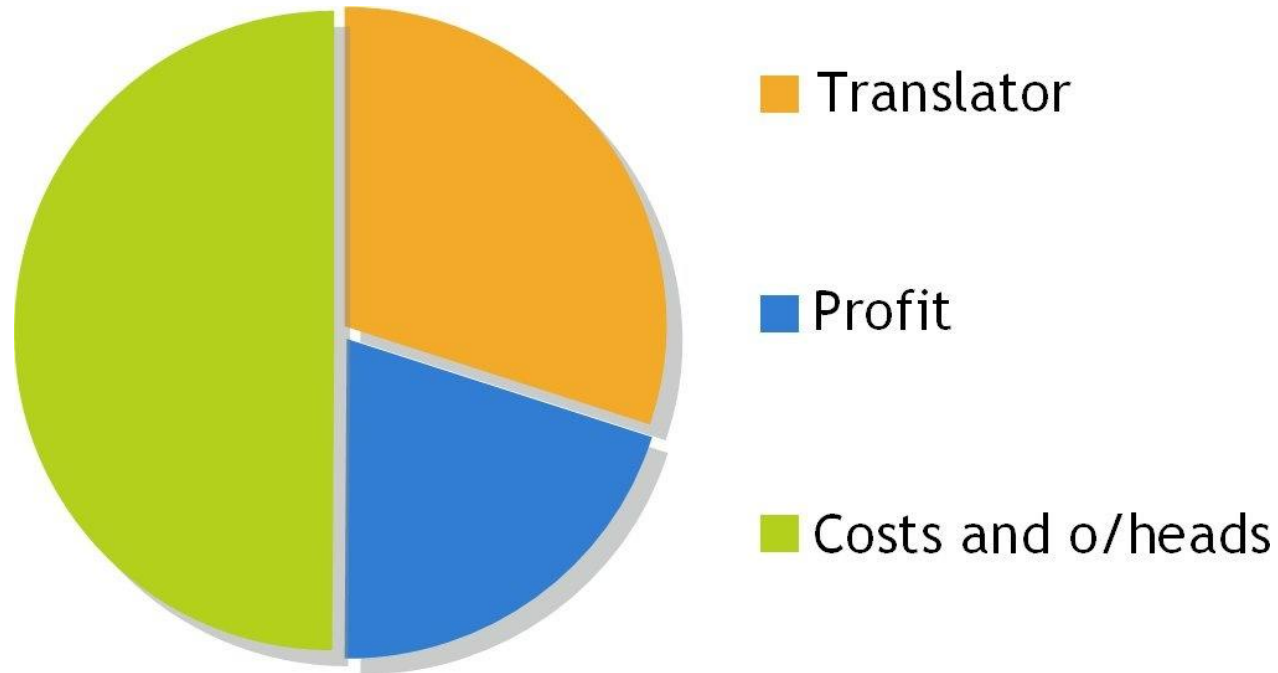




# What are we trying to achieve?

- Reducing costs?
- Increasing productivity?
- Increasing quality?
- It depends on the goal
- Provides an additional solution for a given instance

