RUTGERS

RUTGERS
THE STATE UNIVERSITY
OF NEW JERSEY

Artificial Intelligence: Déjà vu all over again

Miklos A. Vasarhelyi
KPMG Distinguished Professor of AIS

November 8, 2019
47 WCAS

The Singularity in artificial intelligence (AI) is the **inflection point where machines advance beyond human intelligence and thus become self sufficient**. Jon von Neumann first defined the term in the 1950s and it has been further advanced by well-known futurist and technologist Ray Kurzweil.

AI AND THE SINGULARITY

Outline

- The CarLab
- What is intelligence
- Evolution of Artificial Intelligence
- Cognitive computing
- Exogenous Data,
- IPA
- Estimates with machine learning

THE CARLAB


RUTGERS
250
Rutgers
Business
School
EXIT

BRIGHAM YOUNG UNIVERSITY			
The Ranking of Rutgers in the Accounting Areas			
Areas	Ranking 2008-2013	Ranking 2002-2013	Ranking 1990-2013
AIS	#1 out of 179	#1 out of 207	#1 out of 241
Audit	#10 out of 320	#7 out of 370	#11 out of 438
Financial	#70 out of 356	#89 out of 406	#83 out of 470
Managerial	#120 out of 286	#80 out of 346	#66 out of 413
Tax	#53 out of 129	#76 out of 178	#79 out of 246
Other	#35 out of 171	#18 out of 248	#25 out of 341

CarLab Analytic Research

Choosing apps	Predictive Analytics with Weather data	Audit data analytics and EDA	Envisaging the future of audit and Big Data	Text Mining	Monitoring Unibanco's branches
Visualization	Process Mining at Gamma Bank	Expert System for P-Card	Logit regression for control risk assessment	Exceptional Exceptions	Client Retention Project
Litigation prediction	Fraud Risk Assessment using EDA	Detecting duplicate records	Continuity equations	Predictive Audit	Credit card Default prediction
Insurance Analytics	Multidimensional clustering for fraud detection	Rule-based selection for transitory accounts	Continuity Equations at HCA	XBRL	Insurance Analytics
Cognitive Decision Aids	AI: Deep Learning	Robotic Process Automation (RPA)	Intelligent Process Automation (IPA)	Blockchain and Smart contracts	Cluster Analysis of US States

Recent research of PhD students

Name	Title
Abdulrahman Alrefai	Formalization of Internal Control Assessment: A Process Mining Application
Ahmad AlQassar	Resisting Change in the Audit Profession: Two Case Studies from Multi-National Firms
Andrea Rozario	Examination of Audit Planning Risk Assessments Using Verbal Protocol Analysis: An Exploratory Study
Cheng Yin	Privacy-Preserving Information Sharing within an Audit Firm
Deniz Appelbaum	Using Drones in Internal and External Audits: An Exploratory Framework
Feiqi Huang	Audit Evidence Index Project
He Li	Are External Auditors Concerned about Cyber Incidents? Evidence from Audit Fees
Jiahua Zhou	The Survived Companies with Going Concern Are Really Different from Those Bankrupted
Jun Dai	Towards Blockchain-based Accounting and Assurance
Jun Dai	Imagineering Audit 4.0
Zhaokai Yan	Impact of Data Analytics on Managerial Accounting Using Balanced Scorecard Framework
Yunsen Wang	An Application of Blockchain Technology to Fraud Detection
Yue Liu	Risk Analysis Based on 10-K Item 1a
Ting Sun	The Performance of Sentiment Features of 10-K MD&As for Financial Misstatement Prediction
Tiffany Chiu	Apply Process Mining to Evaluate Internal Control Effectiveness Automatically
Qiao Li	Rule-Based Decision Support System for Audit Planning and Audit Risk Assessment
Lu Zhang	Interactive Data Visualization for Error and Fraud Detection: Case Studies and Practice Implications

https://www.youtube.com/playlist?list=PLauepKFT6DK9vKn7-eKxzmxBegpe8v8xw

Content

- **Undergraduate, Graduate, PhD, & Audit Analytics Content**

Undergraduate	Graduate	PhD	Audit Analytics Certificate
• Introduction to Financial Accounting • Introduction to Managerial Accounting • Intermediate Accounting I • Intermediate Accounting II • Advanced Accounting • Auditing Principles • Management and Cost Accounting • Accounting Information Systems • Business Law I • Business Law II • Federal Taxation I • Accounting in the Digital Era • Computer Augmented Accounting • Decoding of Corporate Financial Communications	• Accounting Principles and Practices • Information Technology • Government and Not-for-Profit Accounting • Advanced Auditing and Information Systems • Advanced Accounting • Corporate Taxation • Income Taxation • Income Tax Estate and Trust	• Special Topics in Accounting • Survey of Accounting Information Systems • Current Topics in Auditing • Machine Learning	• Introduction to Audit Analytics • Special Topics in Audit Analytics • Information Risk Management • Tutorials for Risk Management

Rutgers Web | Videos | Playlists | Channels | Discussion | About

Special Topics in Audit Analytics

by Rutgers Web • 26 videos • 145 views • Last updated on Jun 5, 2015

Play all | Share | Save

1 Special Topics in Audit Analytics: Week 1- (Lecture 2: Analytics Big Data Audit Automation)
by Rutgers Web

2 Special Topics in Audit Analytics: Week 1-(Lecture 3: The Audit Ecosystem)
by Rutgers Web

3 Special Topics in Audit Analytics: Week 1-(Lecture 4: Audit Data Standard)
by Rutgers Web

4 Special Topics in Audit Analytics: Week 2-(Lecture 2)
by Rutgers Web

5 Special Topics in Audit Analytics: Week 3-(Lecture 1- Hypothesis Testing)
by Rutgers Web

6 Special Topics in Audit Analytics: Week 3-(Lecture 2- Hypothesis testing 2)
by Rutgers Web

7 Special Topics in Audit Analytics: Week 3-(Lecture 3 : Confidence interval)
by Rutgers Web

8 Special Topics in Audit Analytics: Week 3-(Lecture 4 -Two sample test)
by Rutgers Web

9 Special Topics in Audit Analytics: Week 3-(Lecture 5: two dependent sample test)
by Rutgers Web

10 Special Topics in Audit Analytics: Week 3-(Lecture 6: Introduce R)
by Rutgers Web

WHAT IS INTELLIGENCE?

What is intelligence?

- Is it being very cultured?
  - Being a renaissance person?
- Is it being able to solve problems no one solved before?
- Is it being Human?
- Is it contingent on the moment in time?
- Is it satisfying functionalities?

The Turing Test

Alan Turing: “Can Machines Think?”

Interrogator

Human Being | **Computer**

Written Communication

Loebner prize established in 1990

EVOLUTION OF AI

Evolution of AI

- Dreyfus (1964) classifies traditional Artificial Intelligence (**AI**)work into four main areas:
  - game playing,
  - problem solving,
  - language translation,
  - and pattern recognition.

When?

- In 1964, when Mr. McCarthy established the Stanford Artificial Intelligence Laboratory, the researchers informed their Pentagon backers that the construction of an artificially intelligent machine would take about a decade. Two decades later, in 1984, that original optimism hit a rough patch, leading to the collapse of a crop of A.I. start-up companies in Silicon Valley, a time known as “the A.I. winter.”
- Such reversals have led the veteran Silicon Valley technology forecaster Paul Saffo to proclaim: “never mistake a clear view for a short distance.”

Evolution of AI

- In the early sixties Feigenbaum reoriented the work in AI by focusing not on basic paradigms and pure logical development but on the identification and formalization of human expertise often represented in the form of software systems. This led to the area of Expert Systems (**ES**) which became one of the five main areas of AI.

