Donde empieza la Innovación: El Laboratorio de Investigación de IBM Zurich

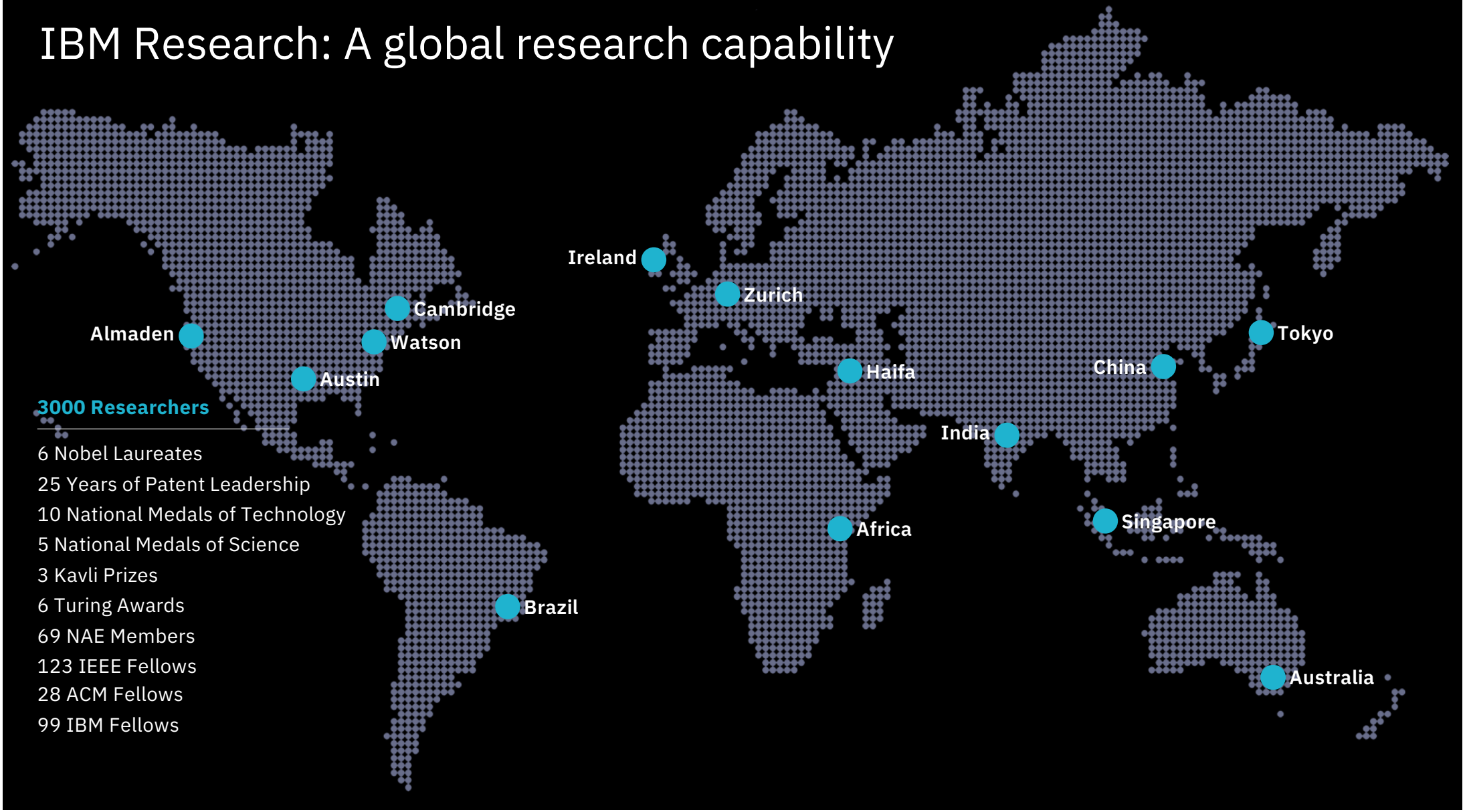
Cristiano Malossi, IBM Research THINKLab

@C_Malossi

#IBMCloud

IBM

IBM Research: A global research capability



3000 Researchers

6 Nobel Laureates
25 Years of Patent Leadership
10 National Medals of Technology
5 National Medals of Science
3 Kavli Prizes
6 Turing Awards
69 NAE Members
123 IEEE Fellows
28 ACM Fellows
99 IBM Fellows

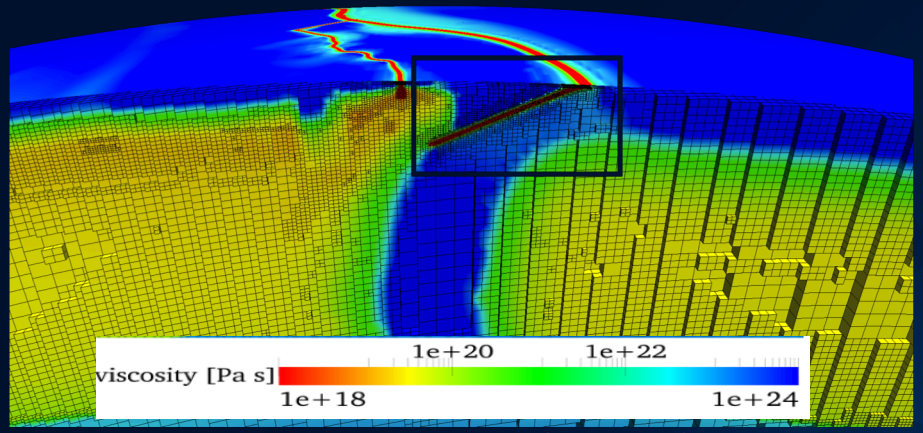
IBM Research - Zurich

- Established in 1956
- 45+ different nationalities
- Open Collaboration:
  - Horizon2020: 43 funded projects and 500+ partners
- Two Nobel Prizes, One Kavli Prize
  - 1986: Nobel Prize in Physics for the invention of the scanning tunneling microscope by Heinrich Rohrer and Gerd K. Binnig
  - 1987: Nobel Prize in Physics for the discovery of high-temperature superconductivity by K. Alex Müller and J. Georg Bednorz
  - 2016: Kavli Prize in Nanoscience for the invention and realization of atomic force microscopy by Gerd K. Binnig and Christoph Gerber
- Binnig and Rohrer Nanotechnology Centre opened in 2011 (Public Private Partnership with ETH Zürich and EMPA)
- 7 European Research Council Grants
- 2017 - Named Historical Site by the European Physical Society

Major Historic Accomplishments

Token Ring	Scanning Tunneling Microscope	High Temperature Superconductivity	Noise-Predictive Maximum-Likelihood	JavaCard (JCOP)	OLED	Millipede

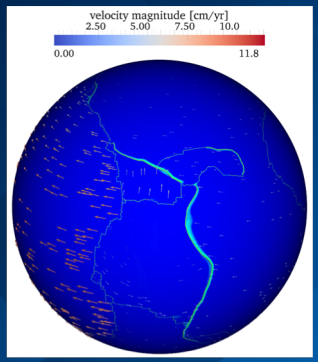
Success in Petascale computing: implicit linear solver to scale!

