Help Wanted: Creating a New Era of Computing

For AAAI Fall Symposium
November 2014

Frank Stein
Director of Analytics Solution Center
Washington, DC
fstein@us.ibm.com
www.ibm.com/ascdc

Jim Spohrer
Director, Global University Programs
San Jose, CA
spohrer@us.ibm.com
Cognitive-science.info
Businesses and Professionals are “dying of thirst in an ocean of data”

- 90% of the world’s data was created in the last two years
- 80% of the world’s data today is unstructured
- 1 in 2 Business leaders don’t have access to data they need
Vision:
Create a new partnership between people and computers that **augments**, **scales** and **accelerates** human expertise.
Help Wanted: Collaborators on Creating Cognitive Assistants

Introduction to the Cognitive Systems Institute

The Cognitive Systems Institute, a new set of IBM university programs launched in conjunction with IBM Research and the Watson Business Unit, focuses faculty collaborators on building and evaluating cognitive assistants for every profession. The Cognitive Systems Institute centers on professional cognitive assistants that exhibit the three Ls – language, learning, and levels to augment and scale human expertise.
Agenda

- The Start of a New Era of Computing
- IBM Watson – What we’ve been up to
- Cognitive Assistance Vision
- Audience Discussion and Input
Agenda

- The Start of a New Era of Computing
- IBM Watson – What we’ve been up to
- Cognitive Assistance Vision
- Audience Discussion and Input
IBM Research
– Innovation that matters – for our company, for the world

1945 – 1980’s
Established hardware/system labs near major universities

1990’s – Today
Expand to software and services, globalize labs, deep client engagements

2014 – Future
Leading IBM’s future growth
Watson enables insights by connecting and analyzing hundreds of internal and external data sources in minutes rather than weeks.
Annotators allow Watson to read and extract appropriate information

...doxorubicin results in extracellular signal-regulated kinase (ERK)2 activation, which in turn phosphorylates p53 on a previously uncharacterized site, Thr55...

- **ERK2** maps to domain of Post Translational Modification
- **p53** extracts entities: ERK2 = Protein, P53 = Protein, Thr55 = Amino Acid
- **Thr55** extracts entities: ERK2 = Protein, P53 = Protein, Thr55 = Amino Acid
- **phosphorylates** extracts verb: Maps to domain of Post Translational Modification, Recognizes subject / object relationships

- **on** extracts preposition: Recognizes preposition location on Thr55
Watson evaluates supporting evidence

**Ingest**
- What genes contribute to developing colon cancer?

**Learn**
- Side Effects
- Lab Notes
- Genes
- Publications
- Drugs
- Animal Models
- Clinical Trial Data

**Test**
- Quantity
- Proximity
- Relationship
- Domain Truths/Business Rules

**Experience**
What we learned in Jeopardy!

- The DeepQA approach can accurately answer single sentence queries with confidence & speed.

- Highly dependent on content, content quality, content formats

- Need a combination of technologies to get satisfactory performance: Semantic Technology, machine learning, information retrieval/search technology, databases, and high performance computing techniques.

- Both structured and unstructured content need to be combined for best results.

- Need to extend Watson to handle richer interactions and continuous training/learning
Watson is ushering in a new era of computing . . .

With the goal to create a new partnership that **enhances**, **scales** and **accelerates** human expertise.
Agenda

- The Start of a New Era of Computing
- IBM Watson – What we’ve been up to
- Cognitive Assistance Vision
- Audience Discussion and Input
Watson has evolved

2011
- 2880 cores
- Single user system
- 2-3 sentences input
- 5+ days to retrain Wikipedia

2013
- 240% faster, 1 Power 750
- 1000s of users
- 20 pages of input
- < day to ingest and train
- Medical corpus

2014
- Watson Cloud Service
- Millions of users
- Dialog chaining
- Hours going to minutes
- Broad industry corpus

Cognitive Ecosystem

Cognitive Products

Jeopardy

size speed corpus flexibility #users
Current Watson Solutions

Engagement
- IBM Watson Engagement Advisor
  - Customers
  - Transform Customer Experiences