Rutgers

Areas of AI 1990’S

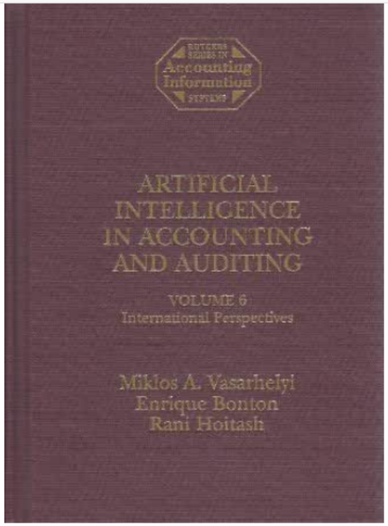
- –Natural Languages,
- –Expert Systems,
- –Cognition and Learning,
- –Computer Vision and
- –Automatic Deduction.

RUTGERS

Evolution of AI

- Expert Systems became the most popular area of AI and eventually the basis of many commercial, semi-commercial and prototype systems.

- Vasarhelyi, M. A. “**Expert Systems in Accounting and Auditing,”** in Artificial Intelligence in Accounting and Auditing, Vols. 1 to 6 , Markus Wiener Publishing Inc., New York. (1989 to 2002) om/s?k=vasarhelyi+artificial+intelligence&i=strip 140% Search



Perspectives (Rutgers Series in Accounting Information...

by Miklos A. Vasarhelyi | Jul 1, 2005

Hardcover

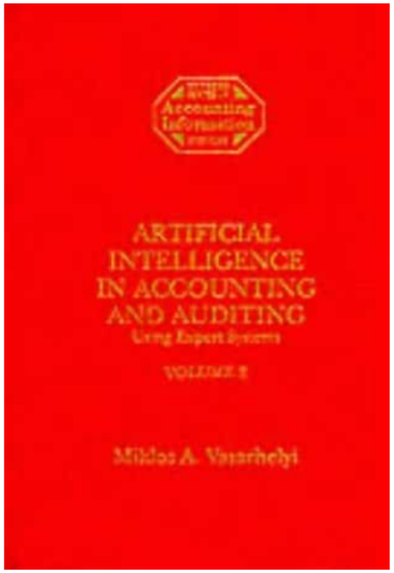
$76.51 ~~$129.95~~

prime FREE Delivery **Mon, Nov 11**

Only 1 left in stock - order soon.

More Buying Choices

$20.00 (6 used & new offers)



Artificial Intelligence in Accounting and Auditing: Using Expert Systems (Rutgers Series in Accounting Information...

by Miklos A. Vasarhelyi | Mar 1, 1996

Hardcover

$89.95

prime FREE Delivery **Mon, Nov 11**

Only 1 left in stock - order soon.

More Buying Choices

$85.93 (3 used & new offers)

Expert Systems Literature

- Shpilberg, D., Graham, L. E., & Schatz, H. (1986). ExperTAXsm: an expert system for corporate tax planing. *Expert Systems*, *3*(3), 136-151.
- Shpilberg, D., and Graham, L. E. (1986). Developing Expertaxsm-An Expert System for Corporate-Tax Accrual and Planning. *Auditing: A Journal of Practice and Theory*, *6*(1), 75-94.
- Graham, L. E., Damens, J., and Van Ness (1991). Developing risk Advisor: An expert System for Risk Identification. *Auditing: a Journal of Practice and Theory*, *10*(1), 69-96.
- Feigenbaum, E. A. (1981). Expert systems in the 1980s. *State of the art report on machine intelligence. Maidenhead: Pergamon-Infotech*.
- Buchanan, B. G., & Duda, R. O. (1983). Principles of rule-based expert systems. In *Advances in computers* (Vol. 22, pp. 163-216). Elsevier.
- Gray, G. L., Chiu, V., Liu, Q., & Li, P. (2014). The expert systems life cycle in AIS research: What does it mean for future AIS research?. *International Journal of Accounting Information Systems*, *15*(4), 423-451.

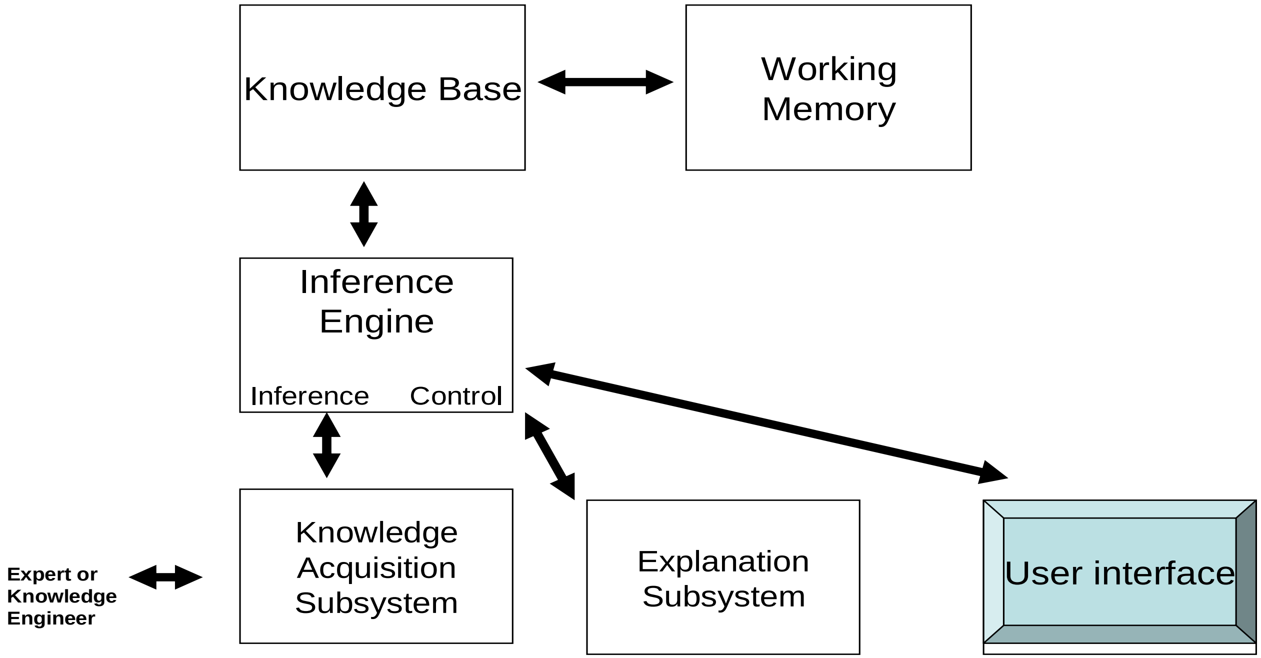
Hype or disapointment?

- Although AI has been through cycles of hype and disappointment before, big technology companies have recently been scrambling to hire experts in the field, in the hope of building machines that can learn even more sophisticated tasks. (Economist, 2014)