# Putting Things Into Perspective



Source Professor Reinhard Schäler LRC - ASLIB 2002

# AI in Localization

## Expert Systems

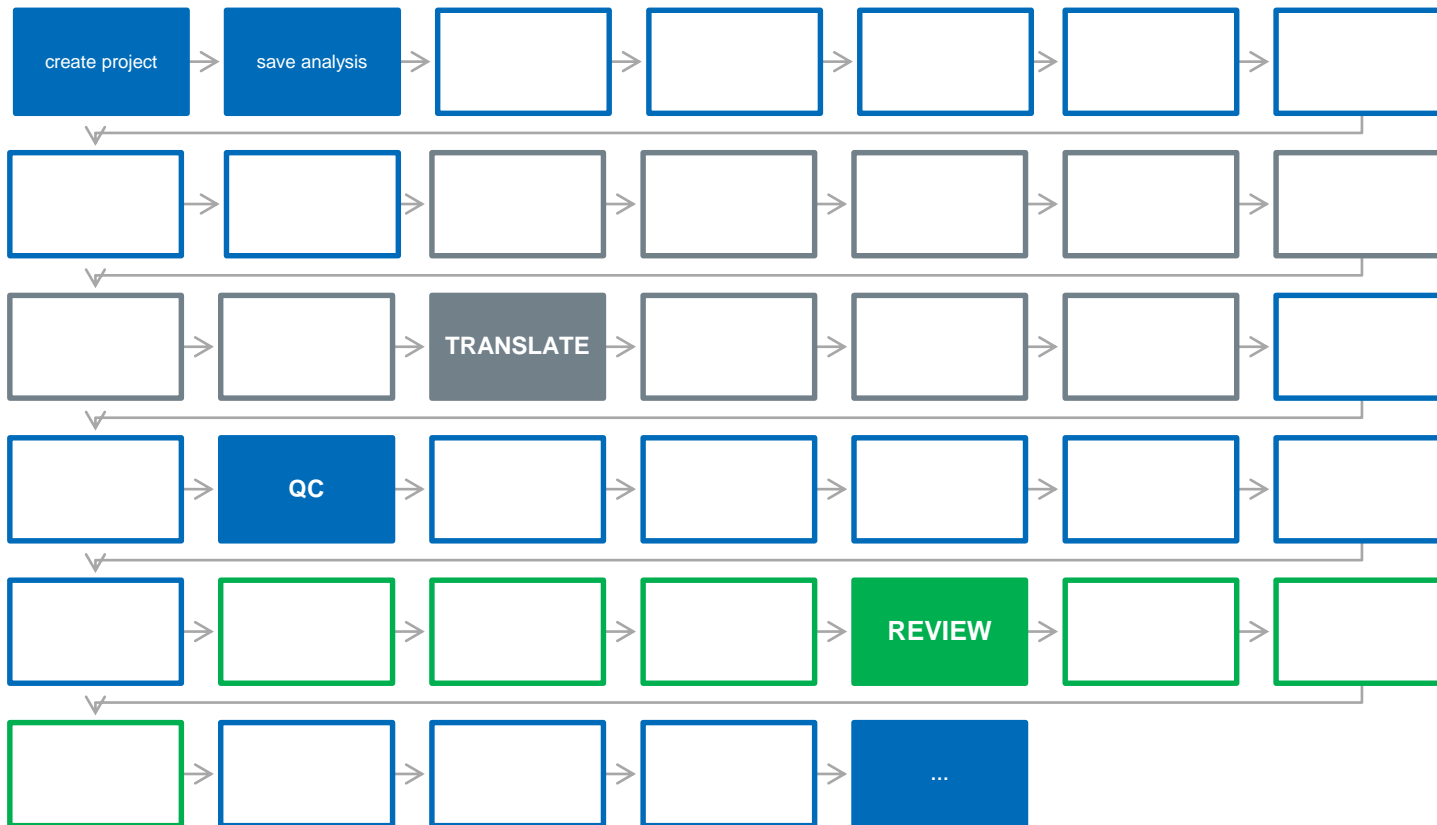
1. Automated Workflow
2. Translation Memory
  - a. ICE matching
  - b. Leveraged
  - c. Fuzzy matching
3. Terminology Extraction, including Bilingual
4. Intelligent Corpus Alignment
5. Intelligent Decision Making: which translations to proofread

# Complex Workflows with Offline Data

Offline Work = Cumbersome File Management



# Automated Workflow



# Translation Memory



Welcome BobTR, Reviewer/Corrector/Translator, Mela Logout

My Inbox Configuration

Project : Dell Laptops2 File : Por que utilizar portátiles de Dell - ES 2.docx Role : Translator Segments : | 1 - 27 |

