

AI in Context: The End of Rules, The Dawn of Data

chris.wiggins@columbia.edu

Fall 2025 - Lecture 3 of 4

Reminder of Lectures 1 & 2

Lecture 1: Intelligence

- Available here
- "intelligence" as "objective" summary (of grades, "cleverness")

Lecture 2: AI 1.0 - Rules Rule

- Available here
- "intelligence" operationalized as engineering test
- fork in the road for AI: rules vs data (rules ruled)
 - "rules" wins: math, interpretability
 - "learning" requires memory, compute, infra, and data!

Lecture 3: Twilight of the Programmers, Dawn of Learners

Today's Topics

- **Data disputes** (data as "low road")
 - Example: the Perceptron story
- **Rules and rule-rot** (recall: knowledge bottleneck)
 - Result: "winter"
- **Precondition for learning: data**
 - The making of big data, WWII→present
- **Next week:** the rebrandings

Was There No Learning work?

What About Cryptography?

- Memory as challenge
- Shannon 1945: "sparks" of data paradigm

Recall: "Neuron nets" appeared in 1955 AI doc

Example Data Dispute: The Perceptron Story

NYT 1958: "NEW NAVY DEVICE LEARNS BY DOING"

"Psychologist Shows Embryo of Computer Designed to Read and Grow Wiser" ...

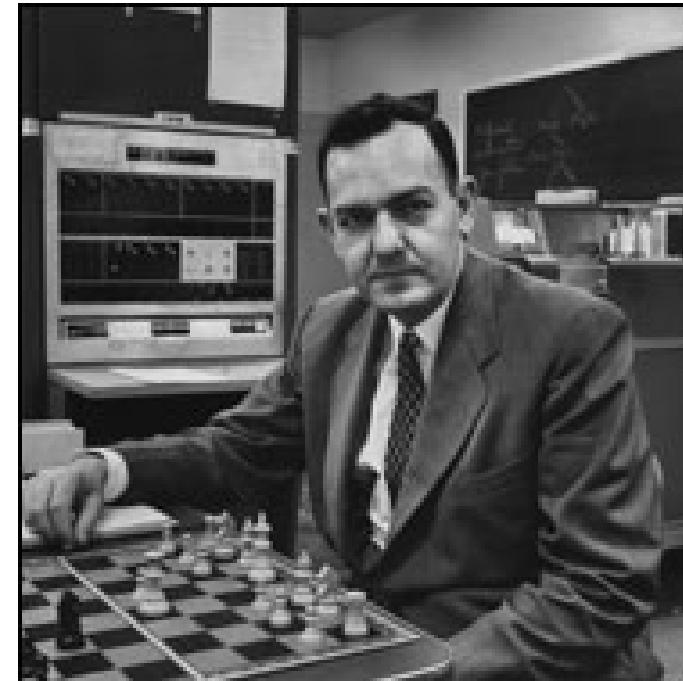
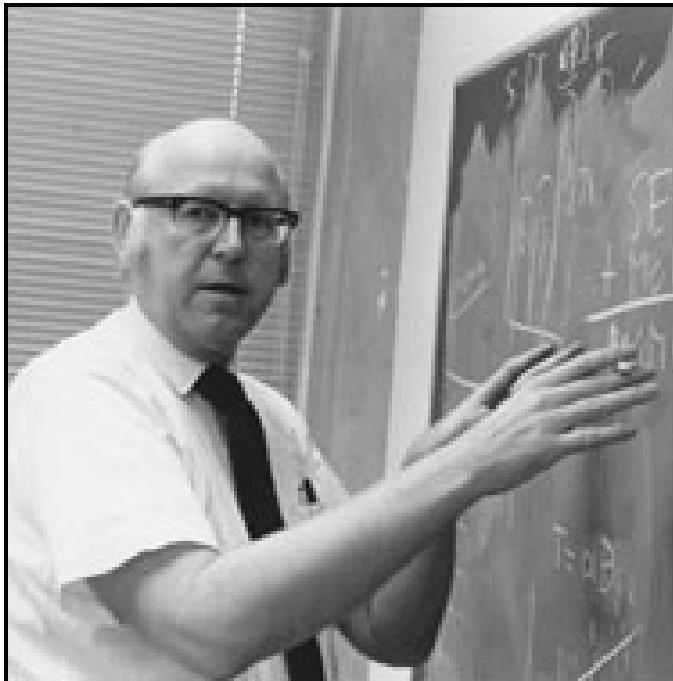
WASHINGTON, July 7 (UPI) -- The Navy revealed the embryo of an electronic computer today that it expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.