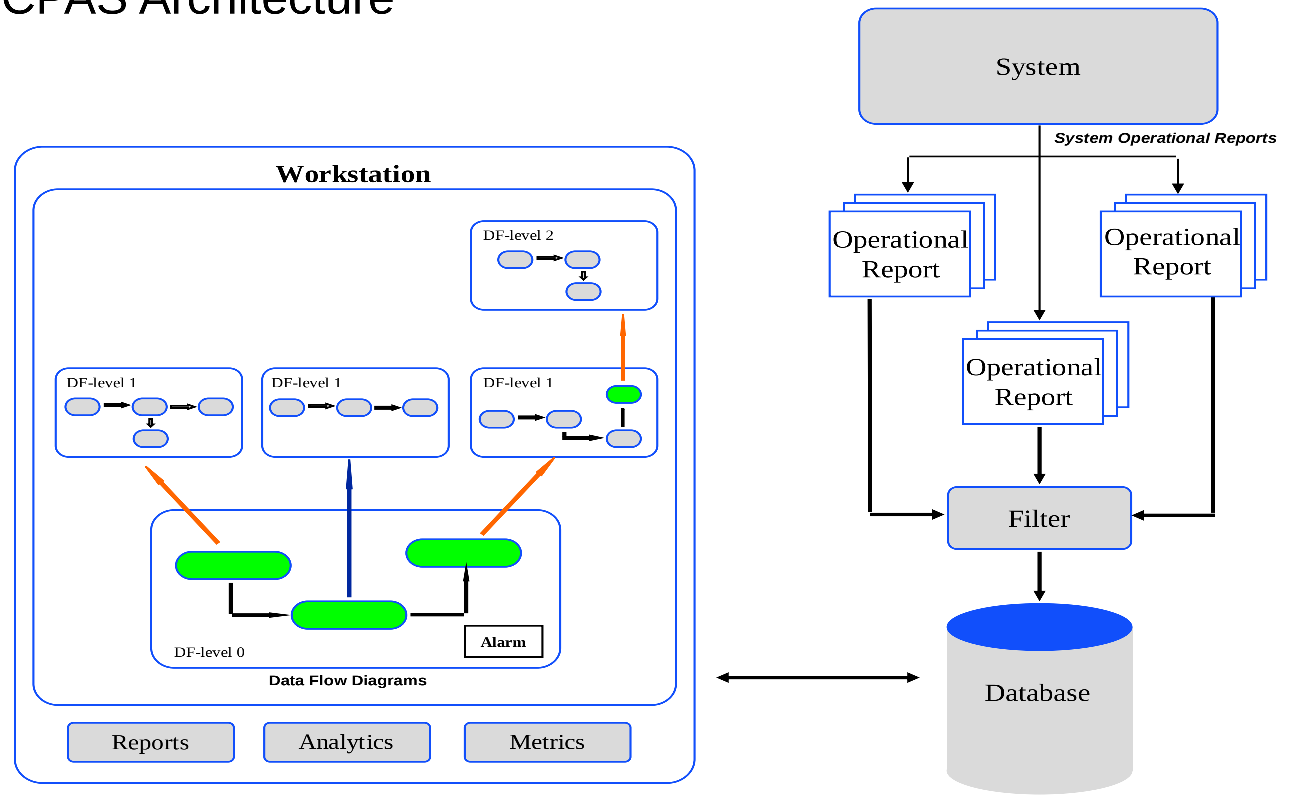
RUTGERS

IN ACCOUNTING AND AUDITING

Expert Systems



CPAS Architecture



Areas of AI 1990'S

- Natural Languages,
- Expert Systems,
- Cognition and Learning,
- Computer Vision and
- Automatic Deduction.

- And Now
  - Deep Learning / Cognitive Computing

RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

Exogenous data analytics for Auditing

Miklos A. Vasarhelyi
Helen Brown Liburd

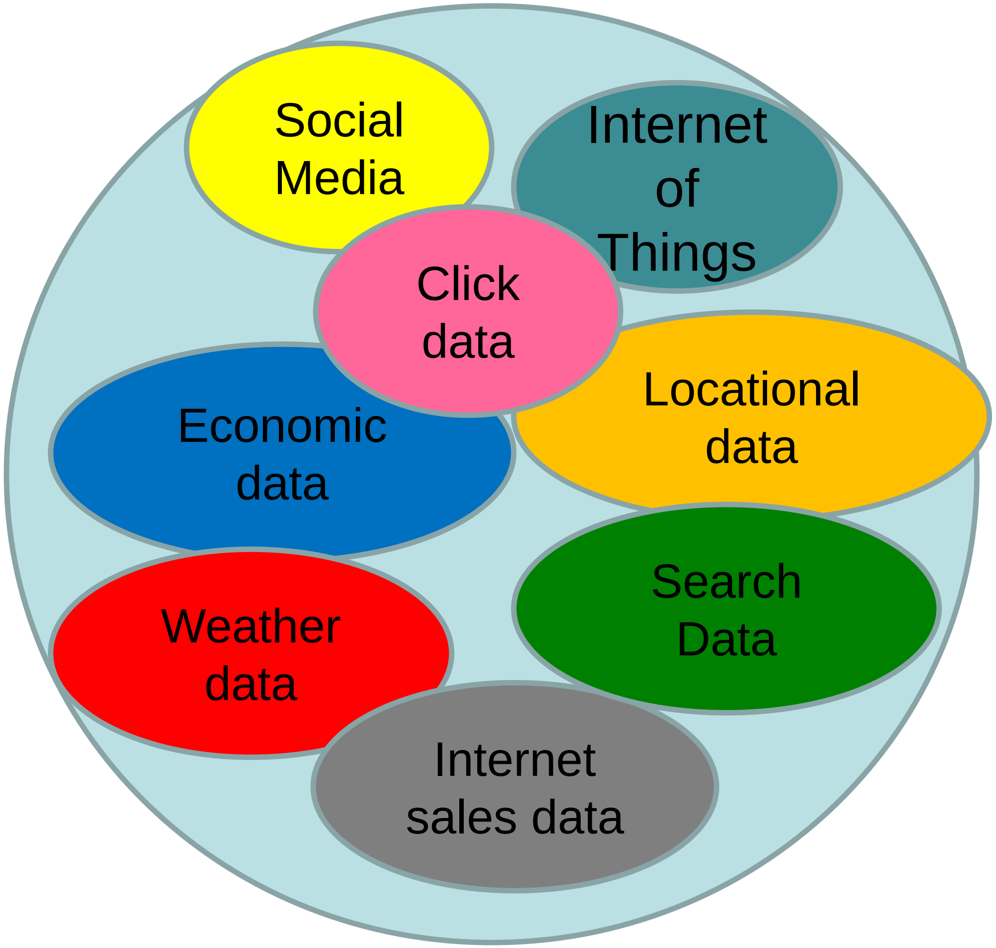
Rutgers Business School

RUTGERS

Some sources

- Amazon sales
- Google searches
- Apps used
- Calls made
- GPS or JEEP location
- Sites accessed
- Car license plates photographed
- Pictures of parking lots
- Face recognition pictures
- Site clickpaths

Exogenous Data



ED may be of easier access

ED is likely less tamperable

ED relationships will be stochastic

ED is a form of confirmation

ED may complement many current procedures

ED may create many new procedures

Some other sources

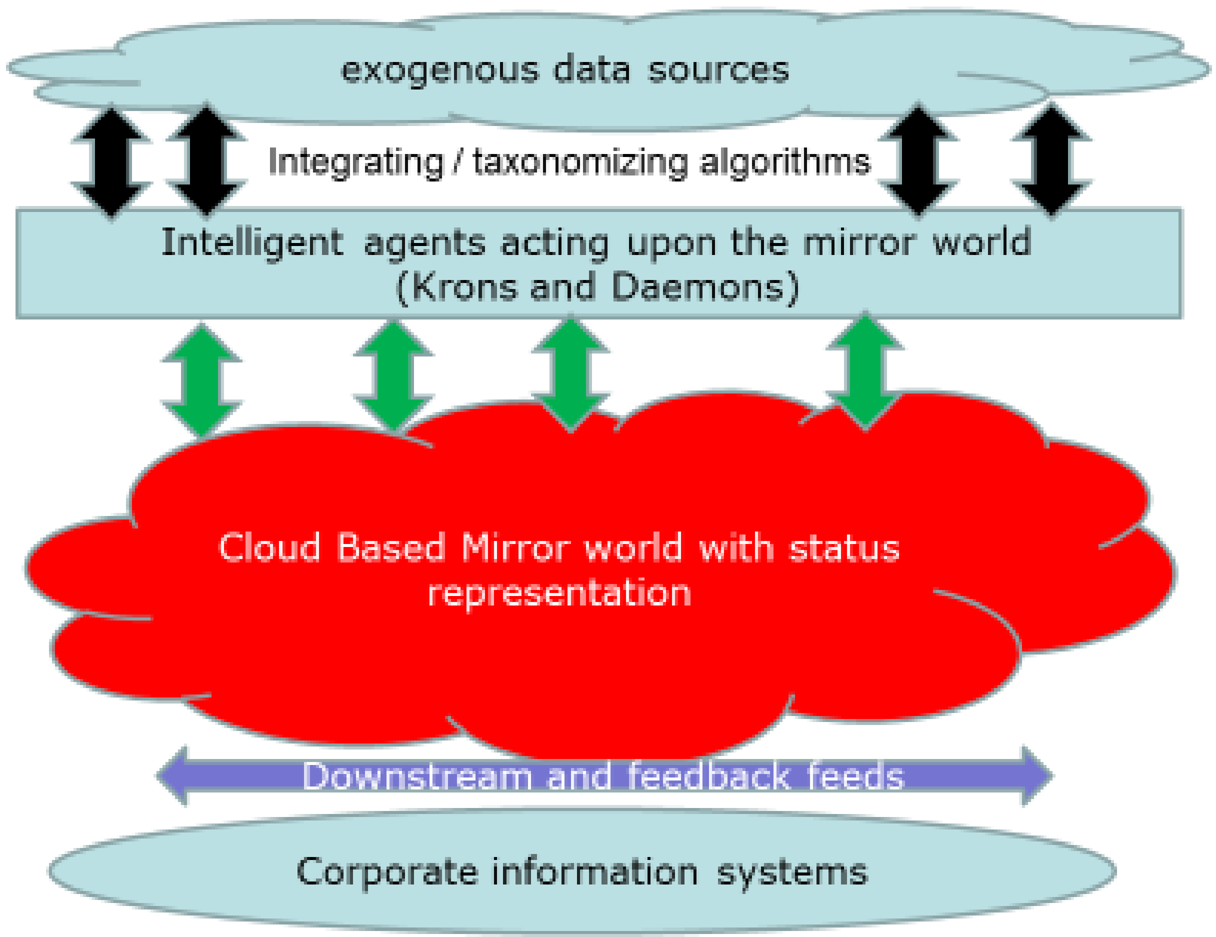
- Security recordings of arrivals and departures of trucks from parking lots for assuring inventory changes
- Telephone records, associated with e-mails, to validate sales, ordering, and discrepancy determinations
- Examination of video streams in network TV to confirm that ads were actually placed. These can be linked to variations in order/ sales to validate the ad efficiency promised by ad agencies and marketing strategies