< 1-10 11-20 21-27 > 1-27 ▾

Translations			
ID	Spanish	English (UK)	Status
1	Por qué utilizar portátiles de Dell	Why use Dell laptops	<span>M</span>
2	Dell ofrece una amplia gama de portátiles, desde sistemas diseñados exclusivamente para la pequeña empresa hasta estaciones de trabajo móviles totalmente certificadas o factores de forma especializados como Tablet PC o portátiles reforzados.	Dell offers a wide range of laptops, from systems for small companies to portable workstations or specialised form factors like Tablet PCs or ruggedised laptops.	<span>M</span>
3	De esta forma podrá encontrar una oferta adecuada a sus necesidades y presupuesto, lo que le permitirá adquirir todos sus sistemas en un único proveedor y reducir su complejidad.	In this way you can find an offer to suit your needs and budget, which will enable you to purchase all your systems from one supplier and reduce complexity.	<span>M</span>
4	Dell tiene presencia mundial y presta un servicio de calidad en 130 países, las 24 horas del día, los 7 días de la semana.	Dell has a worldwide presence and provides a service in 130 countries 24 hours a day 7 days a week.	<span>M</span>
5	Éstas son las razones para elegir un nuevo portátil de Dell:	Here is why you should choose a new laptop from Dell:	<span>M</span>
6	<p>{1}La arquitectura inalámbrica de Dell puede ofrecer conexiones Wi-Fi más rápidas que los productos de la competencia {1}{2}1{2}.</p> <p><b>Source/Fuzzy source:</b></p> <p><b>Source:</b> {1}La arquitectura inalámbrica de Dell puede ofrecer conexiones Wi-Fi más rápidas que los productos de la competencia {1}{2}1{2}.</p> <p><b>Fuzzy source:</b> La arquitectura inalámbrica de Dell puede ofrecer conexiones Wi-Fi más rápidas que los productos de la competencia 1 .</p>	<p>{1}Dell's wireless architecture can provide Wi-Fi connections faster than competitive products {1}{2}1{2}.</p> <p><b>Leveraged:</b></p> <p>{1}Dell's wireless architecture can provide Wi-Fi connections faster than competitive products {1}{2}1{2}.</p> <p>{1}La arquitectura inalámbrica de Dellxx puede ofrecer conexiones Wi-Fi más rápidas que los productos de la competencia {1}{2}1{2}.</p> <p><b>Fuzzy target:</b></p> <p>Dell's laptops help remove more dead spots in the home or office than any other leading small business laptop provider 1 . 97%</p>	<span>M</span>
7	Los ordenadores portátiles de Dell son los mas avanzados en el mundo.	Dell Laptops are the most advanced in the world.	<span>M</span>
8	{1}Con los ordenadores portátiles de Dell se eliminan más puntos sin cobertura en casa o en la oficina que con los productos de la competencia {1}{2}1{2}.	{1}Con los ordenadores portátiles de Dellxx se eliminan más puntos sin cobertura en casa o en la oficina que con los productos de la competencia {1}{2}1{2}.	<span>M</span>
9	Un 93% de los usuarios consiguieron conectarse con éxito a una red utilizando el Asistente de red de Dell.	93% success rate of users using Dell Network Assistant to connect to a network	<span>M</span>
10	El porcentaje de éxito fue del 66% entre los usuarios que no lo utilizan {1}2{1}.	The success rate was 66% among users who do not use it {1}2{1}.	<span>M</span>

# Bilingual Terminology Extraction

	A	B	C	D	E	F	G	H
1	en-US	Surface Forms Frequency Frequencies frequencies frequency	296	FR-small.xif	Context Open the Inventory tab to add the equipment you plan to use as offline devices and configure frequency coordination for that equipment. Though you must tune non-networked equipment to its calculated frequency values manually, Wireless Workbench can take into account the frequencies that the equipment can tune to during frequency coordination. Frequency Coordination	FR fréquences	Surface Forms fréquences	Context Ouvrir l'onglet Inventaire pour ajouter les équipements prévus pour être utilisés comme appareils hors ligne et configurer la coordination des fréquences pour ces équipements. Bien que l'utilisateur doit régler manuellement les équipements hors réseau sur leurs valeurs de fréquence calculées, Wireless Workbench peut prendre en compte les fréquences sur lesquelles les équipements peuvent se régler lors de la coordination des fréquences. Coordination des fréquences
2	Channel	channels Channel channel Channels	172	FR-small.xif	You can see how many channels you will need, and if you have access to scan data, you can try multiple hardware combinations with the available frequencies in the RF environment of the venue. TV channels and other exclusions Use this setting to synthesize frequencies for high-priority audio channels, or when a wide variety of transmitter-to-receiver distances may occur (that is, some transmitters operate in close proximity to the receiver, while others are at a much greater distance).	canaux	canaux	L'utilisateur peut voir le nombre de canaux qui seront nécessaires et s'il a accès aux données de scan, il peut essayer plusieurs combinaisons de matériel avec les fréquences disponibles dans l'environnement RF du lieu de spectacle. Utiliser ce réglage pour synthétiser des fréquences pour des canaux audio à haute priorité ou lorsqu'il peut exister une grande disparité de distances émetteur-récepteur (c'est-à-dire que certains émetteurs fonctionnent à proximité immédiate du récepteur, tandis que d'autres se trouvent beaucoup plus loin).
3	Group	Group Groups groups group	154	FR-small.xif	A user group allows you to limit the spectrum available for Frequency Coordination. When a User Group is selected, the frequency calculator will attempt to place all frequencies within the specified group's included frequencies. Frequencies that lie outside the bounds of the User Group will fail analysis.	groupe	groupe	Un groupe d'utilisateurs permet de limiter le spectre disponible pour la coordination des fréquences. Lorsqu'un groupe d'utilisateurs est sélectionné, le calculateur de fréquences tentera de placer toutes les fréquences dans les fréquences incluses du groupe spécifié. Les fréquences qui se trouveront en dehors des limites du groupe d'utilisateurs ne pourront pas être analysées.
4	Device	devices Device Devices device	148	FR-small.xif	Properties panels display meters and settings for a selected networked device. Device parameters can be edited directly from the Properties panel. Device parameters	appareil	appareil	Les panneaux Propriétés affichent les indicateurs et les réglages pour un appareil en réseau sélectionné. Les paramètres des appareils peuvent être modifiés directement dans le panneau Propriétés. Paramètres des appareils
5	List	Lists lists List list	114	FR-small.xif	Inventory List In the Inventory tab, the list displays the equipment you have added to the show. Inventory list	Liste	Liste	Liste d'inventaire
6	file	file files Files File	82	FR-small.xif	You can save data for an event or show to a show file that you can use later or share: To import a list of single frequencies, select "Import Frequencies from File" from the More Options menu. In order to import frequencies from a file, that file must be specifically formatted.	fichier	fichier	Il est possible d'enregistrer des données pour un événement ou un spectacle dans un fichier de spectacle à utiliser ultérieurement ou à partager : Pour importer une liste de fréquences uniques, sélectionner « Importer des fréquences depuis un fichier » dans le menu Plus d'options. Afin d'importer des fréquences depuis un fichier, celui-ci doit être spécifiquement mis en forme.
7								

