Cognitive Computing

Progress & Challenges

David Nahamoo,

IBM Fellow

11/15/2016

A Point of View (So What?)

- Application – Checkpoint is more than Sensing
  - Big Data Analytics (People, Activity, Behavior, Emotion, …)
  - From Passive Data to Engaging Data
- Delivery – System of Systems is extremely difficult
  - Single point of integration by One organization
  - From Biometrics to Biomarkers
- Technology – End of Super Compressed Formulas
  - Deep Neural Networks
  - Vector Representation (Language, Reasoning!)

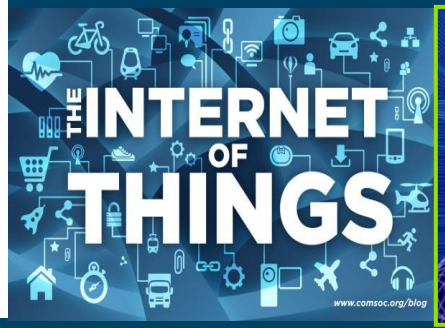
Digital Transformation: Evolution of Computing

Digitization of transaction

Digitization of enterprise

Digitization of commerce

Digitization of interactions

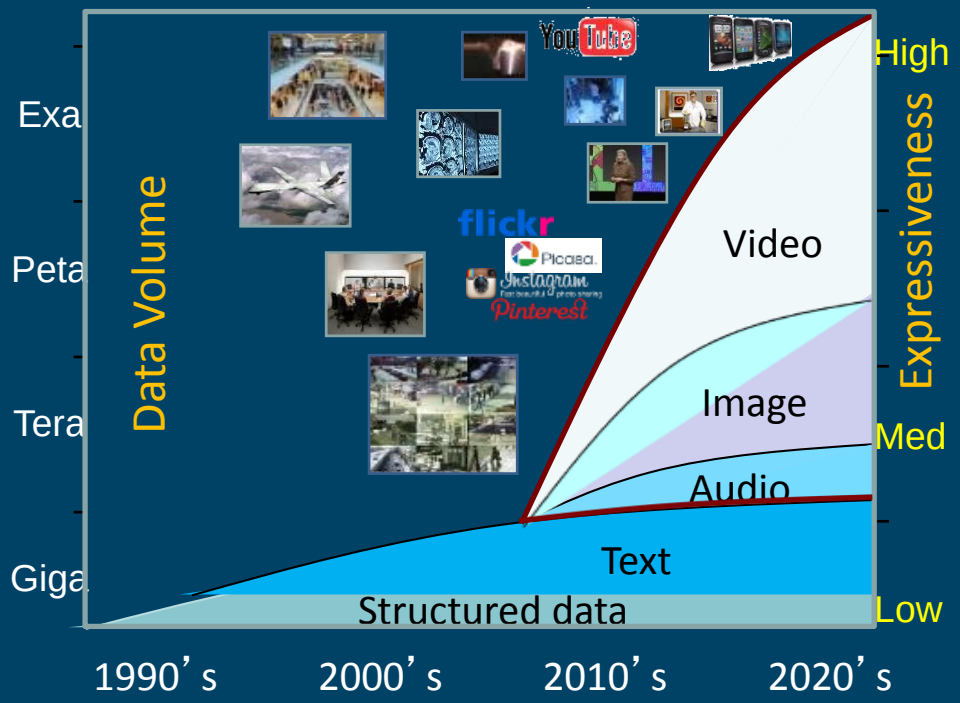


Digitization of environment

Digitization of reasoning, learning, and knowledge

A New Era of Computing…

Cognitive Systems learn and interact naturally with people to amplify what either humans or machines could do on their own.