RUTGERS

RADAR: external data that were mentioned in firm interviews

- Bloomberg data, Twitter (for estimating warranty liability)
- Economic indicators
- Capital IQ
- Credit Ratings

The new data ecosystem: Cho, Vasarhelyi & Zhang



RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Developing A Cognitive Assistant For Audit Plan Brainstorming Sessions

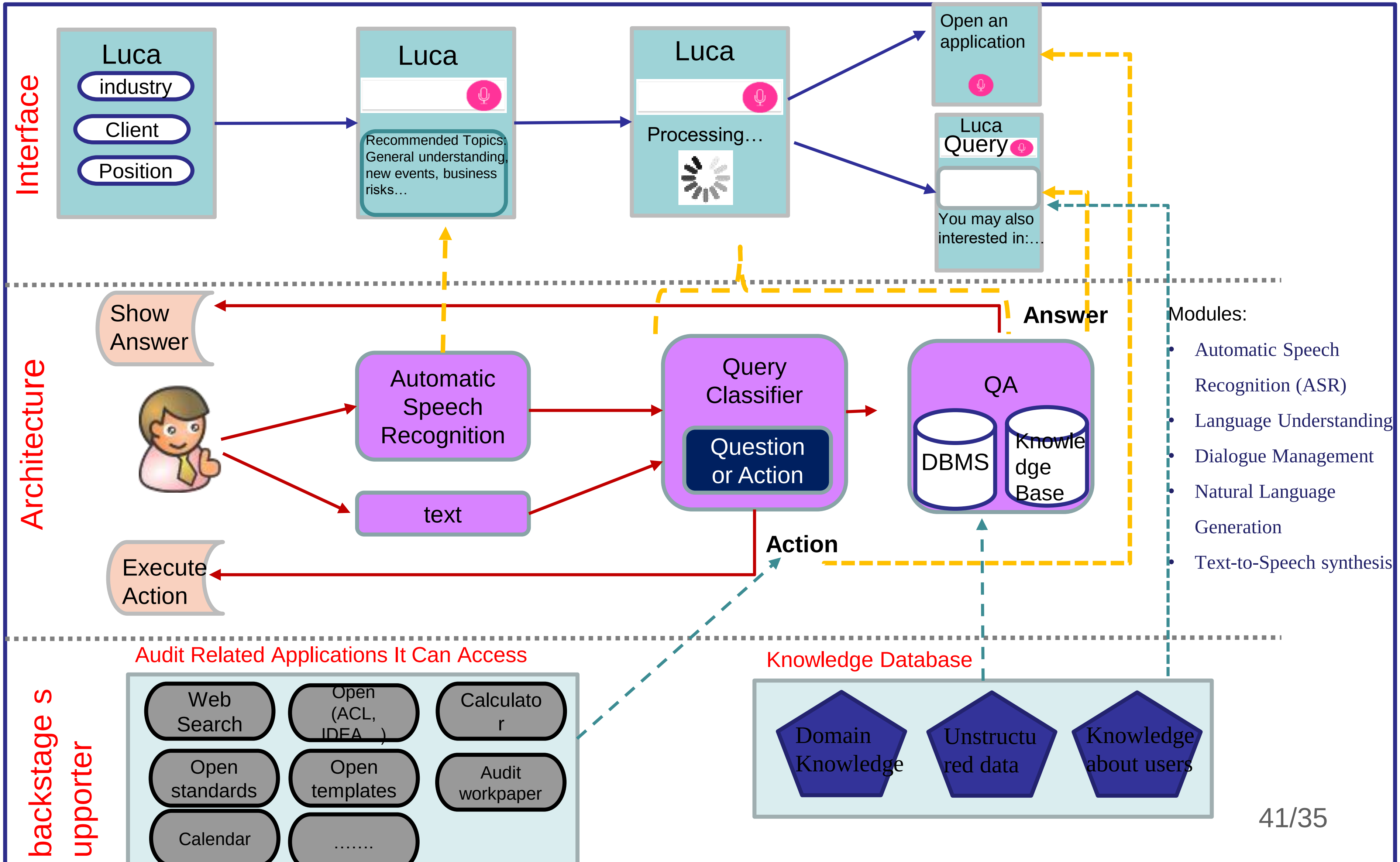
Qiao Li

Rutgers Business School

THE STATE UNIVERSITY OF NEW JERSEY

RUTGERS

Architecture of the Proposed Audit Cognitive Assistant



Jimmy Chin Skiing down the Everest

RUTGERS
THE STATE UNIVERSITY
OF NEW JERSEY

Apply Deep Learning to Analyze Big Data for Predictive Auditing

Ting Sun & Miklos Vasarhelyi

Motivation

- **AlphaGo beats European Go Champion.**
- **Deep Learning**

1. Using a vast collection of Go moves from expert players(about 30 million moves in total)→trained their system to play Go on its own→ as good as the best humans

2. (self-reinforcing)matched the system against itself → generate a new collection of moves →train a new AI player that could be better than human

How could deep learning be used in audit?

Why deep learning?