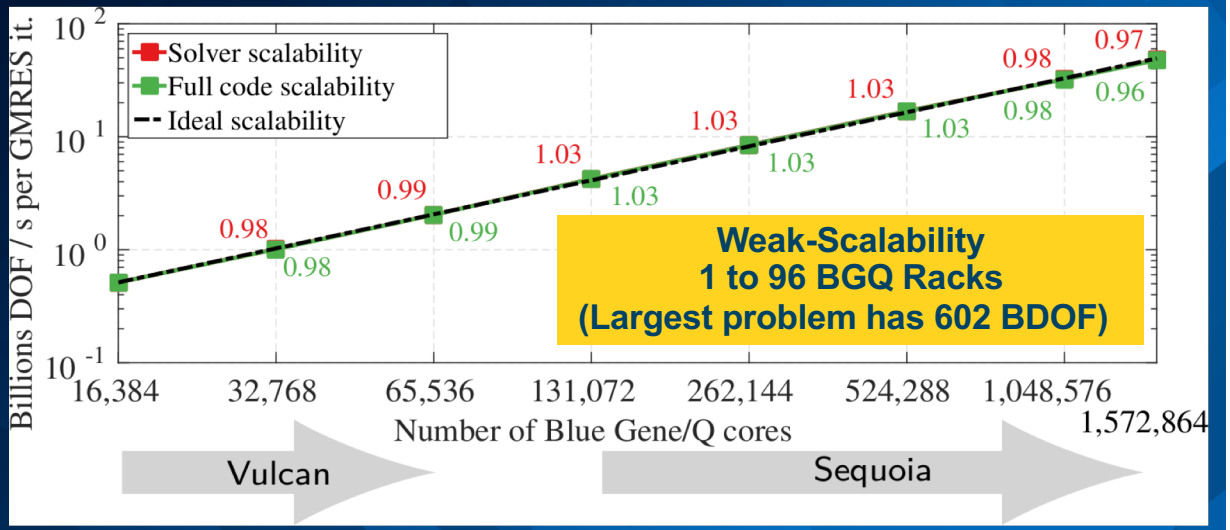


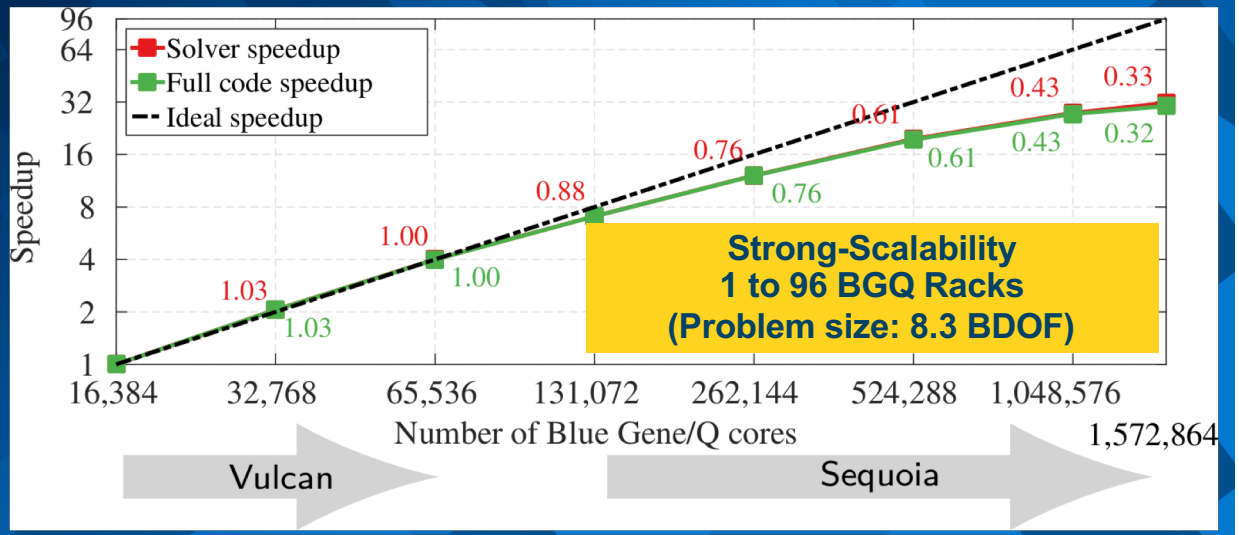
An Extreme-scale Implicit Solver for Complex PDEs: Highly Heterogeneous Flow in Earth's Mantle

J. Rudi[1], A.C.I. Malossi[2], T. Isaac[1], G. Stadler[3], M. Gurnis[4], P.W.J. Staar[2], Y. Ineichen[2], C. Bekas[2], A. Curioni[2], O. Ghattas[1]

1: The University of Texas at Austin
2: IBM Research – Zurich
3: New York University
4: California Institute of Technology