Modeling Intelligence

Humans	• Fast Thinking	• Slow Thinking	• Emotions
Machines	• Conversational Automated Experts	• Creative Discovery & Innovation Machines	• Interpersonal Skills

Computing	Knowledge	Processing	Build	Form
Past	• Structured • Certain	• Accurate • Precise	• Program	• Rules
Future	• Unstructured • Ambiguous	• Inaccurate • Imprecise	• Learn	• Statistics

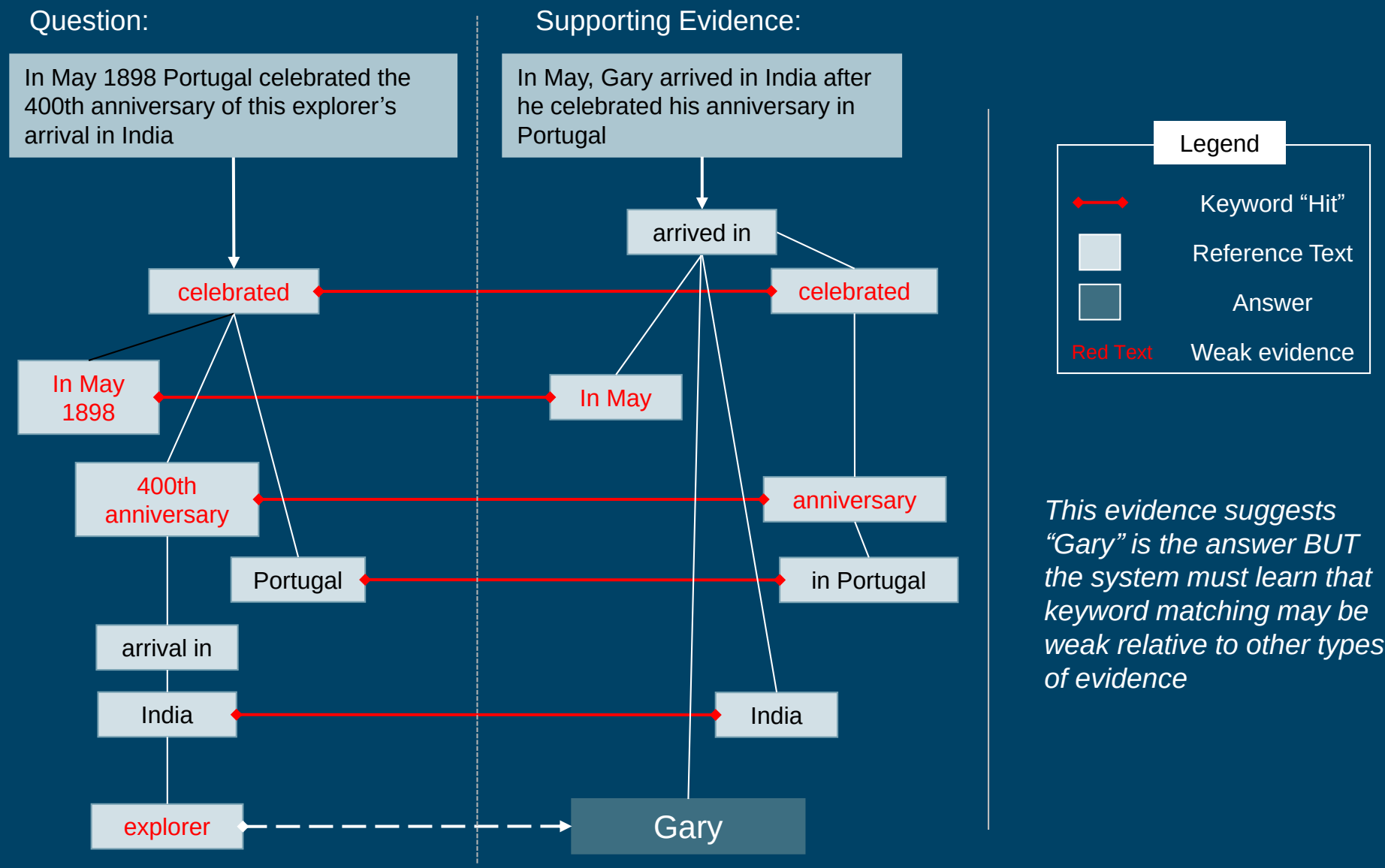
Sensing	**Reasoning**	**Learning**
Language		

Understanding language is critical to making effective use of the *rest*

- Noses that run and feet that smell?
- Ship by truck and send cargo by ship?
- How can a slim chance and a fat chance be the same, while a wise man and a wise guy are opposites?
- How can a house burn up as it burns down?
- Why do we fill in a form by filling it out?
- How does an alarm go off by going on?