- The future of Big data is deep learning

the biggest part of Big Data is the unstructured part, and it contains valuable information and learnable patterns

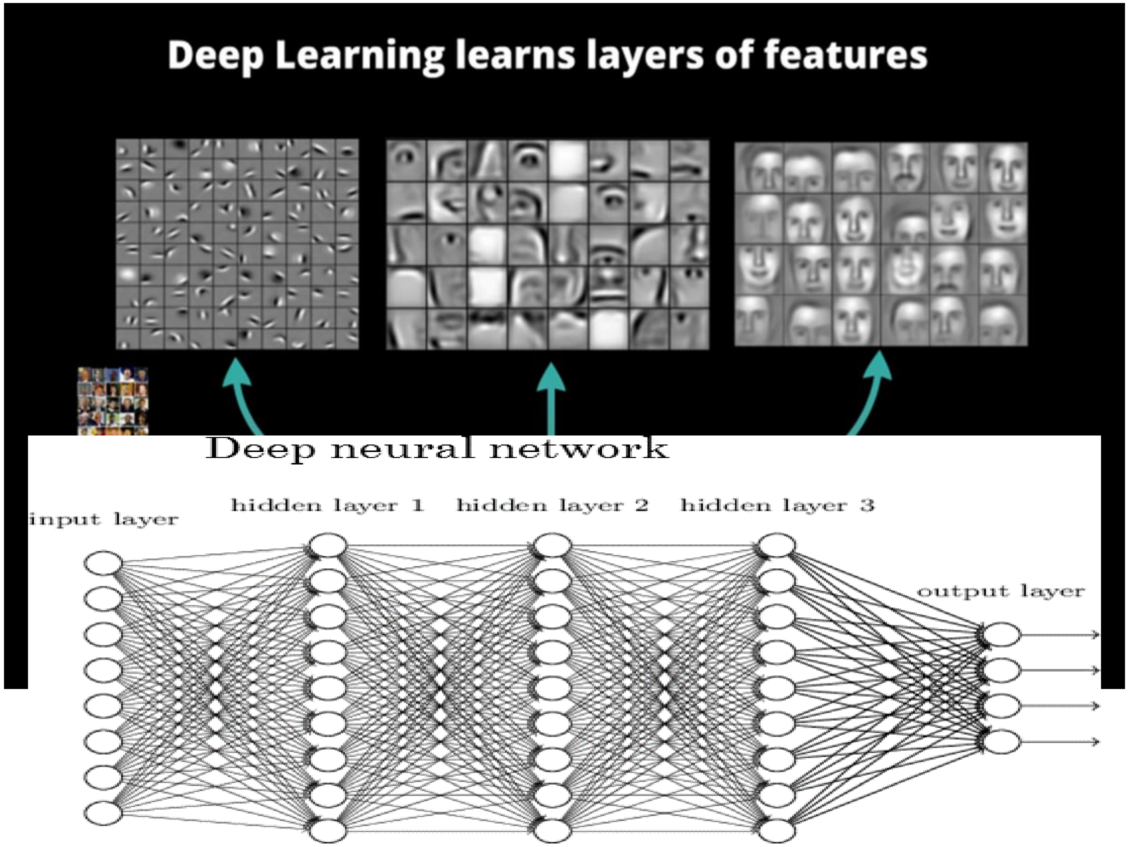
impossible and costly for human to extract features from unstructured data or label the structured data (whether there is a fraud or not)

- Traditional way (i.e., SVM, LR) of applying machine learning to predictive audit: need human to identify and detect features(attributes) of data and label the data
- Deep learning:

The computer learns the inner structure and features of data itself

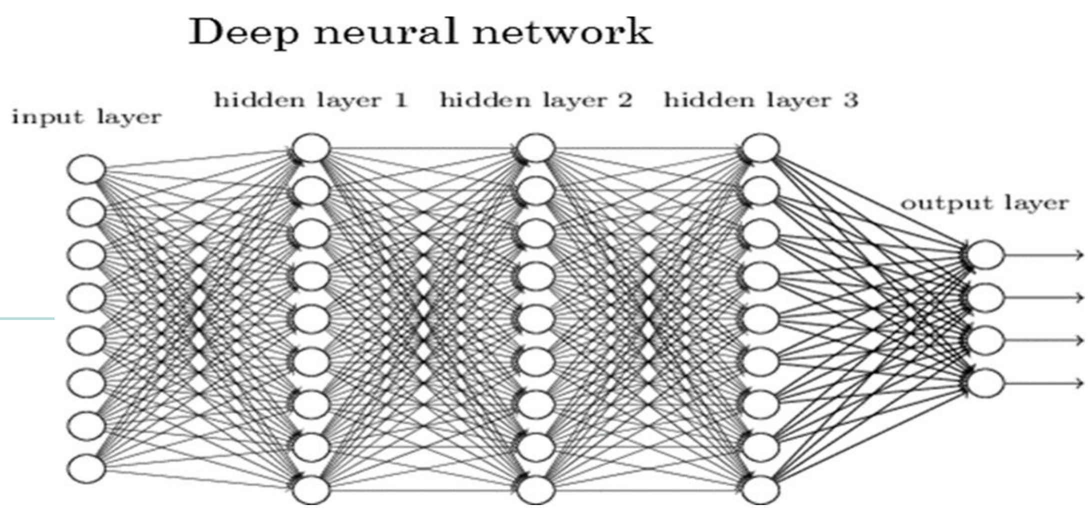
the technology could mimic the human intuition and think like human brain and automatically developing new ways of representing the data （often based on ANN）

An illustration: face recognition



Apply deep learning to audit

Raw data: Mixture of structured and unstructured data



Normal behavior

Irregularity against assertion A

Irregularity against assertion B

Irregularity against assertion C

How to apply

- For a given assertion, auditor's objective is to detect irregularities for this assertion:
- train huge volume of data (past data), including regular numerical data, semi-structured data, and unstructured data (video, audio, text)→Machine extracts features → find the patterns → generate the model(classifier)→auditors use the model to predict irregularities
- Reinforced learning: self-correction

Example: financial audit

- Step 1: train the data (historical records, video, audio, text)
- Step 2: get the output : p($fraud_i$| $State_k$)
- Step 3: calculate possible $loss_i$
- Step 4: rank risk level based on p ($fraud_i$| $State_k$) × $loss_i$
- Step 5: take audit actions for top leveled risks (audit recommendation system)
- For the long term: reinforced learning → self-correction as the state changes/new data are included

RUTGERS

How to apply ?

1. train huge volume of semi-structured or/and unstructured raw data (e.g., video, audio, text data)
2. Machine uses Deep Learning to extract features from the raw data (within the black box)
3. find the patterns (characteristic types)
4. generate model (classifier) A
5. auditors use model A to identify characteristic types from big data
6. auditors combine information identified from last step with regular structured data (like financial data) as audit evidence
7. The past audit evidence can be used as training data to develop model B (use supervised shallow learning) and use model B to predict frauds
8. As new data is collected, machine uses Reinforced Learning technique to improve the accuracy of model A (a continuous self-correction process)

Example: The Securities and Exchange Commission (SEC) comment letter

- Step 1: collect and train the raw data, SEC comment letters after reviewing registrants' filings
- Step 2: develop model A to extract features from data: words->phrases->sentences->paragraphs->meaning of entire letter (in this step, machine learns to understand the letter on its own)
- Step 3: use model A to classify different characteristic types of letters based on the meaning the machine understood in step 2

RUTGERS
THE STATE UNIVERSITY
OF NEW JERSEY

IPA

Intelligent Process Automation (IPA)

IPA is “an emerging set of new technologies that combines fundamental process redesign with RPA and machine learning” (McKinsey, 2017).

“It is a suite of business-process improvements and next-generation tools that assists the knowledge worker by removing repetitive, replicable, and routine tasks.” (McKinsey, 2017).