Discovery
- IBM Watson Discovery Advisor
  - Analysts, Academics
  - Accelerate Research and Insights

Decisions
- IBM Watson Policy Advisor
  - Interactive Care Guide & Reviewer
  - Health Plans & Providers
  - Clinicians
- Interactive Care & Insights for Oncology
  - Streamline authorization of procedures
  - Improve Diagnosis and Treatment

Transform Customer Experiences
Accelerate Research and Insights
Streamline authorization of procedures
Improve Diagnosis and Treatment
Watson Engagement Advisor

Delivering the next wave of customer engagement

What it does

- Automates customer interaction to increase customer engagement in sales and service
- Transforms customer engagement by knowing, engaging and empowering clients
- Develops customer relationships through a transformative user experience

How it does it

- Provides answers not links and webpages
- Answers with evidence not guesses
- Not restricted to a predefined question-answer set
- Learns from every interaction
 Accelerate the discovery of new insights by synthesizing information in seconds

- Take advantage of massive sources of data
- Move beyond keyword search
- Find answers to questions that have not been asked yet or answered before
- Find insights into hidden relationships and dig deeper
- Generate leads to valuable insights and provide evidence to substantiate new claims
- Stay current
### Watson Discovery Advisor success stories

#### Expert
- **Baylor College of Medicine Genomics Researcher**
- **Large Pharma Infectious Disease Researcher**
- **Large Pharma Toxicologist**
- **Janssen Outcomes Researcher**

#### Challenge
- Explore New Ways to Target P53
- Find Potential in Approved Drugs to Treat Malaria
- Identify Safety Issues Faster
- Compare Treatments Across thousands of Studies

#### What Watson Did
- **Watson analyzed:**
  - thousands of articles instantly
  - hundreds of kinases
- **Watson saw:**
  - a universe of complex relationships
  - connected known MOAs to disease pathways
- **Watson extracted:**
  - insights from hundreds of toxicology reports
- **Watson will compare:**
  - many treatments quickly and simultaneously

#### Result
- **7 New Potential Targets**
- **11 Drug Candidates**
- **Safety Assessment Time Shortened**
- **Discover Key Differences Faster**
Watson Discovery Advisor for Oncologists

- **Evidence-based test and treatment suggestions**
  Drawn from 600K+ pieces of evidence and 2M pages of text from 42 publications

- **Expert training by Memorial Sloan Kettering Oncologists**
  5000+ physician and analyst training hours

- **Evolves with the fast-changing field**
  Improves over time as a learning system with usage and training

- **Full transparency into sources behind suggestions**
  Including journal articles, physicians' notes, NCCN guidelines and best practices
Watson Decision Advisor for Oncology in action

video
Reusable services form the basis for Watson cognitive solutions

The same services are used by business partners, customers, and IBM Developers.
Watson Beta services are available through Watson Developer Cloud hosted on IBM Bluemix

Watson Services for Bluemix
Rapidly prototype and build powerful cognitive apps in the cloud

SIGN UP FOR A FREE BLUEMIX TRIAL

Get started with the bold new partnership between people and computers today.

30 day trial is free and is followed by standard Bluemix rates
Watson Beta services available now with more in plan

**Available today**

- **User Modeling**
  Personality profiling to help engage users on their own terms.

- **Message Resonance**
  Communicate with people with a style and words that suits them

- **Concept Expansion**
  Maps euphemisms to more commonly understood phrases

- **Relationship Extraction**
  Intelligently finds relationships between sentences components

- **Machine Translation**
  Translate text from one language to another.

- **Question and Answer**
  Direct responses to users inquiries fueled by primary document sources

- **Visualization Rendering**
  Graphical representations of data analysis for easier understanding

- **Language Identification**
  Identifies the language in which text is written

**Coming**

- Concept Analytics
- Question Generation
- Speech Recognition
- Text to Speech
- Tradeoff Analytics
- Medical Information Extraction
- Semantic Expansion
- Policy Knowledge
- Ontology Creation
- Q&A in other languages
- Policy Evaluation
- Inference detection
- Social Resonance
- Answer Assembler
- Relationship identification
- Dialog
- Machine Translation (French)
- Smart Metadata
- Visual Recommendation
- Industry accelerators
Question and Answer Service

What is it?
Direct responses to users inquiries fueled by primary document sources

How does it work?
Interprets and answers user questions directly based on primary data sources (brochures, web pages, manuals, records, etc.) that have been selected and gathered into a body of data or ‘corpus’. The service returns candidate responses with associated confidence levels and links to supporting evidence. The current data corpora on BlueMix focuses on the Travel and Healthcare industries.