# Intelligent Corpus Alignment

A	B	C	D	E
Src Seg.	en-GB	hu-HU	Tgt Seg.	Probability
	Unmatched <span style="background-color: #f8d7da; border: 1px solid #f5c6cb; padding: 2px;"> </span> Not sure <span style="background-color: #fff3cd; border: 1px solid #ffeeba; padding: 2px;"> </span> Merged <span style="background-color: #d1ecf1; border: 1px solid #bee5eb; padding: 2px;"> </span> Very good <span style="background-color: #d4edda; border: 1px solid #c3e6cb; padding: 2px;"> </span>			
0	Before Adam by Jack London	NNCL1041-452v1.0	0	0.0000
1	1906	JACK LONDON Ádám előtt	1,2	1.0000
		Fordította: SZÁSZ IMRE A fedél és a belső illusztráció VISZT GYÖRGY munkája (c) Szász Imre, 1985 Hungarian translation	3	0.0000
		ISBN 963 11 3917 4	4	0.0000
		Móra Ferenc Ifjúsági Könyvkiadó, Budapest Felelős kiadó: Sziládi János igazgató Alföldi nyomda (5783.66-15-2), Debrecen, 1985 Felelős vezető: Benkő István igazgató Felelős szerkesztő: Szathmáry Éva Műszaki vezető: Szakálos Mihály Műszaki szerkesztő: Beszédes Natasa 67 000 példány, Terjedelem: 10,47 (A/5) ív.	5	0.0000
2	"These are our ancestors, and their history is our history.	IF 5331 "Ezek a mi őseink, az ő történetük a mi történetünk.	6	0.9287
3	Remember that as surely as we one day swung down out of the trees and walked upright, just as surely, on a far earlier day, did we crawl up out of the sea and achieve our first adventure on land."	Ne feledjük el: amilyen biztos az, hogy egy napon leereszkedtünk a fákról, és egyenesen kezdtünk járni, éppoly biztos, hogy egy sokkalta régebbi napon kimásztunk a tengerből, és átéltük első kalandunkat a szárazföldön."	7,8	1.0000
4	CHAPTER I	1.	9	1.0000
5	Pictures!	Képek!	10	1.0000
6	Pictures!	Képek!	11	1.0000
7	Pictures!	Képek!	12	1.0000
8	Often, before I learned, did I wonder whence came the multitudes of pictures that thronged my dreams; for they were pictures the like of which I had never seen in real wake-a-day life.	Míg meg nem tudtam, gyakran csodálkoztam, honnan származik az a rengeteg kép, ami előzőnötte álmaimat; mert ilyen képeket sose láttam a nappal, az ébrenlét valóságában.	13	1.0000
9	They tormented my childhood, making of my dreams a procession of nightmares and a little later convincing me that I was different from my kind, a creature unnatural and accursed.	Már gyerekkoromban gyötörtek, álmaimat lidércnyomások sorozatává tették, s kissé később azt a meggyőződést ültették el bennem, hogy más vagyok, mint a fajtám: természetellenes, elátkozott lény.	14,15	0.9594
10	In my days only did I attain any measure of happiness.	Csak nappal érezhettem némi kis boldogságot.	16	0.6523
11	My nights marked the reign of fear - and such fear!	Éjszakáim a félelem uralmát jelentették - és kicsoda félelemém!	17	1.0000