IPA

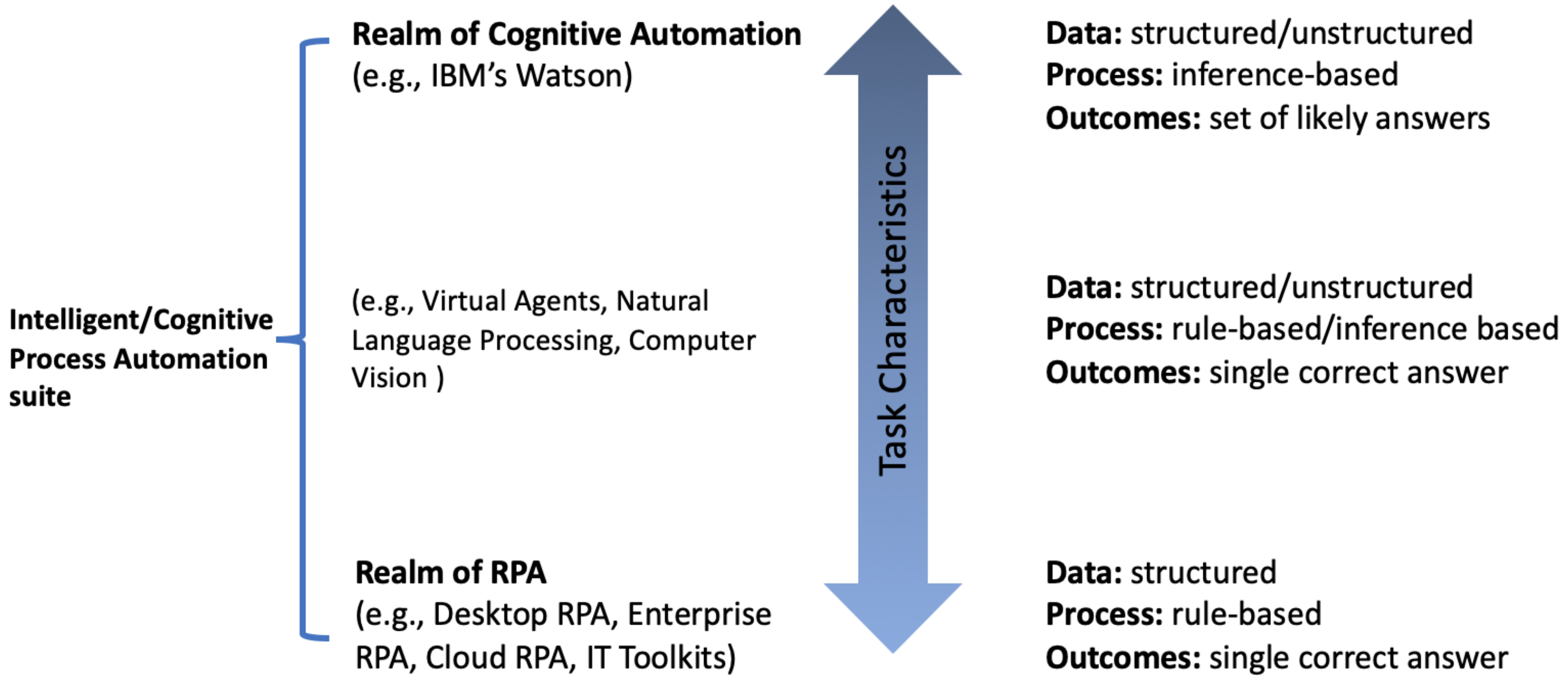
RPA

Artificial intelligence

Other technologies

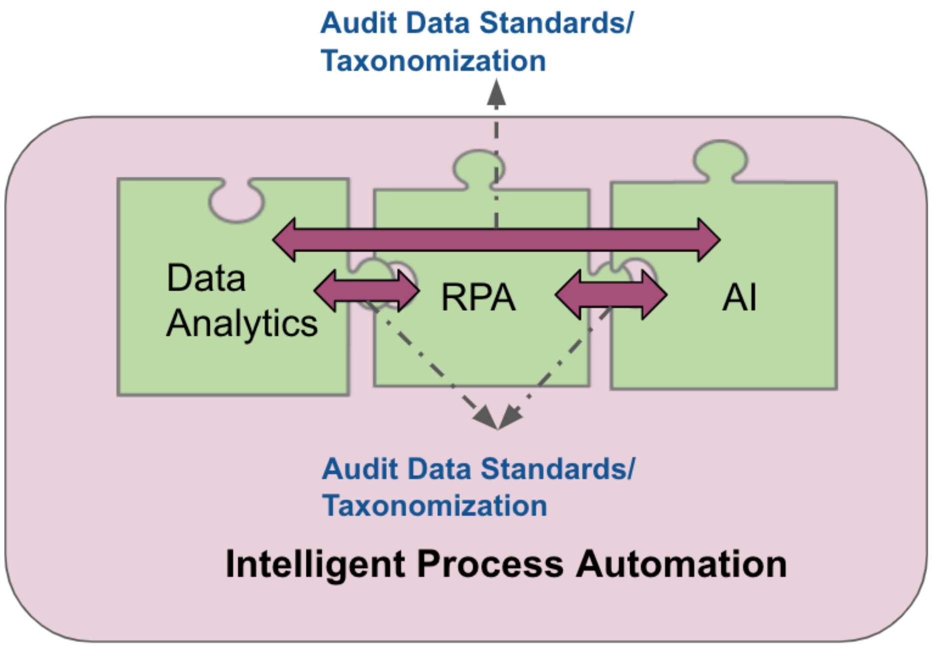
Intelligent Process Automation (IPA)

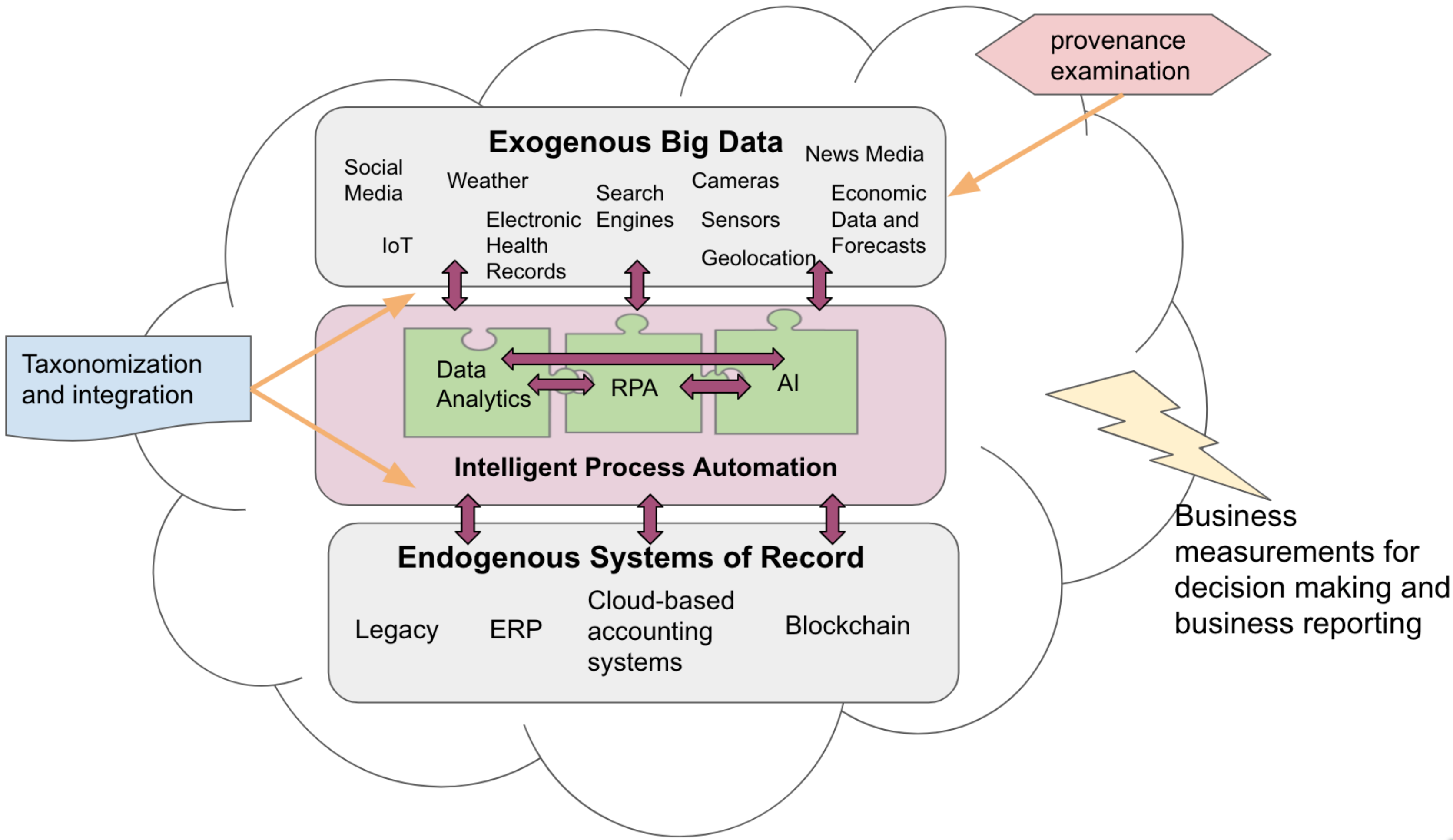
IPA can cover the Automation Continuum



(Adapted from Lacity & Willcocks, 2017)