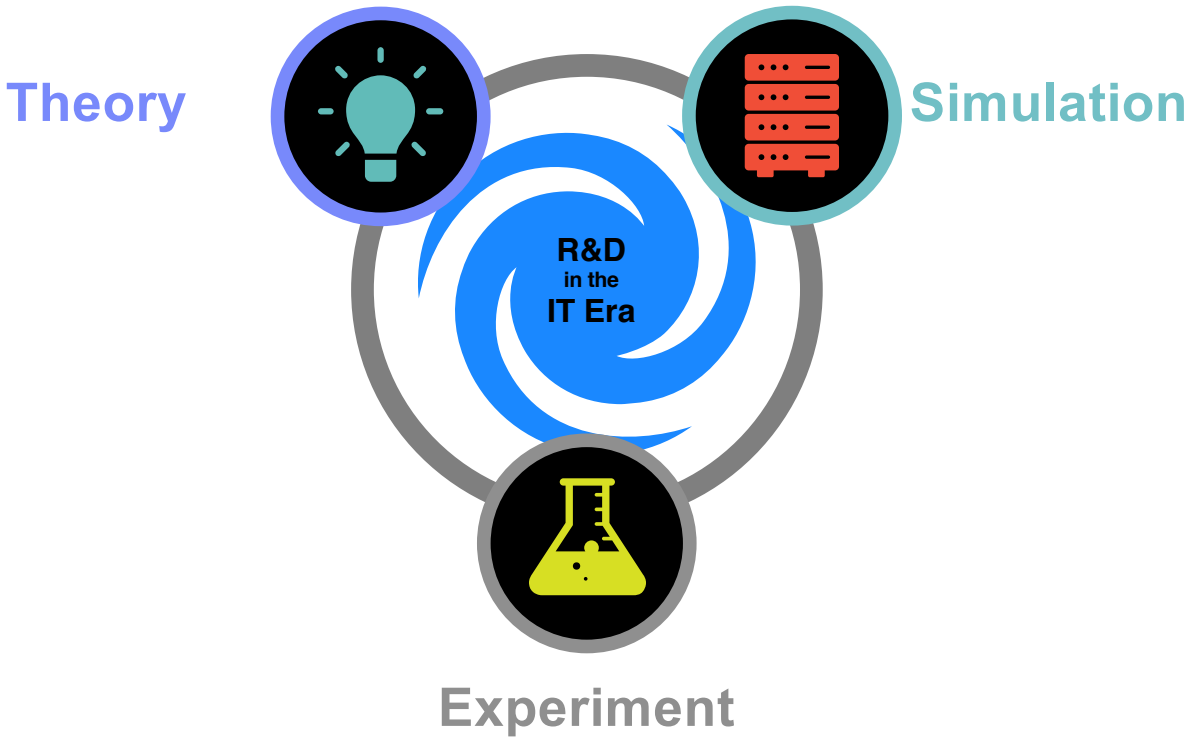




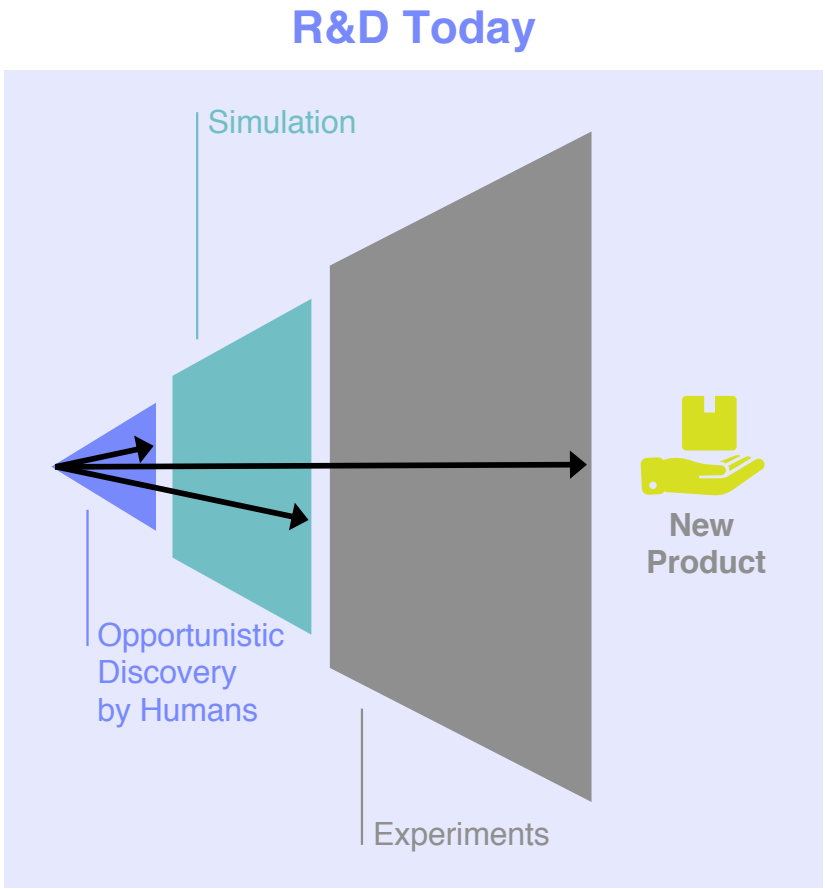


Cognitive Discovery

Technical R&D today: The three pillars

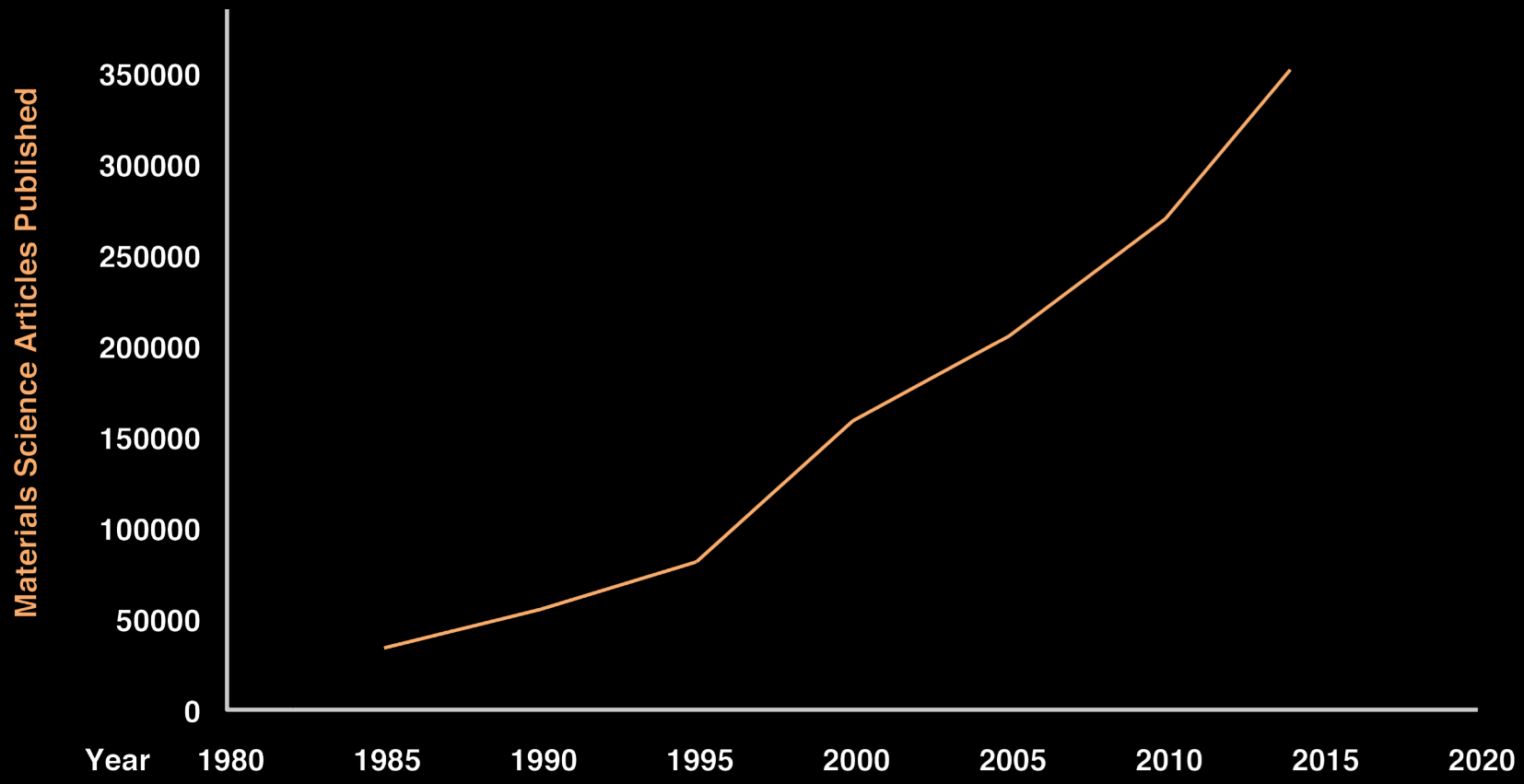


Traditional R&D has limits

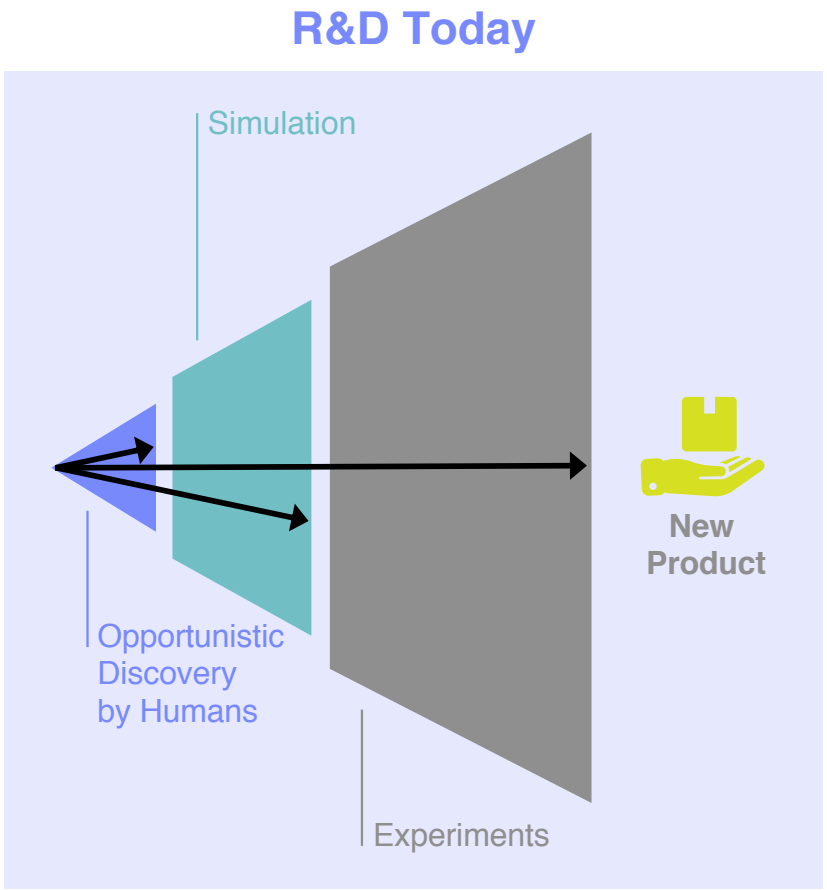


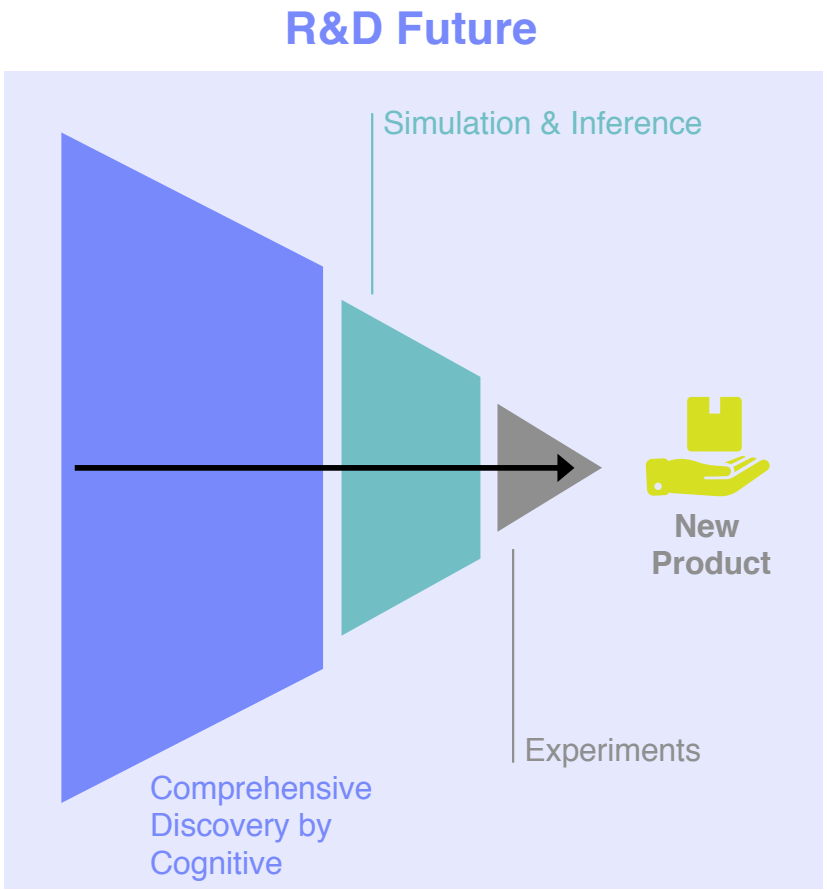
- **We cannot beat complexity with brute force simulation**
- **We need a new, data driven, holistic approach**

Unstructured Data Deluge in Peer Review Publications



Cognitive Discovery based on AI to reverse the pyramid





Scalable Knowledge Ingestion

PDF Parser → PDF Interpretation → Semantic representation

PRL 108, 190603 (2012) PHYSICAL REVIEW LETTERS week ending 11 MAY 2012

Exact Solution for the Stationary Kardar-Parisi-Zhang Equation

Takashi Imamura[1,*] and Tomohiro Sasamoto[2,†]

[1]*Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro, Tokyo 153-8904, Japan*
[2]*Department of Mathematics and Informatics, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan*
(Received 23 November 2011; published 10 May 2012)

We obtain the first exact solution for the stationary one-dimensional Kardar-Parisi-Zhang equation. A formula for the distribution of the height is given in terms of a Fredholm determinant, which is valid for any finite time t. The expression is explicit and compact enough so that it can be evaluated numerically. Furthermore, by extending the same scheme, we find an exact formula for the stationary two-point correlation function.

DOI: 10.1103/PhysRevLett.108.190603 PACS numbers: 05.10.Gg, 05.40.−a, 64.70.kj

The dynamics of a growing surface has long been a subject of intense study because of its wide appearance in nature and relevance to material science; it is also of great theoretical interest as an important example of nonlinear nonequilibrium phenomena [1,2]. For a surface growing only through local interaction, a prototypical equation was proposed by Kardar, Parisi, and Zhang (KPZ) in 1986 [3]. In this Letter, we focus on its one-dimensional version,

$$\frac{\partial h(x,t)}{\partial t} = \frac{\lambda}{2}\left(\frac{\partial h(x,t)}{\partial x}\right)^2 + \nu\frac{\partial^2 h(x,t)}{\partial x^2} + \eta(x,t). \quad (1)$$

Here $h(x,t)$ represents the height of the surface at position $x \in \mathbb{R}$ and time $t \geq 0$. The last term $\eta(x,t)$ is taken to be the Gaussian white noise with covariance $\langle \eta(x,t)\eta(x',t')\rangle = D\delta(x-x')\delta(t-t')$.