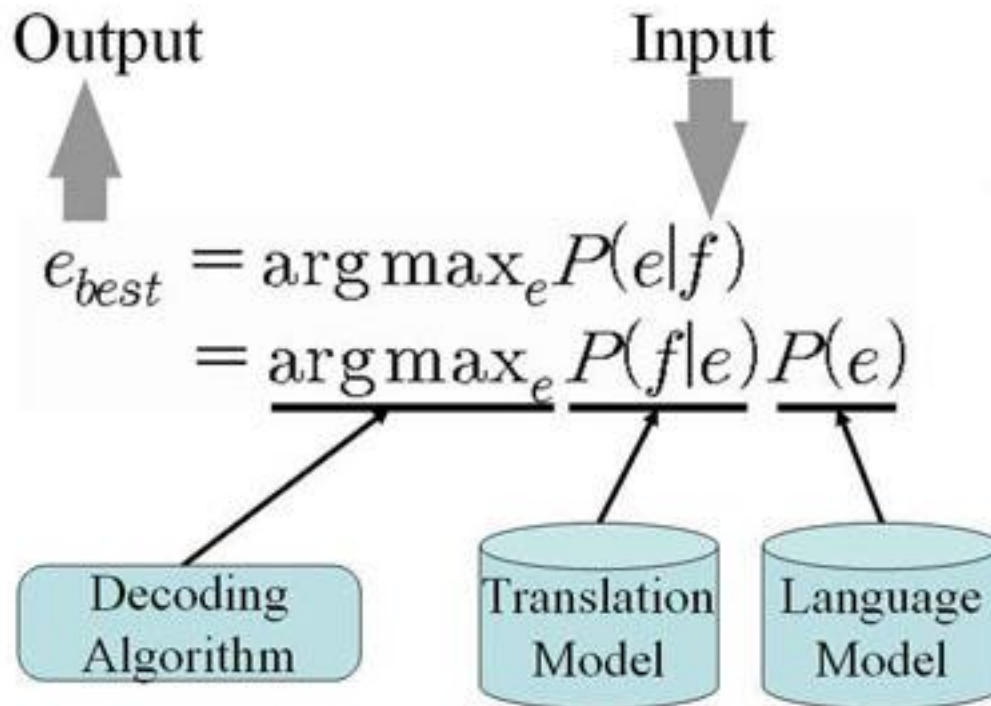


# AI in Localization

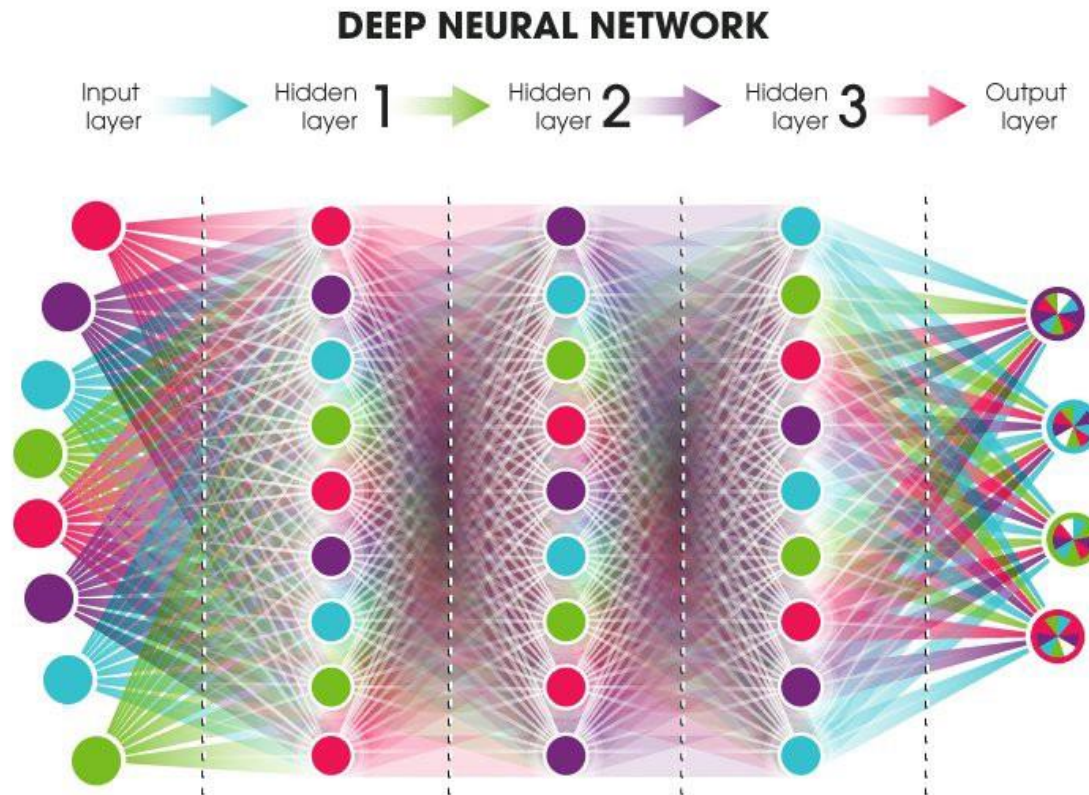
## Bayesian probabilities

1. MT post editing: recognizing common patterns
2. Automated placement of inline elements
3. Dynamic sub-segment phrase matching
4. Learning translation

# Statistical MT



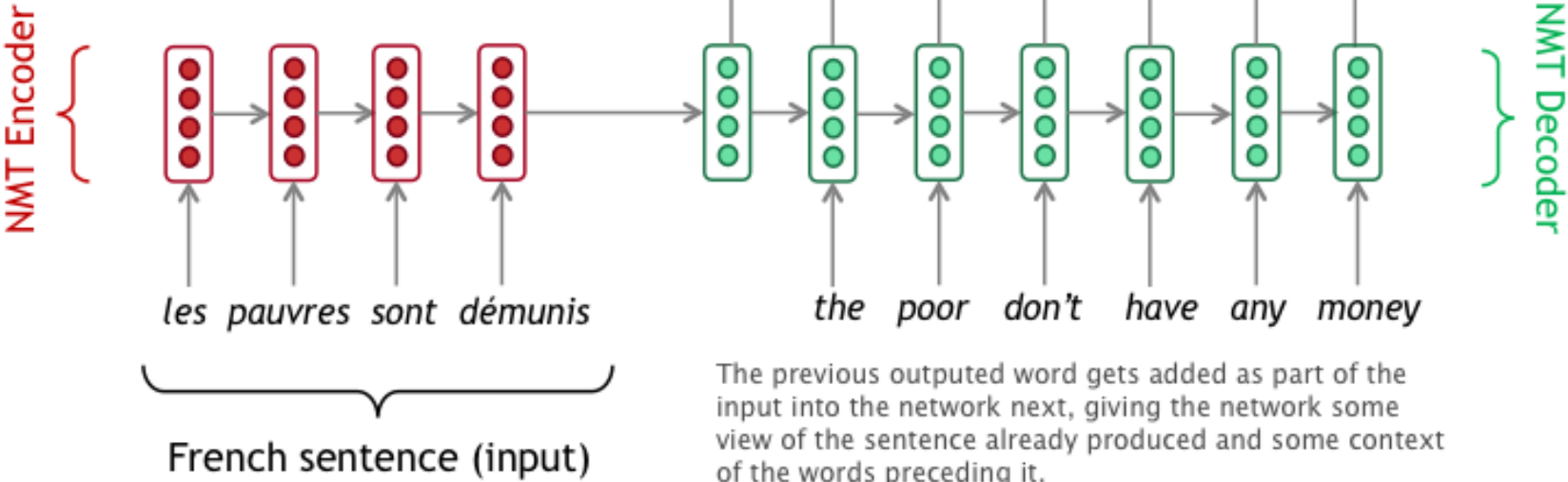
# Machine Learning: Neural Networks



neuralnetworksanddeeplearning.com - Michael Nielsen, Yoshua Bengio, Ian Goodfellow, and Aaron Courville, 2016.

# Neural MT: Deep Learning

We feed in each word from left to right, one at a time. By the end, the NMT system has encoded information about the whole sentence in a numerical format.