Use Cases-
Healthcare: What is a stroke? What is the cause of Wilson Disease?
Travel: Where is the best place to stay in Prague?
User Modeling

What is it?
Personality profiling to help engage users on their own terms.

How does it work?
The user Modeling service uses linguistic analytics to extract a set of personality and social traits from the way a person communicates. The service can analyze any communication the user makes available such as their text messages, tweets, posts, email, and more. Users of the service can understand, connect, and communicate with people on a more personally tailored level by analyzing personality and social traits.

Use Cases-
The service can analyze text based on a customer’s twitter stream to help a travel agency decide between leading with a budget or luxury trip offer. Anywhere improving a customer engagement can help create an organization differentiate itself.
Language Identification

What is it?
Identifies the language in which text is written

How does it work?
The Language Identification service detects the language in which text is written. This helps inform next steps such as translation, voice to text, or direct analysis. The service can be used in tandem with the Machine Translation service. Today, the service can identify 15 languages – Arabic; Chinese (Simplified); Chinese (Traditional); Cyrillic; Danish; Dutch; English; Farsi; Finnish; French; German; Greek; Hebrew; Hindi; Icelandic; Italian; Japanese; Korean; Norwegian (Bokmal); Norwegian (Nynorsk); Portuguese; Spanish; Swedish; Turkish; Urdu

Use Cases-
A building block for Machine Translation and future tech
Machine Translation

What is it?
Globalize on the fly. Translate text from one language to another.

How does it work?
The Machine Translation service converts text input in one language into a destination language for the end user. Translation is available among English, Brazilian Portuguese, Spanish and French.

Use Cases-
A French speaking help desk representative is assisting a Portuguese speaking customer through a chat session and is able to interact through the translation service.
Message Resonance

What is it?
Communicate with people with a style and words that suits them

How does it work?
The message resonance service analyzes draft content and scores how well it is likely to be received by a specific target audience. This analysis is based on content that’s been written by the target audience itself such as fans of a specific sports team or new parents. Today, analysis can be done against people active in cloud computing or discussions but future versions will let users provide their own community data.

Use Cases-
Among people active in cloud computing discussions, option A content is likely to resonate very well, option B poorly, and option C moderately well.
Concept Expansion

What is it?
Maps euphemisms or colloquial terms to more commonly understood phrases

How does it work?
The Concept Expansion service analyses text and interprets its meaning based on usage in other similar contexts. For example, it could interpret “The Big Apple” as meaning “New York City”. It can be used to create a dictionary of related words and concepts so that euphemisms, colloquialisms, or otherwise unclear phrases can be better understood and analyzed.

Use Cases-
“drugs” can be expanded to:

start seed terms → motrin, aspirin, Keflex

post expansion → allegra, lisinopril, metformin, aspirin, equagesic, cimetidine, fiorinal, vancomycin, avelox, protonix, glimepiride,
Relationship Extraction

What is it?
Intelligently finds relationships between sentences components (nouns, verbs, subjects, objects, etc.)

How does it work?
Parses sentences into their various components and detects relationships between the components. It can process new terms (like people’s names in a news feed) it has never analyzed before through contextual analysis. Sentence components include parts of speech (noun, verb, adjective, conjunction, etc.) and functions (subjects, objects, predicates, etc.). The service maps the relationships between the components so that users or analytics engines can more easily understand the meaning of individual sentences and documents.

Use Cases-
The service can analyze a news article and pull out the relevant people, organization, event type, date/time, or location of actions taken. It then can relate the components such as what adjective describes what noun and what action a subject takes upon what object.
Visualization Rendering

What is it?
Graphical representations of data analysis for easier understanding

How does it work?
The service takes input data and graphically renders it as an interactive visualization which can range from a common business chart to more advanced layouts. The visualizations can be easily modified to match user needs, visual styling, and types of data being analyzed.