Perceptrons: Consequences ("Internal" Dispute)

The Attack on Rosenblatt

- [NYT 1958](#): "NEW NAVY DEVICE LEARNS BY DOING"
- [Perceptrons, 1969](#): Attack on Rosenblatt's feed forward neural net
- [Rosenblatt](#): Died July 11, 1971 (aged 43) Chesapeake Bay

Example: Simon on Learning (1980, 1983)



Context: Simon

- Attended Dartmouth '56
- Turing award '75
- Econ Nobel '78
- Prof of psych and CS at CMU

**WHY SHOULD
MACHINES LEARN?**

Herbert A. Simon
Carnegie-Mellon University

Example: Simon on Learning (1980, 1983)

Chapter 2. In R.S. Michalski,
J.G. Carbonell and T.M. Mitchell
(Eds.), *Machine learning, an
artificial approach*. Palo Alto,
CA: Tioga Publishing Co., 1983.
Bib. #552.

2

WHY SHOULD MACHINES LEARN?

Herbert A. Simon
Carnegie-Mellon University

The Skeptical Challenge

At the Carnegie-Mellon Machine Learning Workshop in July, 1980, Herbert Simon was asked to deliver the keynote address, where he chose to play the role of devil's advocate and ask the question "Why Should Machines Learn?"

[His analysis] concluded that, with the exception of cognitive modeling, some rethinking of long-term objectives was in order. After dispelling some common myths, Simon concluded with a clarified and more appropriate set of reasons why one ought to pursue machine learning research.

-- Mitchell

Simon 1980/1983

The Skeptical Challenge

"skeptical challenge to learning as the road to the future in AI... (sometimes called cognitive simulation, or information processing psychology)... -- H. Simon

"case against AI research in learning" -- H. Simon

"the most important kinds of learning research to carry out in AI are those that are oriented toward understanding human learning. Here as elsewhere, Man seems to be the measure of all things."

Simon'83: Tension

"Just Program It" vs. Learning

"Why Machine Learning? just program it"

"just sat down and programmed it"

vs.

- Schema (13x) & representations (16x)
- (calls out Perceptrons as not working)

Simon'83: The Core Question

"If we understand the domain ourselves, if we understand physics, why don't we just choose an internal representation and provide the problems to the system in that internal representation? What's all this learning and natural language understanding about?"

And Yet... Rule-Rot

The Knowledge Bottleneck Problem

- Mycin: practical?
- Mycin: code

People started to notice it wasn't quite working...