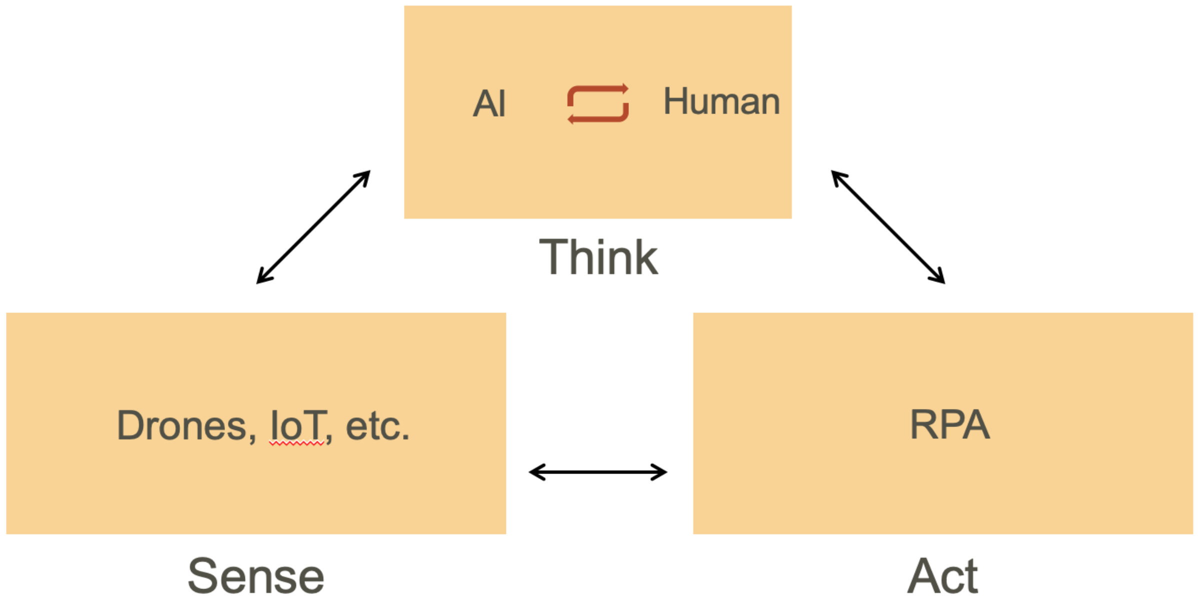
IPA


Audit Data Standards/
Taxonomization
Data
Analytics
RPA
AI
Audit Data Standards/
Taxonomization
Intelligent Process Automation


provenance
examination
Exogenous Big Data
Social Media
Weather
Search Engines
Cameras
News Media
Economic Data and Forecasts
IoT
Electronic Health Records
Sensors
Geolocation
Taxonomization and integration
Data Analytics
RPA
AI
Intelligent Process Automation
Business measurements for decision making and business reporting
Endogenous Systems of Record
Legacy
ERP
Cloud-based accounting systems
Blockchain

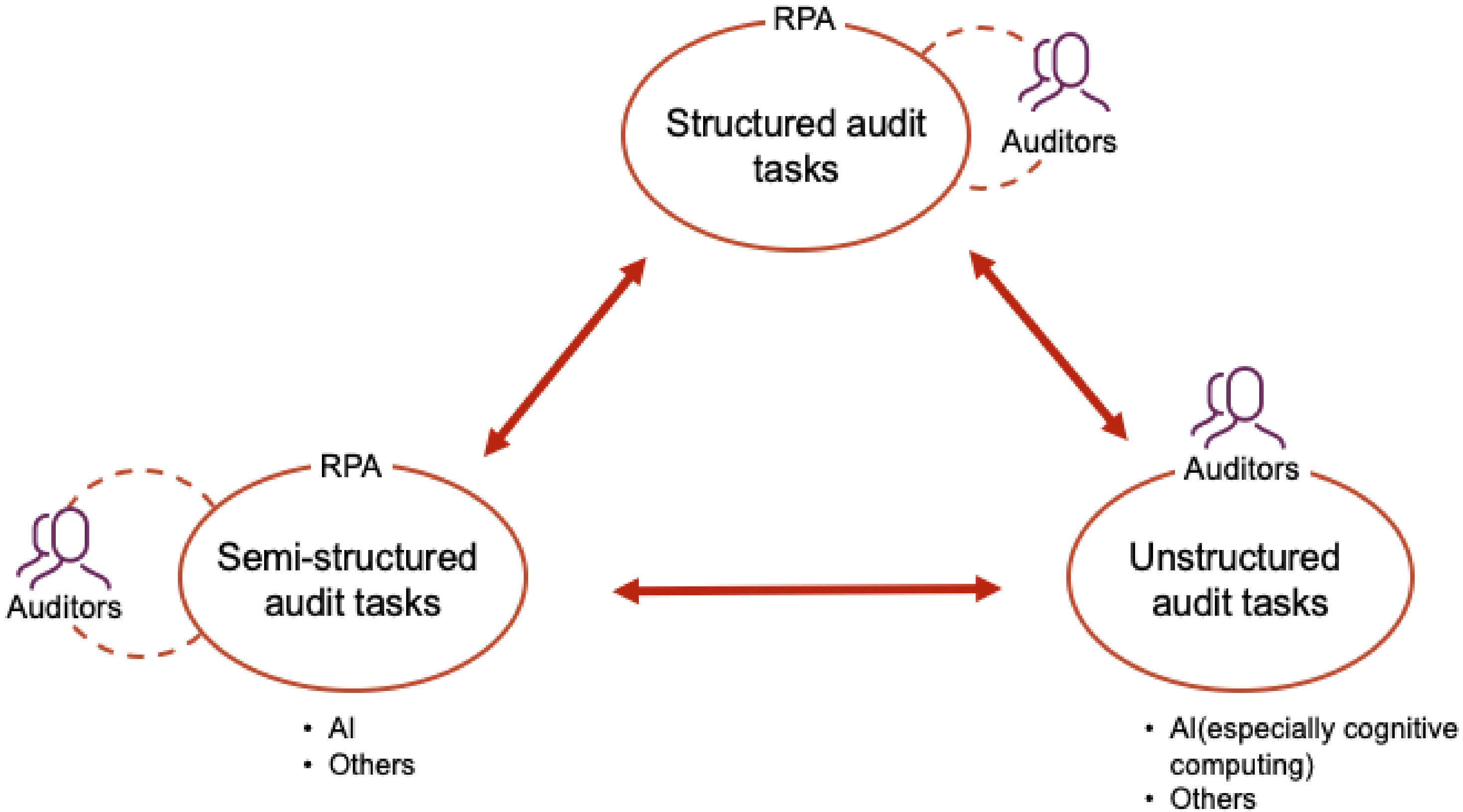
Intelligent Process Automation (IPA)

The "Sense-Think-Act" loop
of IPA



(Adapted from UiPath, 2017)

The "Auditor-in-the-Loop" IPA Ecosystem



RUTGERS
THE STATE UNIVERSITY
OF NEW JERSEY

Accounting Estimates Using Machine Learning

Keshing Ding, SWUFE
Baruch Lev, NYU
Xuan Peng, SWUFE
Miklos A. Vasarhelyi, Rutgers University