	Structured Data	*Unstructured Data*
Welch ran this?	**Person** / **Organization** L. Gerstner / IBM J. Welch / GE W. Gates / Microsoft	*"If leadership is an art then surely Jack Welch has proved himself a master painter during his tenure at GE."*

Answering complex questions requires more than keyword evidence



This evidence suggests "Gary" is the answer BUT the system must learn that keyword matching may be weak relative to other types of evidence

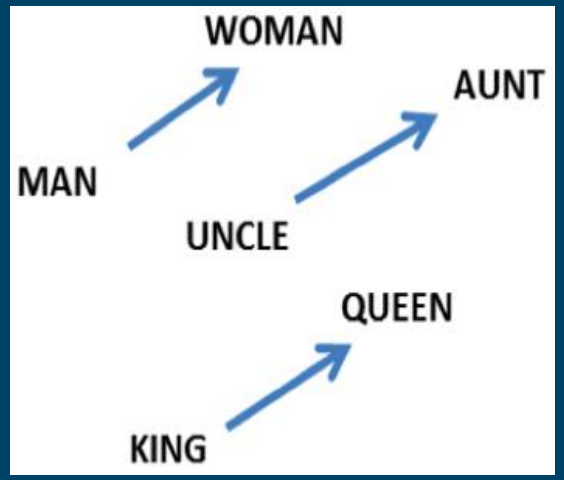
Projecting Words/Letters as vector representations

Learning vector representations of word by exploiting its *various contexts* observed in large amounts of text

- Words are no longer discrete symbols, brightly is more related to shining than other random words.
- Semantic relations appear as linear relationships in the space of learned representations

King – Queen ≈ Man – Woman

Paris – France + Italy ≈ Rome



- More robust for foreign languages (e.g, CJK).
- More robust to sparse data (and spelling errors)

Visual data provides key insights that can transform industries

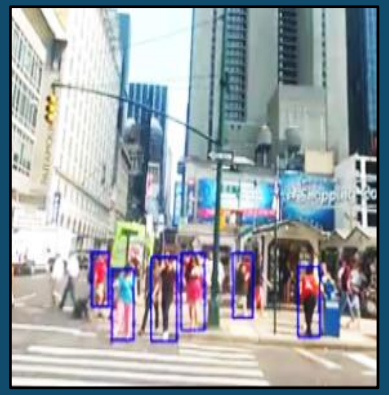
Medical Imaging

- Diagnosis?
- Clinical Features?
- Similar Cases?

Health and Wellness

- What food?
- Portion size?
- Nutrition?

Safety and Security



- Persons?
- Activities?
- Behaviors?

Retail

- Fashions?
- Products?
- Trends?

Real Estate / Insurance

- Condition?
- Style?
- Value?

Satellite

- Counts?
- Patterns?
- Environment?