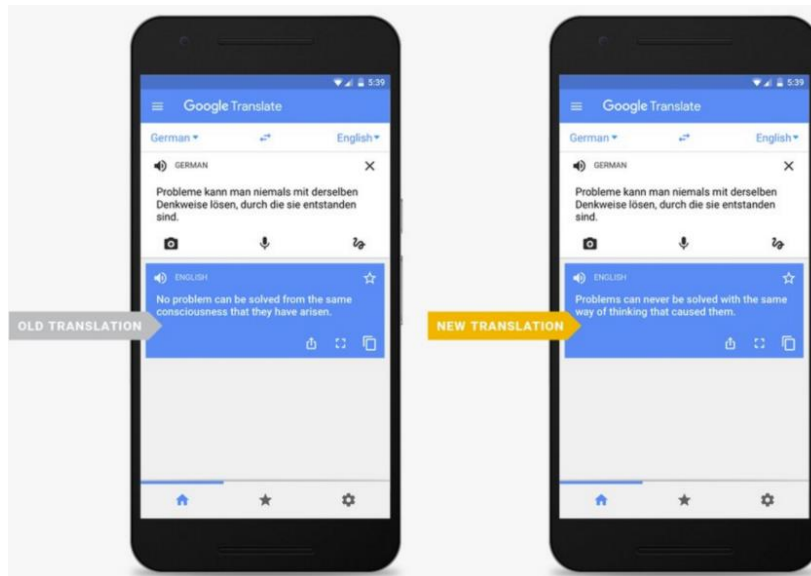
# Neural MT

## SMT:

No problem can be solved from the same consciousness that they have arisen

## NMT:

Problems can never be solved with the same way of thinking that caused them.



# NMT Assessment

- *Improved translation quality for morphologically rich languages*
- *Fluency is improved, word order errors are fewer, specially when translating from English into languages with rich morphologies*
- *Fewer segments require editing*
- *Fewer morphological errors*
- *Much smaller output engine*
- *No clear improvement for omission or mistranslation*

Joss Moorkens (DCU, ASLING TC38):

# NMT Assessment

- *Mistakes can be harder to spot: cognitive 'fixing'*
- *Issues when translating from morphologically rich languages to ones with primitive morphology*
- *Cannot cope with OOV words*
- *Very sensitive to spelling and grammatical errors in source*
- *Tuning is effectively impossible*
- *Very poor performance on out-of-domain text*

# NMT Assessment

French source: avez-vous de la sauce de salade au bleu ?

SMT translation: do you have sauce of salad in blue?

NMT translation: do you have any chicken salad?

French source: c'est le contrat d'achat de mes chèques de voyage.

SMT translation: it's the purchase agreement of my checks.

NMT translation: it's the seniority wage system.

French source: les adolescents japonais aiment les jeux vidéos .

SMT translation: the adolescents japanese love electronic vidéos .

NMT translation: japanese teenagers are interested in fashion .

French source: je voudrais manger de la vraie nourriture indienne

SMT translation: I'd like to have true food indienne

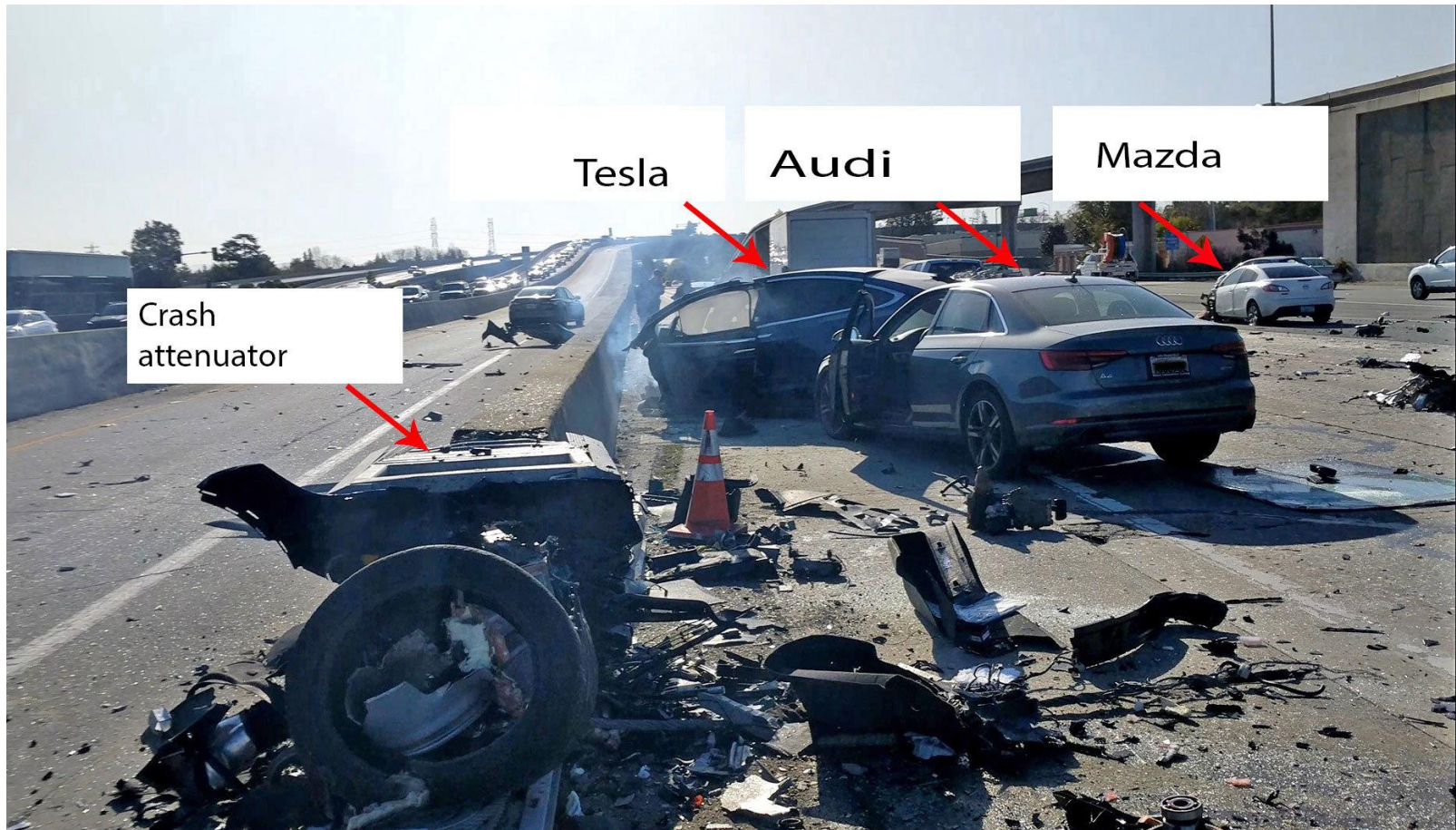
NMT translation: I'd like to eat some food food