Use Cases-
The service could represent neighborhood demographic data as mini pie charts showing income levels centered on geographic locations on maps, or as tree maps that can switch from looking at income by age to house size or by education level.
Watson APIs will be available on Bluemix Development Environment

- Get a Bluemix account
- Try the Watson services free of charge for 30 days
- Take the next step toward further development or production deployment

BETA versions of the Watson cognitive service are now available to registered Bluemix users through the Bluemix catalog.

LEARN MORE ABOUT THE WATSON PLATFORM
BROWSE THE CATALOG
VIEW THE DOCS
DARPA SyNAPSE: Disruptive Architecture that complements von Neumann architecture

TrueNorth Chip (2014)
- 1M Neurons
- 250M synapses
- 70mWatts
Agenda

- The Start of a New Era of Computing
- IBM Watson – What we’ve been up to
- Cognitive Assistance Vision
- Audience Discussion and Input
Create a new Partnership for Professionals:

Augmenting

Scaling

Accelerating

Human Expertise and Knowledge with a cognitive assistant
Extending human cognition with Cognitive Assistants

**Discovery**
- Create new insights and new value

**Decision**
- Provide bias-free advice semi-autonomously

**Understanding**
- Build and reason about models of the world, of the user, and of the system itself

**Question Answering**
- Leverage encyclopedic domain knowledge in context
Key technologies for a new era of computing

Context and Learning

Visual Analytics and UI

Data-Centric Systems

Cognitive Architectures

Atomic and Nano-scale

© 2013 International Business Machines Corporation
Future Learning Systems

- Training and Learning Engines
  To Build Models and Define Insight
- Policy Engine
  Business, Legal and Ethical Rules
- Hypothesis Engines
  To Understand and Plan Actions
- Outcome Engine
  Actuation and Validation

- Active Learning
- Verification Engines
  (e.g. Simulations)

(Natural Interfaces)
IBM Cognitive Assistant Efforts

- Cognitive Assistance for various professions: (in progress)
  - Oncology Expert Advisor
  - Watson Electronic Medical Records (EMR) Assistant
  - On-line shopping Advisor
  - Chef Watson
  - New Hire Chatbot
  - IBM Sales Assistant

- Cognitive Work Assistant in IBM Research
  - Focused on helping with office and work task, processing information surrounding a worker in the work environment, and offering intelligent suggestions and proactively/reactively acting on behalf of the person it is representing
  - Using NLP, knowledge management, machine learning and cognitive computing techniques
Watson @Bon Appetit – Chef Watson
Concept: Start with Task Definitions for Standard Occupations
(courtesy of O*net Job Library)

Summary Report for:
17-2199.01 - Biochemical Engineers

Develop usable, tangible products, using knowledge of biology, chemistry, or engineering. Solve problems related to materials, systems, or processes that interact with humans, plants, animals, microorganisms, or biological materials.

Sample of reported job titles: Engineering Director, Process Engineer

Tasks

- Devise scalable recovery, purification, or fermentation processes for producing proteins or other biological substances for human or animal therapeutic use, food production or processing, biofuels, or effluent treatment.
- Read current scientific or trade literature to stay abreast of scientific, industrial, or technological advances.
- Design or conduct studies to determine optimal conditions for cell growth, protein production, or protein or virus expression or recovery, using chromatography, separation, or filtration equipment, such as centrifuges or bioreactors.
- Develop biocatalytic processes to convert biomass to fuels or fine chemicals, using enzymes of bacteria, yeast, or other microorganisms.
- Prepare technical reports, data summary documents, or research articles for scientific publication, regulatory submissions, or patent applications.
- Confer with research and biomanufacturing personnel to ensure the compatibility of design and production.
- Design or direct bench or pilot production experiments to determine the scale of production methods that optimize product yield and minimize production costs.
- Develop methodologies for transferring procedures or biological processes from laboratories to commercial-scale manufacturing production.
- Design or conduct follow-up experimentation, based on generated data, to meet established process objectives.
- Maintain databases of experiment characteristics or results.
Building a Society of Cognitive Agents