E.g., Dreyfus 65

1965 Dreyfus' *Alchemy* (then books, '72,'79,'86,'92)

ALCHEMY AND ARTIFICIAL INTELLIGENCE

Hubert L. Dreyfus

December 1965

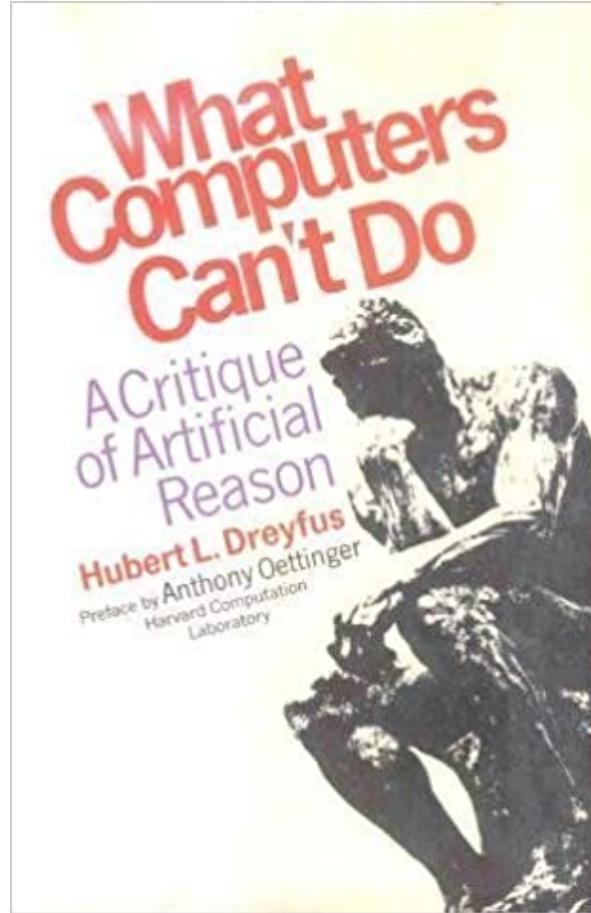
Dreyfus 65 on GOFAI Claims

"Within ten years, they claimed,

- a computer would be world chess champion,
- a computer would compose aesthetically valuable music,
- a computer would discover and prove an important unknown mathematical theorem, and
- most psychological theories would take the form of computer programs."

See also: Strong/AGI vs Weak/narrow AI

E.g., Dreyfus Books



In Translation (Military Goal): Dud

1966: ALPAC Report by Pierce (Bell)

"The Department of Defense, the National Science Foundation, and the Central Intelligence Agency have supported projects in the automatic processing of foreign languages for about a decade; these have been primarily projects in mechanical translation."

"there is no immediate or predictable prospect of useful machine translation."

Lighthill 1973

The Disappointing Reality

"Some very interesting researches have been carried out to develop general problem-solving programs, and such work can be of research interest to psychologists, but the performance of these programs on actual problems has always been disappointing. Students of all this work have generally concluded that it is unrealistic to expect highly generalised systems that can handle a large knowledge base effectively in a learning or self-organising mode to be developed in the 20th century."

"Winter" (1st Known Usage, 84)

Drew McDermott, AAAI 1984

To sketch a worst case scenario, suppose that five years from now the strategic computing initiative collapses miserably as autonomous vehicles fail to roll. The fifth generation turns out not to go anywhere, and the Japanese government immediately gets out of computing. Every startup company fails. Texas Instruments and Schlumberger and all other companies lose interest. And there's a big backlash so that you can't get money for anything connected with AI. Everybody hurriedly changes the names of their research projects to something else. This condition, called the "AI Winter" by some...

Ok But Was There a Winter?

Norvig's View: From Book (1995,2003,2009,2020)

"Overall, the AI industry boomed from a few million dollars in 1980 to billions of dollars in 1988, including hundreds of companies building expert systems, vision systems, robots, and software and hardware specialized for these purposes. Soon after that came a period called the 'AI winter,' in which many companies fell by the wayside as they failed to deliver on extravagant promises. It turned out to be difficult to build and maintain expert systems for complex domains, in part because the reasoning methods used by the systems broke down in the face of uncertainty and in part because the systems could not learn from experience."

Norvig's View: Personal Communication

"I'm a winter skeptic as well. I haven't seen good data.

"there was a winter in the sense that over time not everyone automatically got a job at MIT or Stanford or Bell Labs, but everyone got good jobs throughout.

"So essentially, we define the winter as the death of specific companies, not a widespread drop in students or professionals, and blame the decline on not enough Pearl and not enough ML. The timing is not quite right, as there were reports of winter before 1988, but that's the date we found a revenue number for. Later, we describe the current boom that has taken the industry to trillions of dollars."

Ok But Was There a Winter? (Continued)

- Japan **Fifth generation computer** (1982-1992, 400M 1992USD)
- IPTO **Strategic Computing Initiative**, 1B USD though S. Amarel and J. Schwartz cut back
- **Shift to VC \$\$** (wait until week 13)

But Things Changed Somewhere/Somewhen...

The Evolution of AI Approaches

- Remember: "computers" in 1956 were v. different (and limited)
- 1950s: Setting goals (but also: Sputnik, ARPA)
 - "In effect, ARPA reincarnated the World War II OSRD." --Edwards
- 1960s: Transformative funding, [shakey](#) victories
- 1970s: "winter" (cancellation of certain high-visibility projects)
- 1980s: Data curious AI researchers challenge rules

The Evolution Continues...

- **As late as 1990s:** It was believed that *rules* were the road
- **As late as mid 2000s:** It was believed that domain expertise mattered for predictive performance
- **2010s:** Shatter this via "deep nets", an echo of ANN's from 1943
 - "Overwhelmingly, machine learning systems are oriented towards one specific task: to make accurate predictions." -- SD2019

For This We Need Data; Where Could You Find It?