This KPZ equation has been well known and has already been studied extensively [4–8], but interest in this equation has revived [9]. First, the efforts to have a better understanding of the equation, which are from various perspectives and have lasted for more than two decades, finally culminated in the achievement of an exact solution for its height distribution [10,11]; second, it has turned out that such fine information as the height distribution can be measured experimentally. In [12], Takeuchi and Sano employed a turbulent liquid crystal as a material and found a clear agreement with theoretical predictions.

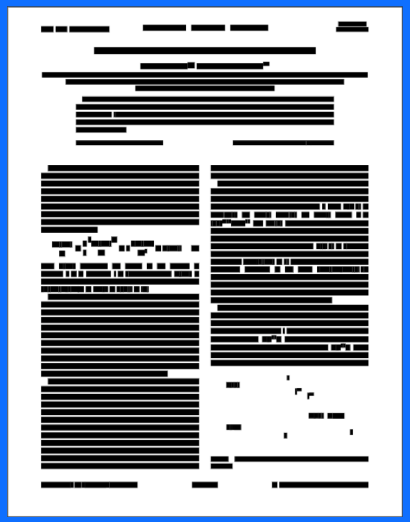
In the studies of growing surface, the fluctuations of the height are of primary importance. In many examples one observes the "kinetic roughening." As time goes on, the surface becomes rougher and rougher, even if it starts from a completely flat substrate. This concerns the transient behaviors which are certainly interesting, but even more natural and important is the stationary situation. Most systems eventually reach their own steady states after relaxation times, and our understanding of the nonequilibrium steady state (NESS) is in general so limited that concrete results for a simple system should give us valuable information. Many papers on the KPZ equation have addressed its stationary properties with various methods, but all attempts to treat this case exactly have so far failed.

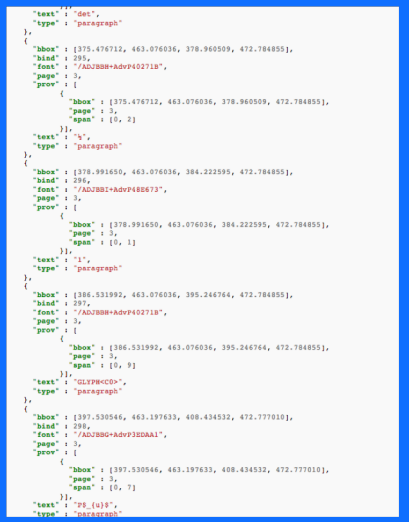
In this article, we provide the first exact results for the KPZ equation in the stationary situation. For the KPZ equation, the stationary state is given by the two-sided Brownian motion (BM) with respect to z [13]: $h(x,0) = \alpha B_+(-x)$ for $x < 0$, $\alpha B_-(x)$ for $x > 0$, where $\alpha = (2\nu)^{-1/2}\lambda D^{1/2}$ and $B_\pm(x)$ are two independent standard BMs, as seen in Fig. 1. Since the system is translationally invariant, one can set without loss of generality the initial height at the origin to be zero, i.e., $h(0,0) = 0$. In the following, we will show the explicit formulas for the height distribution $\mathrm{Prob}[h(x,t) \geq s]$ and the two-point space-time correlation functions of the slope $\langle \partial_x h(x,t)\partial_x h(0,0)\rangle$ (Fig. 1). The obtained expressions are explicit enough so that they can be evaluated numerically. We would say that the main achievements in our paper are that we could draw their figures as given in Figs. 2 and 3 below.

Let us briefly recall the theoretical developments so far. Early works using renormalization group ideas and other analytic methods have identified the critical exponents. It was established that when t is large, the fluctuations of the height grow like $O(t^{1/3})$ and the nontrivial equal-time correlations in space are observed on the $O(t^{2/3})$ scale. About a decade ago, it was realized that a few growth models admit a mapping to a combinatorial problem,

FIG. 1. The two-sided BM initial condition and the two-point function.