Use of Semantics in Visual Processing

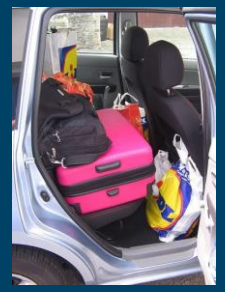
Image Captioning by Natural Language Generation

1. a close up of a cat in a pool of water -0.741016
2. a close up of a cat laying in the water -0.755236
3. a close up of a cat laying on the water -0.827383
4. a close up of a cat laying on a bed -0.872144
5. a close up of a cat laying on a water -0.899716

Image Similarity Search

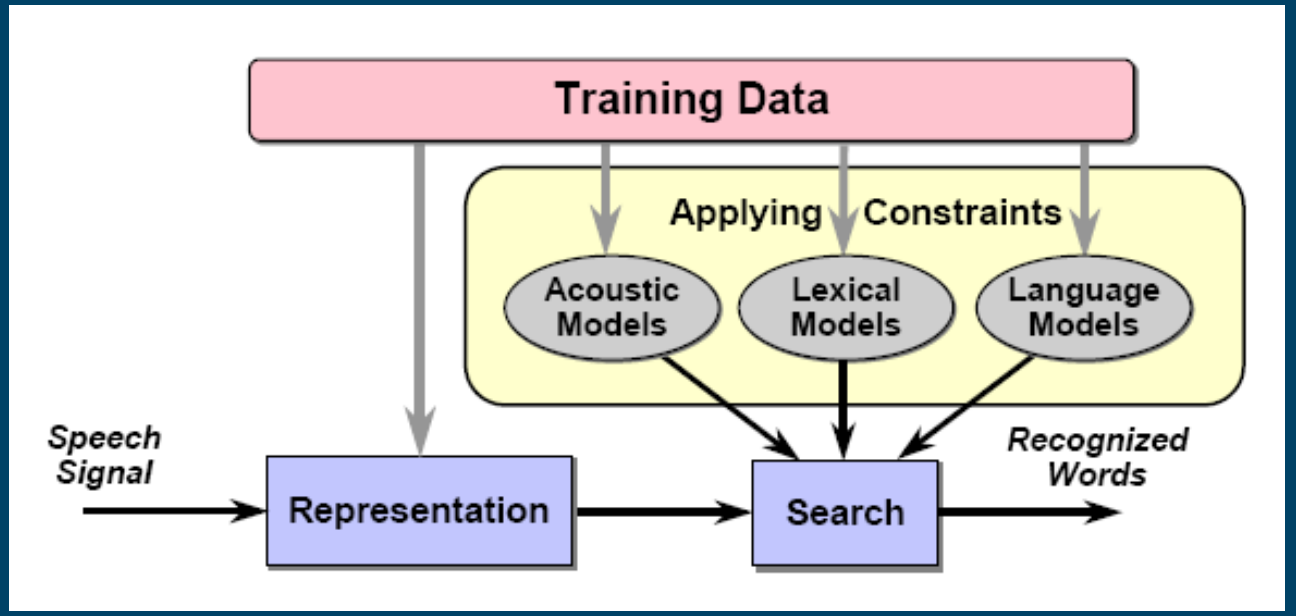
baseline

semantic



Deep Learning in Speech Recognition

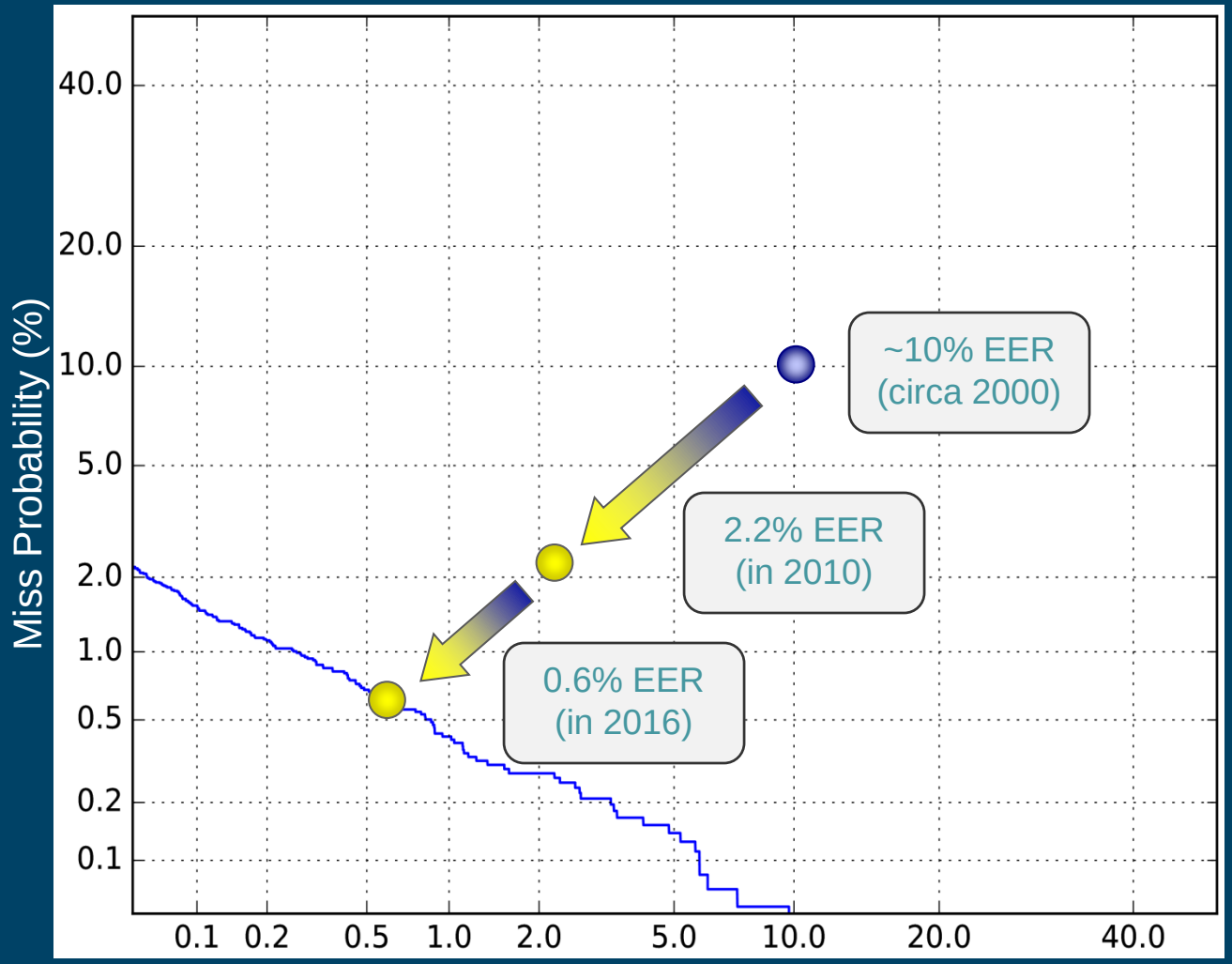
- Deep learning is being applied to more and more aspects of speech recognition with complete end to end systems purely based on deep learning.



Error Rate reduction due to Deep Learning (2011 – now)

Model (trained on SWB-300 hours)	Hub5-2000
Baseline GMM/HMM	14.5
DNN, Cross-entropy Trained (Microsoft)	14.2
DNN Sequence Trained (IBM)	12.4
CNN (IBM)	11.8
Recurrent Neural Networks-RNN (IBM)	11.3
Joint CNN/DNN (IBM)	11.2
Joint CNN/DNN + iVector features (IBM)	10.4
Joint CNN/DNN + RNN + NNLM, 2000 h (IBM)	8.0
RNN + VGG + NNLM + ModelM, 2000 h (IBM)	**6.9**

Speech Biometrics: Speaker, Gender, Age and Language



- Results on NIST 2010 SRE (Telephony)
- Most publicly contested data set
- >16X improvement in 16 years

Common Challenges

- Training Time
- Training Data
- Modeling
- Objective Function

Thought-intensive games are one benchmark of progress in Cognitive Systems

Checkers (1956)

Backgammon (1994)

Chess (1997)

Jeopardy (2011)

Thank You

David Nahamoo
nahamoo@us.ibm.com