For This We Need Data; Where Could You Find It (and store and process it)?

Timeline: the "7 ages"

7 ages

7 Ages

- Retirement (2010) digital first
- Fulfillment (2000s) publishing, multimedia
- Mid-Life Crisis (1990s) WWW
- Maturity (1980s) libraries, research, AI boom/winter, CD-ROM vs online
- Adulthood (1970s) time-sharing, commercial systems, funding decline
- The Schoolboy (1960s) experiments, AI/statistics
- Childhood (1945-1955) microfilm, V. Bush

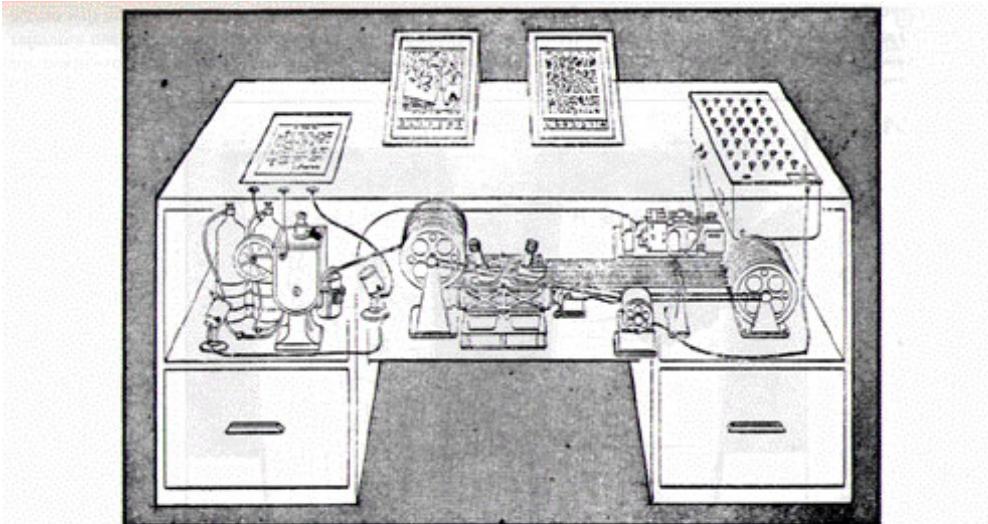
Vannevar Bush



Vannevar Bush

- Dangers of "great man" histories
- Reminder:
 - Differential analyzer, funded by Weaver
 - Proposed NSF
 - Students included Shannon, F. Terman (son of IQ; advisor of HP)
 - OSRD, (3x all pre-war science funding)
 - HUGE document post-war effort "by end of 1947...5M docs from Germany alone"
 - (Related: automated MT, e.g., Soviet, Japanese, German science → English)

Memex, 1945



Memex in the form of a desk would instantly bring files and material on any subject to the operator's fingertips. Slanting translucent viewing screens magnify supermicrofilm filed by code numbers. At left is a mechanism which automatically photographs longhand notes, pictures and letters, then files them in the desk for future reference (*LIFE* 19(11), p. 123).

"As We May Think" (1939/1943/1945) Big Ideas...

Futurism

"Consider a future device ... in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory."

Information Sciences

"When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass. It can be in only one place, unless duplicates are used."

"As We May Think" (Continued)

Demobilization

"It has been exhilarating to work in effective partnership. Now, for many, this appears to be approaching an end. What are the scientists to do next?"