0031-9007/12/108(19)/190603(5) 190603-1 © 2012 American Physical Society



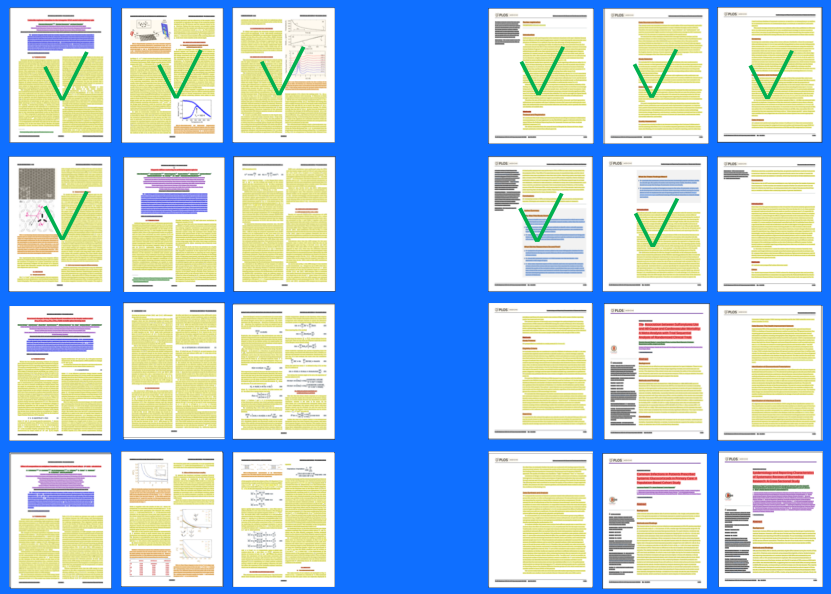



```
        {},
    "text" : "det",
    "type" : "paragraph"
},
{
    "bbox" : [375.476712, 463.076036, 378.960509, 472.784855],
    "bind" : 295,
    "font" : "/ADJBBH+AdvP40271B",
    "page" : 3,
    "prov" : [
        {
            "bbox" : [375.476712, 463.076036, 378.960509, 472.784855],
            "page" : 3,
            "span" : [0, 2]
        }],
    "text" : "k",
    "type" : "paragraph"
},
{
    "bbox" : [378.991650, 463.076036, 384.222595, 472.784855],
    "bind" : 296,
    "font" : "/ADJBBI+AdvP4BE673",
    "page" : 3,
    "prov" : [
        {
            "bbox" : [378.991650, 463.076036, 384.222595, 472.784855],
            "page" : 3,
            "span" : [0, 1]
        }],
    "text" : "1",
    "type" : "paragraph"
},
{
    "bbox" : [386.531992, 463.076036, 395.246764, 472.784855],
    "bind" : 297,
    "font" : "/ADJBBH+AdvP40271B",
    "page" : 3,
    "prov" : [
        {
            "bbox" : [386.531992, 463.076036, 395.246764, 472.784855],
            "page" : 3,
            "span" : [0, 9]
        }],
    "text" : "GLYPH<C0>",
    "type" : "paragraph"
},
{
    "bbox" : [397.530546, 463.197633, 408.434532, 472.777010],
    "bind" : 298,
    "font" : "/ADJBBG+AdvP3EDAA1",
    "page" : 3,
    "prov" : [
        {
            "bbox" : [397.530546, 463.197633, 408.434532, 472.777010],
            "page" : 3,
            "span" : [0, 7]
        }],
    "text" : "P$_{u}$",
    "type" : "paragraph"
```


1. SYSML 2018 (https://www.sysml.cc/doc/76.pdf)
2. KDD 2018 (https://doi.org/10.1145/3219819.3219834)

Knowledge Extraction: PDF – Accelerate via ML

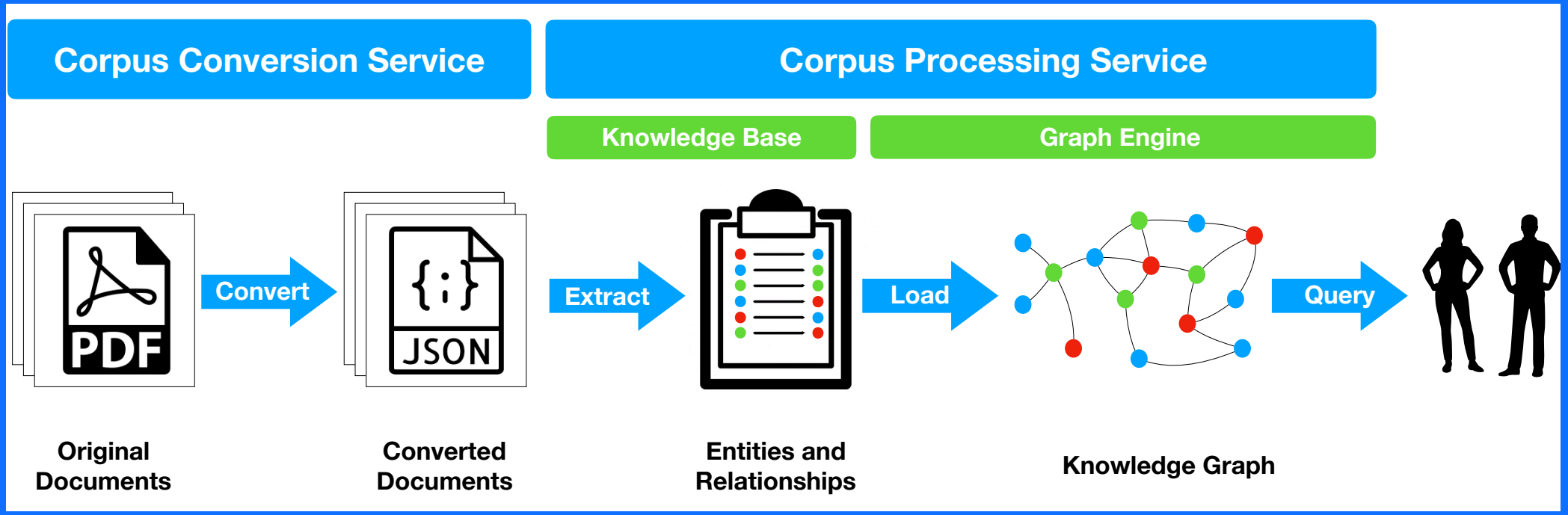
1. Categorize documents into similar layouts
2. Annotate a few pages
3. A Machine Learning model is trained automatically
4. Huge boost in the annotation process



https://samaster.smartannotator-production-v2.zc2.ibm.com/manager/projects/afb2ddec29d5179ab415e1061e627837f886e271

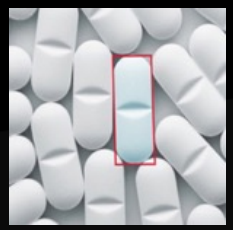
Accuracy > 98%

Corpus Processing Service



AI Automation

Narrow AI: Initial Value Creation

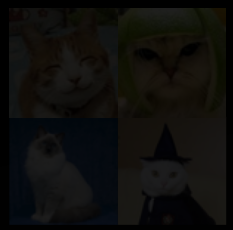


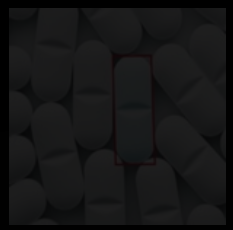
2010 *2015*

AI learns to solve specific tasks, or focuses on **individual domains** or modalities primarily using **human-curated , training data sets** & **manually-crafted architectures**

Narrow AI:
Initial Value
Creation

General AI:
Revolutionary





2010

2015

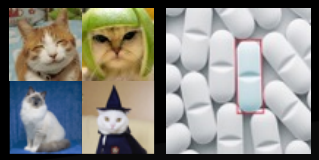
2050–Beyond

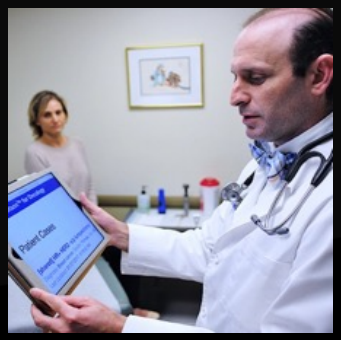
AI acquires knowledge by **reading, discussion, observation, experiments**. **Broad transfer of knowledge** across tasks. **Cross-domain reasoning** is common. **Broad autonomy** within human-managed teams.