Systems of cognitive agents that collaborate effectively with one another

Cognitive agents that collaborate effectively with people through natural user interfaces

A nucleus from which an internet-scale cognitive computing cloud can be built
Cognitive Systems Institute

- **Vision:** Augment and scale human expertise
  - Website (cognitive-science.info)
  - LinkedIn Group

- **Platforms:** Cognition as a Service
  - BlueMix & SoftLayer & CCAMSS
  - DEEPQA Semantic Technologies
  - Watson Developer Cloud
  - Watson Platform Next (IBM Research)
  - Corelet Programming & TrueNorth

- **Researchers in Residence**
  - Grand Challenges
  - Co-Create Grant Proposals
  - Publications, Guest Lectures
  - Recruiting Interns, Co-ops, etc.
  - Conferences & Cognitive Colloquia
Agenda

- The Start of a New Era of Computing
- IBM Watson – What we’ve been up to
- Cognitive Assistance Vision
- Audience Discussion and Input
Help Wanted Areas

- User Interface
- Learning and Training
- Task and Cognitive Modeling
- Test and Evaluation
- System Issues
User Interface

1. How does the User Train the Cog?
2. How does the User Delegate specific tasks to the Cog?
3. User interface issues related to acceptance & trust
   - Can I trust that the Cog did what I told/taught/think the Cog did?
4. On what conditions should the Cog ask for help from User?
5. Modeling User behavior and preferred modes of interaction
6. Cog understanding Human biases and how to counter?
7. How can Cog communicate results to achieve User trust – explaining & showing statistical correlations, providing assumptions and evidence (what if Cog uses techniques it learned that are beyond what Users do?)
8. Cog understanding of figures, graphs, lab notes, pictures, videos, xrays…
Learning and Training

1. How do we train the Cog so that it is reliable?
2. Does the User train the Cog (personal assistant) vs. Enterprise (Organization trains with best practices)?
3. Training by demonstration
4. Training by ingesting documentation, training manuals, text books,
5. Interactive Learning – Cog asks questions of Trainer/User
6. Teaching the Cog Goals – Cog improves over time towards defined goals
7. Adaptation to environmental changes, new data sources, etc
8. Can Cog understand what is Routine and how to automate without User training?
9. Can the Cog train new Users? Other Cogs?
Task and Cognitive Modeling

1. Task & Subtask composition and reuse
   - What data sources to use for each task
   - Next steps based on results
2. Initiating actions based on incoming data
3. Modeling the world of the user being served, including all context around the work/task, and being able to use the contextual and environmental awareness to proactively and reactively act on behalf of the user
4. User definition of persistent tasks (e.g., persistent search of incoming information)
5. Information Restrictions – Could Cog access PII (HIPPA) data that User couldn’t? Does Cog have same access authority as User for classified data?
Testing & Evaluation

1. Testing the value provided to the User
   - How to do controlled experiments?
   - Is it possible to simulate the complexity in a lab environment?
2. How do we test the reliability of the Cog and the combined H-C system?
3. Can we develop some Challenge problems to test the User/Cog system?
   - How to judge results? Answer or Approach? Use of right data?
4. Metrics for a Cog/User system:
   - User Adoption, Speed to result, Accuracy, Precision, Stability, Robustness, Finding Results the User didn’t find alone, Transparency, resource usage
System Issues

1. How to build cooperative cogs working as a team with teams of Users?
2. Symbiosis Questions – What is best for the Cog to do and what is best for the User to do? How to assign appropriately?
3. Methodology – How to build an iterative approach of build/test to quickly evolve?
So what should universities be asking themselves?

**By 2017, 10 percent of computers will be learning rather than processing.**

- Will your researchers, faculty, students be benefitting from cognition as a service?
“The hope is that, in not too many years, human brains and computing machines will be coupled together very tightly, and that the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today”

*Man-Computer Symbiosis*, J. C. R. Licklider IRE Transactions on Human Factors in Electronics, volume HFE-1, pages 4-11, March 1960
Thank You