# AI Fail!



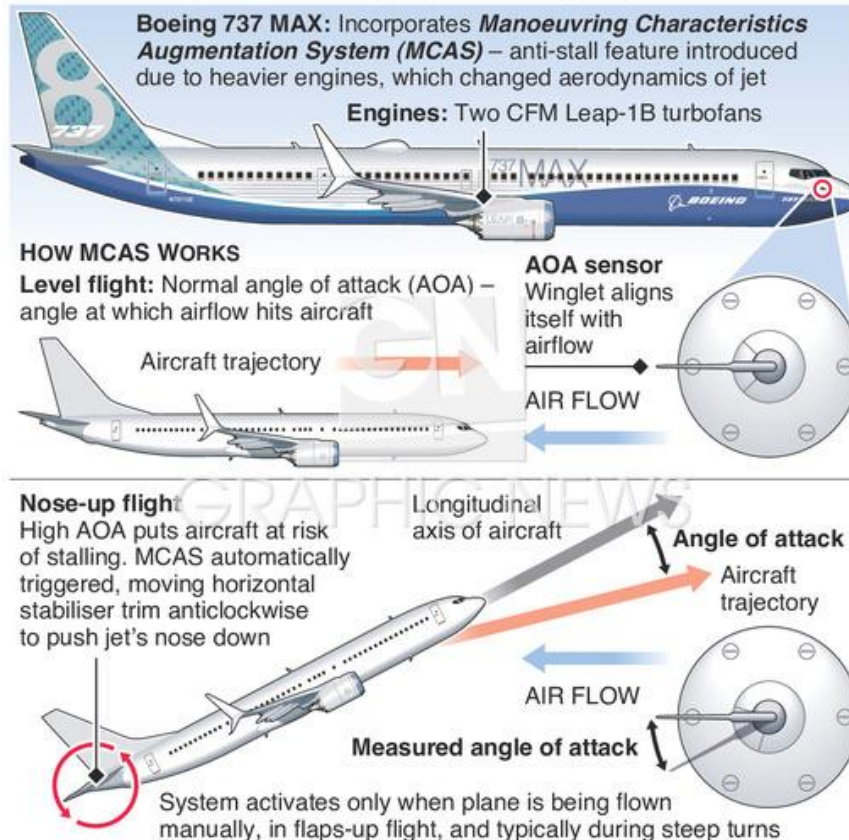
# 80% rule: Tesla Autopilot Fail!



# What were Boeing thinking?

## Boeing “withheld information” on 737 MAX

Boeing failed to tell airline pilots about features of a new flight-control system suspected of playing a role in the fatal crash of an Indonesian 737 MAX plane, according to pilots who fly the jet in the U.S.



In Oct 29 Lion Air crash that killed 189 people in Indonesia, investigators have determined that AOA sensor was feeding bad data to jet's flight computer, repeatedly activating MCAS when there was no risk of stall

# Question and Answer session



# Contact Details



XTM International  
[www.xtm-intl.com](http://www.xtm-intl.com)

Register for future Webinar sessions



[www.xtm-intl.com/demos](http://www.xtm-intl.com/demos)

Contact



[azydron@xtm-intl.com](mailto:azydron@xtm-intl.com)  
+44 (0) 1753 480 479