Narrow AI:
Initial Value
Creation

Broad AI:
Disruptive and
Pervasive

General AI:
Revolutionary





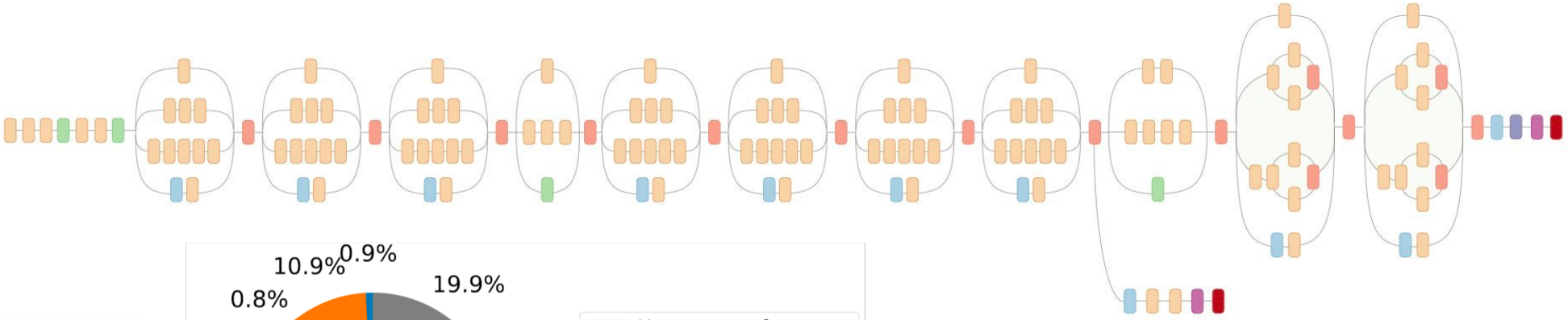
2010 *2015*

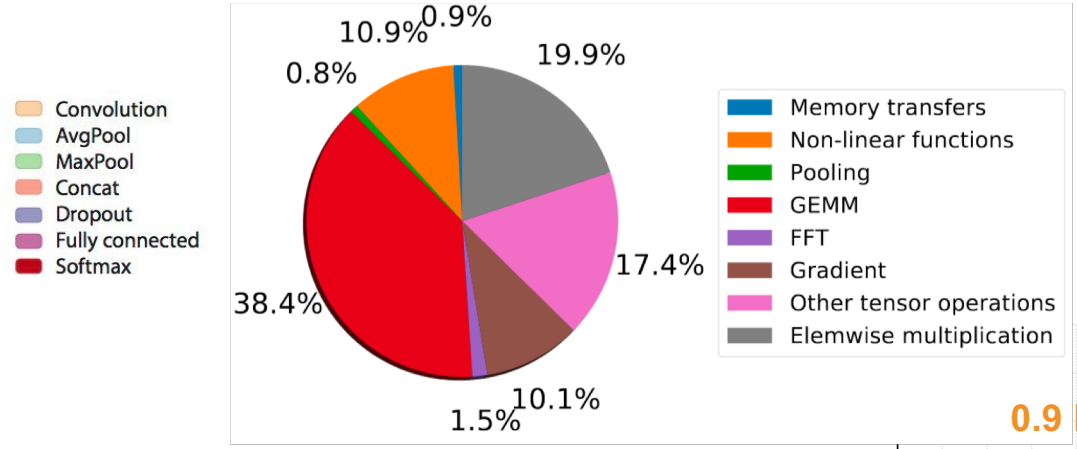
We are here

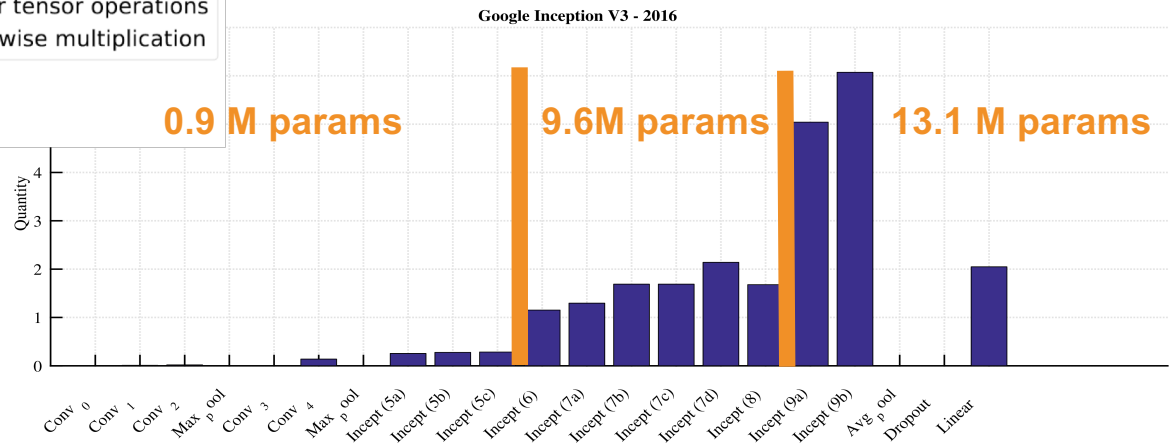
2050–Beyond

AI selectively acquires and integrates knowledge from **multiple modalities and sources**, including interaction; develops and **retains skills that it adapts and combines** to complete new tasks; learning is adaptive, using automatically-constructed architectures.

Today's neural networks: how do they look like?





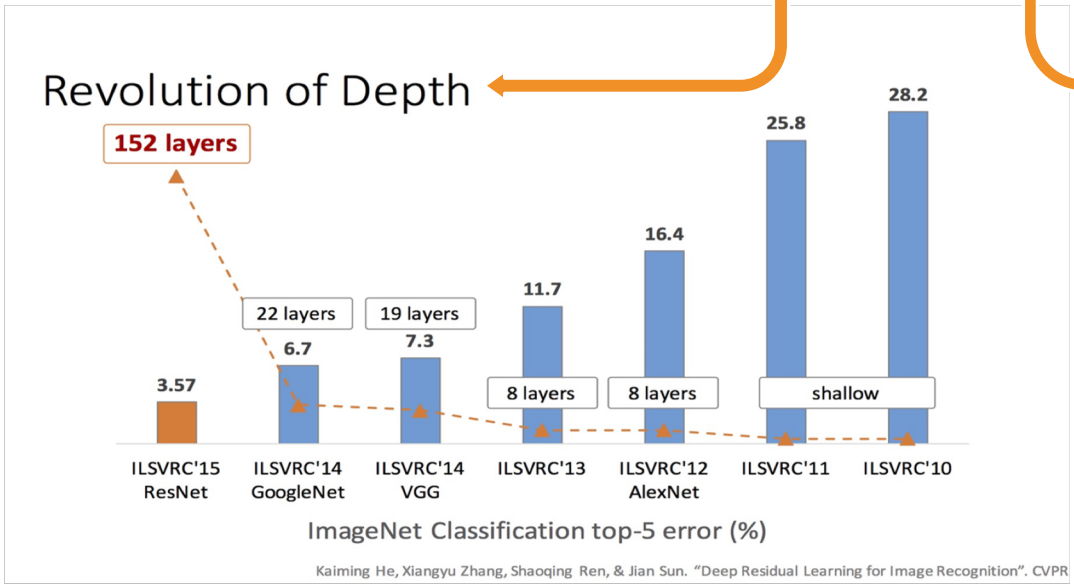


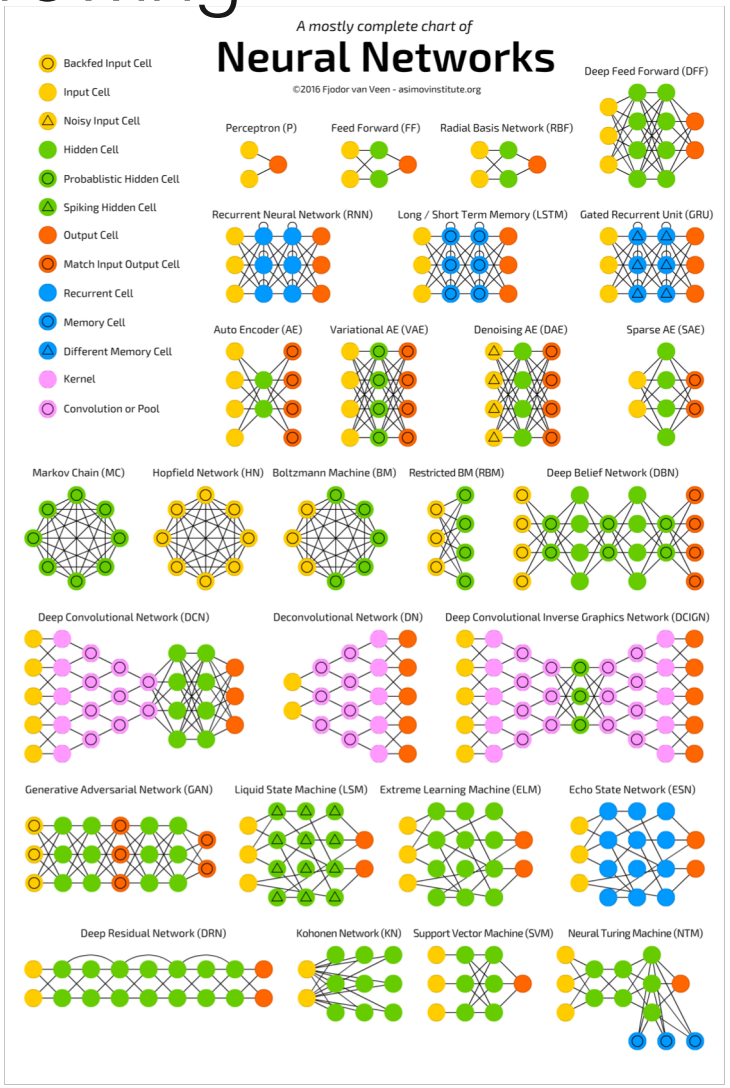
Inception V3 architecture:

C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna. Rethinking the Inception Architecture for Computer Vision, 2015.