November 21, 2019 Toronto

RUTGERS

Accounting estimates

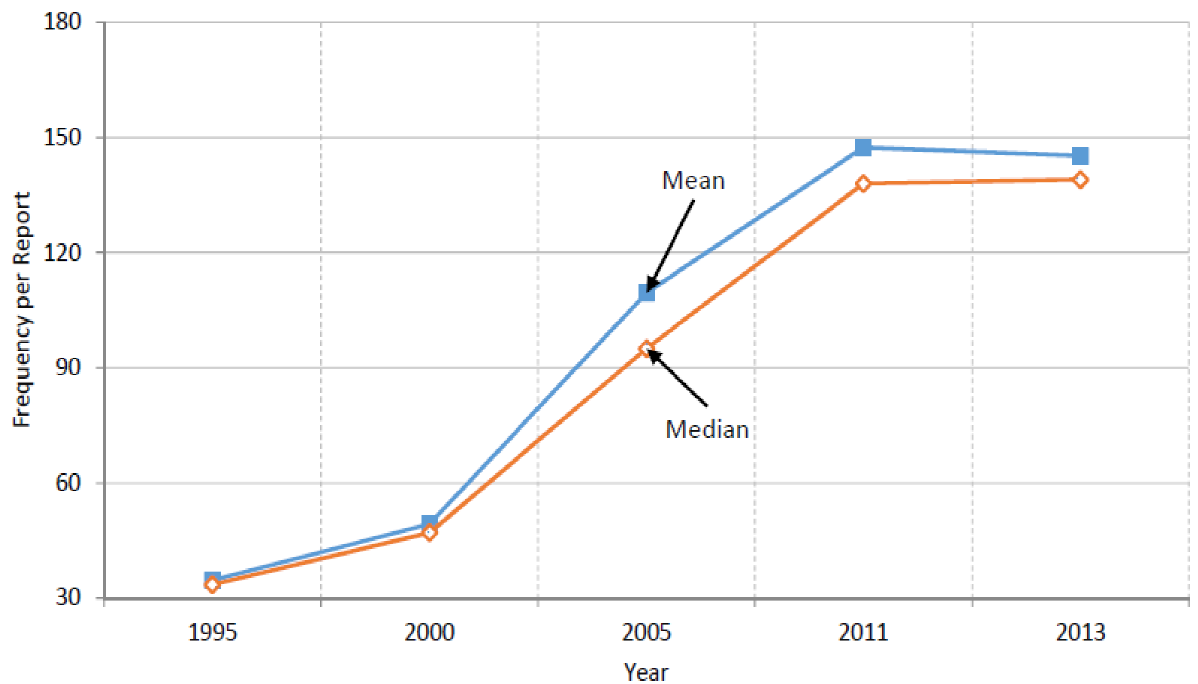
"Accounting estimates are pervasive in financial statements, often substantially affecting a company's financial position and results of operations… " (PCAOB 2018, p.3).

Accounting estimate examples:

- fixed assets
- accounts receivable
- pension expenses and incomes

Figure 5: Increasing Frequency of Estimates-related Terms in Financial Reports

For a sample of 50 S&P 500 companies (from: *The End of Accounting*)



Accounting estimates

- General Electric Example
  - 2016 net earnings is $8.2 billions.
  - Half came from a change in managers' estimates.

"Contract assets increased $4,006 million in 2016, which was primarily driven by <u>a change in estimated profitability</u> within our long-term product service agreements …"

Improve estimates

- Causes of estimation errors
  - environment uncertainty
  - managers' manipulation
- Machine learning
  - decreases manipulation: an independent, less-bias estimates generator
  - decreases uncertainty: take into account more factors in prediction
- Our Research
  - use machine learning algorithms to estimate losses for property & casualty insurance companies
  - compare machine learning estimates with managers' estimates

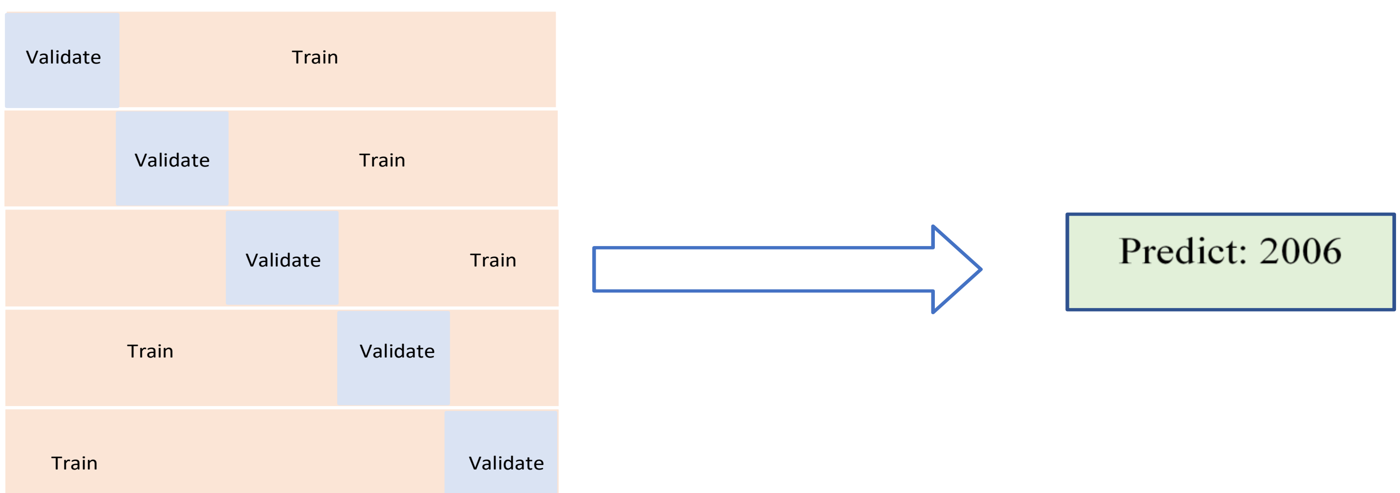
Research design

- Business lines (cumulative payment percentage)

Business Line	Year 0	Year 1	Year 2
Private Passenger Auto Liability	40.64%	72.44%	86.76%
Commercial Auto Liability	25.03%	50.74%	70.90%
Workers' Compensation	24.99%	56.11%	72.90%
Commercial Multi-Peril	44.52%	69.22%	80.03%
Homeonwer/Farmowner	72.62%	93.50%	96.83%

- Training/Validation/Testing approach

Cross Validation: 1996-2005



Cross-validation results

- The percent accuracy improvement of the ML loss estimates over managers' estimates in 5-fold cross validation.

Business line	Sample	Obs	Accuracy Edge
Private Passenger Auto Liability	1996-2005	5949	12%
	1996-2006	6298	13%
	1996-2007	6602	26%
Commercial Auto Liability	1996-2005	5383	42%
	1996-2006	5661	36%
	1996-2007	5957	37%
Workers' Compensation	1996-2005	4183	35%
	1996-2006	4398	43%
	1996-2006	4398	48%
Commercial Multi-Peril	1996-2005	5235	33%
	1996-2006	5457	34%
	1996-2007	5846	42%
Homeowner/Farmowner	1996-2005	6121	-12%
	1996-2006	6544	24%
	1996-2007	6946	24%

Holdout test results

- The percent accuracy improvement of the ML loss estimates over managers' estimates in holdout test.

Business line	Sample	Obs	Accuracy Edge
Private Passenger Auto Liability	2006	670	26%
	2007	659	14%
	2008	637	37%
Commercial Auto Liability	2006	620	20%
	2007	609	20%
	2008	592	49%
Workers' Compensation	2006	499	54%
	2007	498	55%
	2008	473	19%
Commercial Multi-Peril	2006	582	50%
	2007	570	22%
	2008	563	-18%
Homeowner/Farmowner	2006	697	51%
	2007	692	38%
	2008	678	52%

Conclusion

- Accuracy edge: accounting estimates generated by machine learning are potentially superior to managerial estimates.
- Benchmark: estimates generated by machine learning can be used by managers and auditors as benchmarks against which managers' estimates will be compared. Large deviations will suggest a reexamination of managers' estimates.
- Potential: machine learning could be used to generate estimates to be report in the first place.
  - enhance the reliability (no manipulation) and consistency of accounting estimates.

RUTGERS

The FASB could

- Create a machine learning estimate for a very narrow industry corresponding to reporting lines of business
  - Determine estimate based on an allocated percentage or and adjusted percentage of the business
- Allow businesses to do their computations and estimates with
  - A pre-set estimation methodology with machine learning or the machine learning done by the standard setter
  - The inputs to the estimation methodology (variables) be auditable values