Millions of parameters to tweak and growing

- Highly skilled researchers/data-scientists are needed to hand-craft custom neural networks
- Hand-crafting complex networks is time-consuming, error prone, and does not scale with time and resources
- Neural networks continue to grow in size and complexity





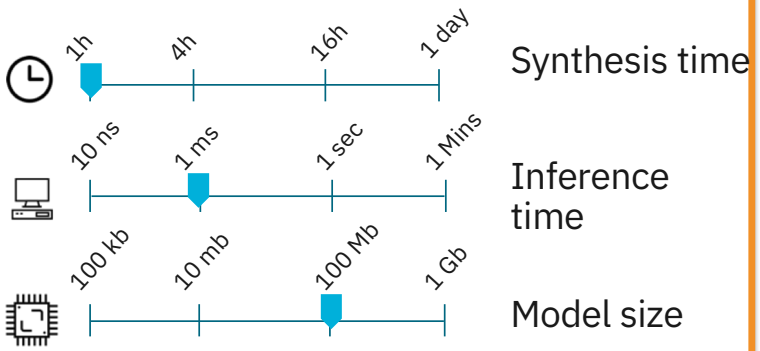
How does it work?

1) USER INPUT

Dataset

Images, Text, …

Budget & requirements (hidden in beta)



2) AI AUTOMATION: FROM DATA PREPARATION TO MODEL SYNTHESIS

DATA PREPARATION

- Ingestion
- Cleansing
- Augmentation

MODELING

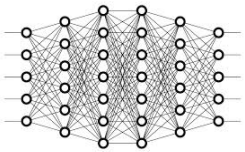
- Prediction
- Arch. Search
- Model selection

SYNTHESIS

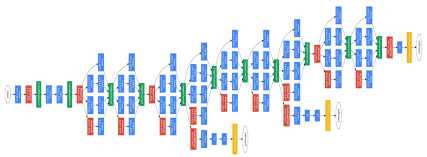
- Training
- HPO
- Validation

3) DEPLOY/DOWNLOAD MODEL

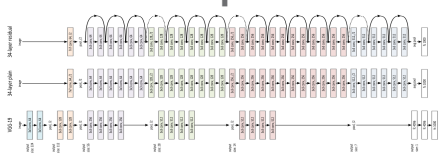
Example #1



Example #2



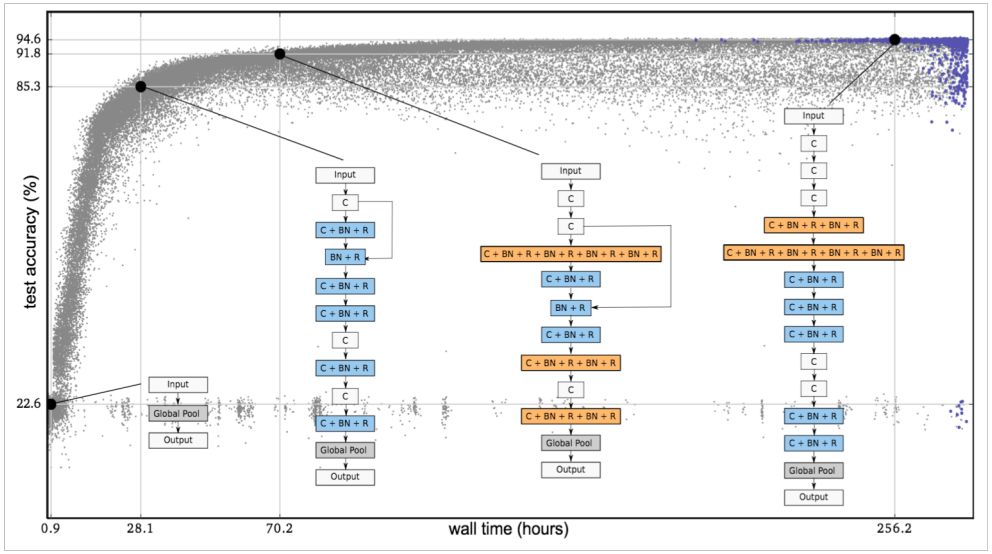
Example #3



…

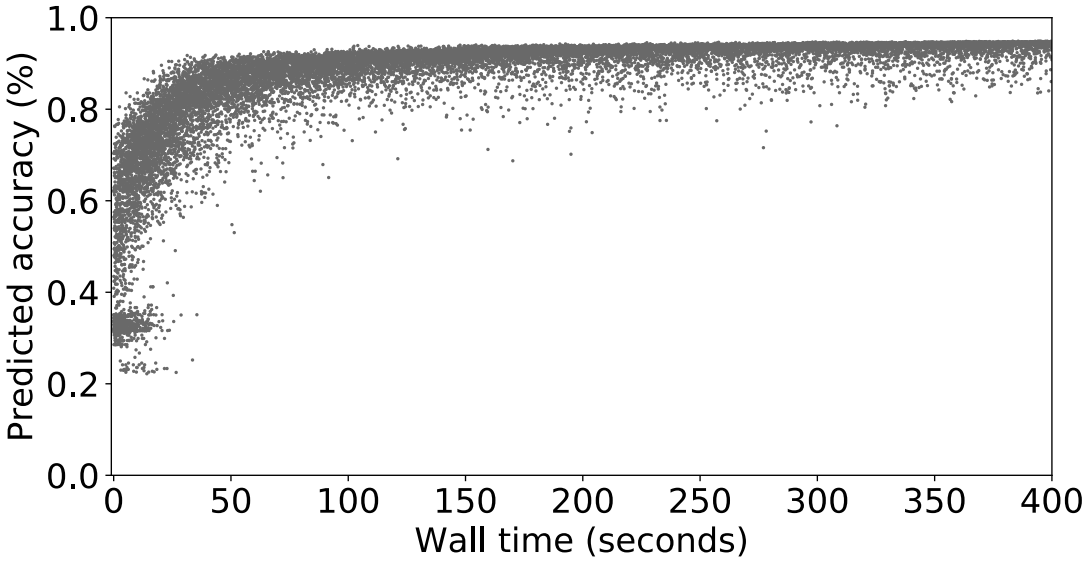
Classical LSE

(Real et al. 2017)



TAPAS-NeuNetS

(Istrate et al. 2018)

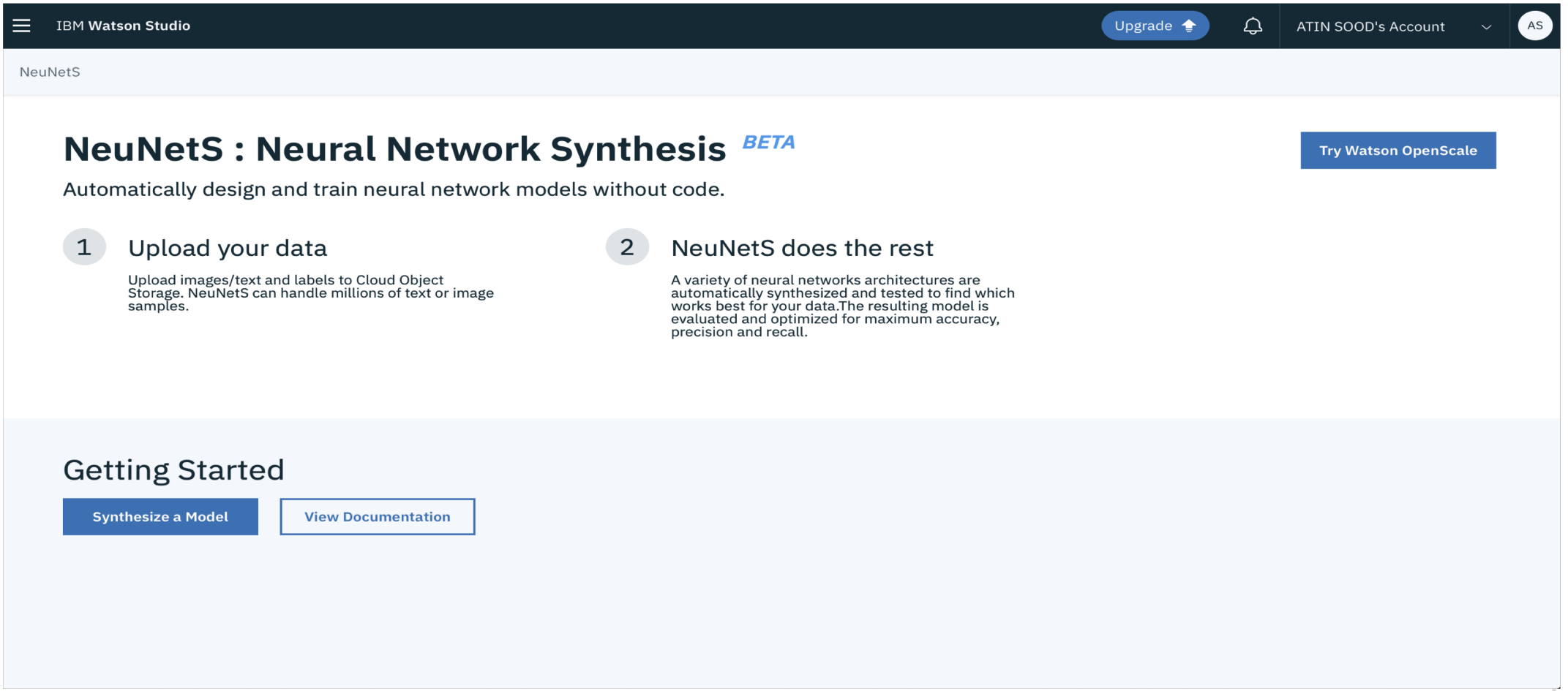


- Search time: 256h
- Resources: 250 workers (~250-1000 GPUs)
- CIFAR-10 top accuracy: 94.6%
- CIFAR-100 top accuracy: 77%

- Search time: 400s (2300x faster)
- Resources: 1 GPU (250-1000x less expensive)
- CIFAR-10 top accuracy: 93.67% (1% worse)
- CIFAR-100 top accuracy: 81.01% (4% better)

SPEED UP ORDER: 1M (architecture discovery) – 100k (trained model)

NeuNetS: Neural Network Synthesis

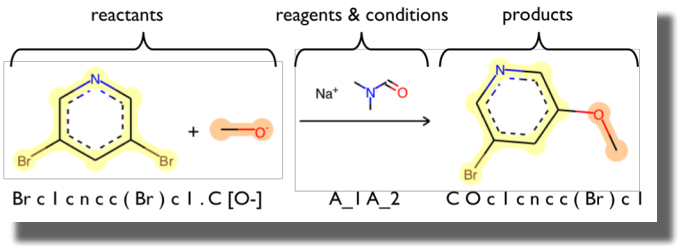


Try it now! http://ibm.biz/neunets

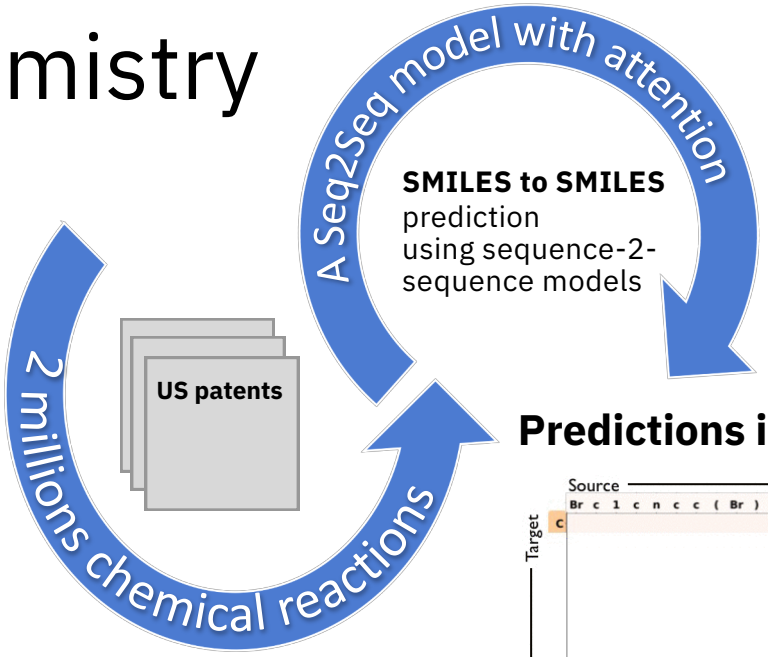
AI for Chemistry

IBM **RXN** for Chemistry

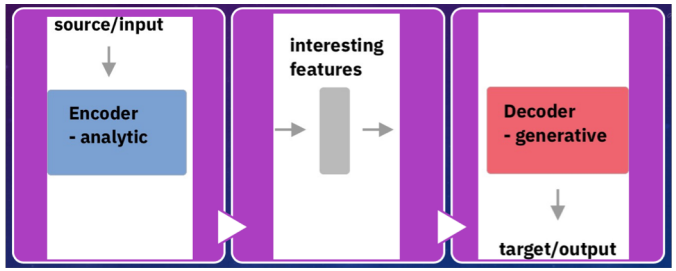
Forward Chemical reaction prediction



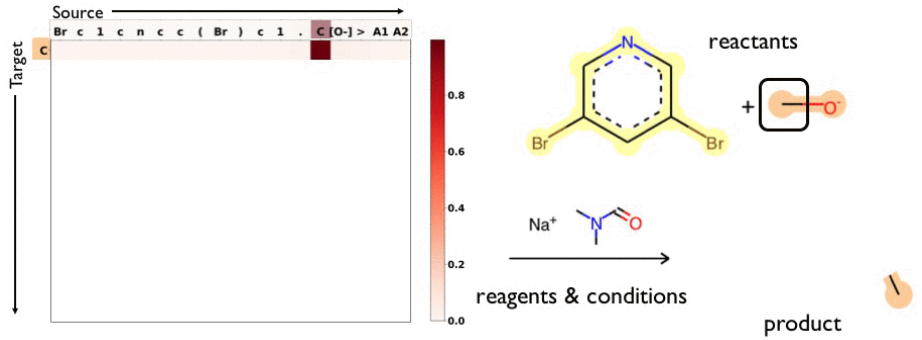
SMILES to SMILES



SMILES to SMILES
prediction
using sequence-2-sequence models



Predictions in short (how a Se2Seq works)



Schwaller et. al. Chem. Sci., 2018, 9, 6091-6098

Freely available for everyone:
http://www.research.ibm.com/ai4chemistry

Released on August 19th at ACS Boston. In 6 months (Dec. 2018):

- 5150 active users
- 20000 chemical reactions predictions
- 8000 projects opened
- 60000 molecules generated

Performance

USPTO*		S2S [13]	WLDN [28]	ELECTRO [30]	GTPN [31]	WLDN5 [29]	our work
_MIT	separated	80.3	79.6		82.4	85.6	**90.4**
_MIT	**mixed**		74				**88.6**

Currently, best method in FORWARD chemical reaction prediction

Muchas gracias

Cristiano Malossi, IBM Research THINKLab

@C_Malossi

#IBMCloud

IBM