Most well known for **memex** influence on hypertext

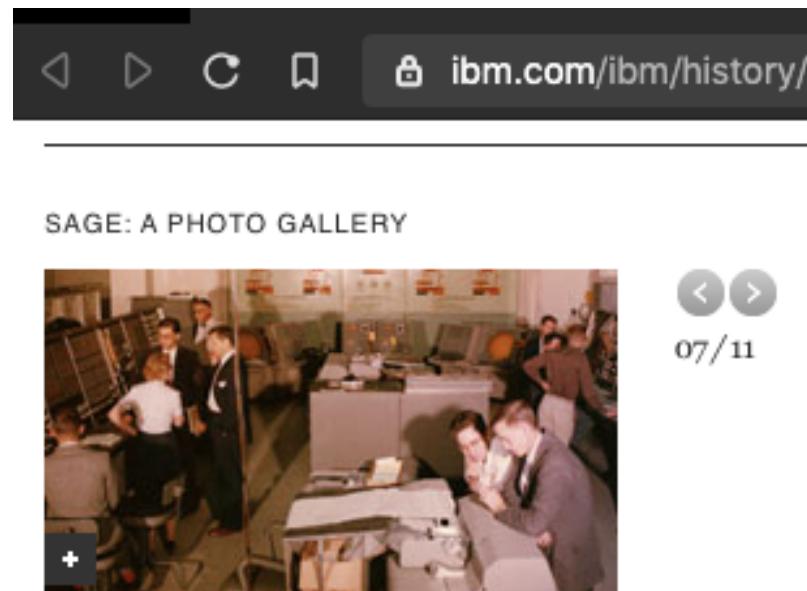
Timeline: 1950s, Growth of State Power

Military: SAGE Drives Complex Info

- Semi-Automatic Ground Environment
- .edu tie, esp MIT
- .com tie: IBM wins SAGE bid
- Streams = automated record-keeping, presages surveillance state

Commercial: Business Intelligence

- From records and punch cards to (digital, computational) "business intelligence"



Courtesy of The MITRE
Corporation

E.g., Luhn of IBM (1896-1964)



E.g., Luhn of IBM (1896-1964)

The Business Intelligence Pioneer

- **"Business"**: who was his employer?
- **Labor**: what were his patents?
- **Impact**: promoting BI to a function, centralizable, a skill
- **IR (information retrieval)** less emphasized than in Bush's Memex

Before Luhn, The Deluge

"Present methods of accomplishing this are inadequate and the general practice is to disseminate information rather broadly to be on the safe side. Since this method tends to swamp the recipients with paper, the probability of not communicating at all becomes great. The Business Intelligence System provides..."

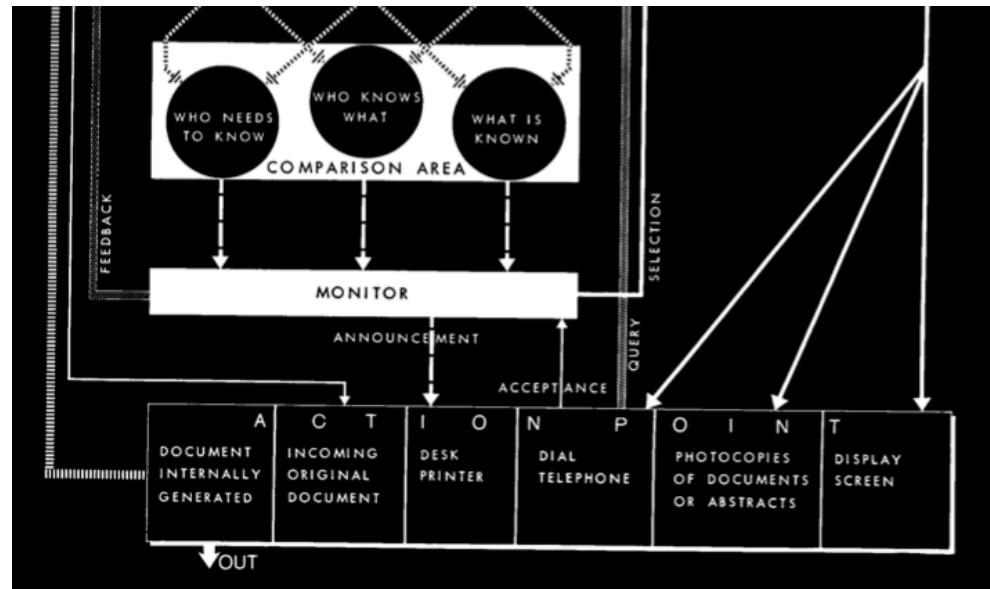
Infosec as Afterthought/ Move Fast, Break Things

*There are many **details** which might have to be provided to adjust the general form of the system to specific applications. One such requirement might be classification, by an editor, of documents with regard to security, proprietary interests, and proper utilization of information...*

Partial Democratization/ Product Mindset

5. The system is not to impose conditions on its user which require special training to obtain its services. Instead the system is to be operated by experienced library workers. Thus, in the case of an inquiry, the user will be required only to call the librarian, who will accept the query and will ask for any amplification which, in accordance with his experience, will be most helpful in securing the desired information.

Luhn the Inventor + Tinkerer



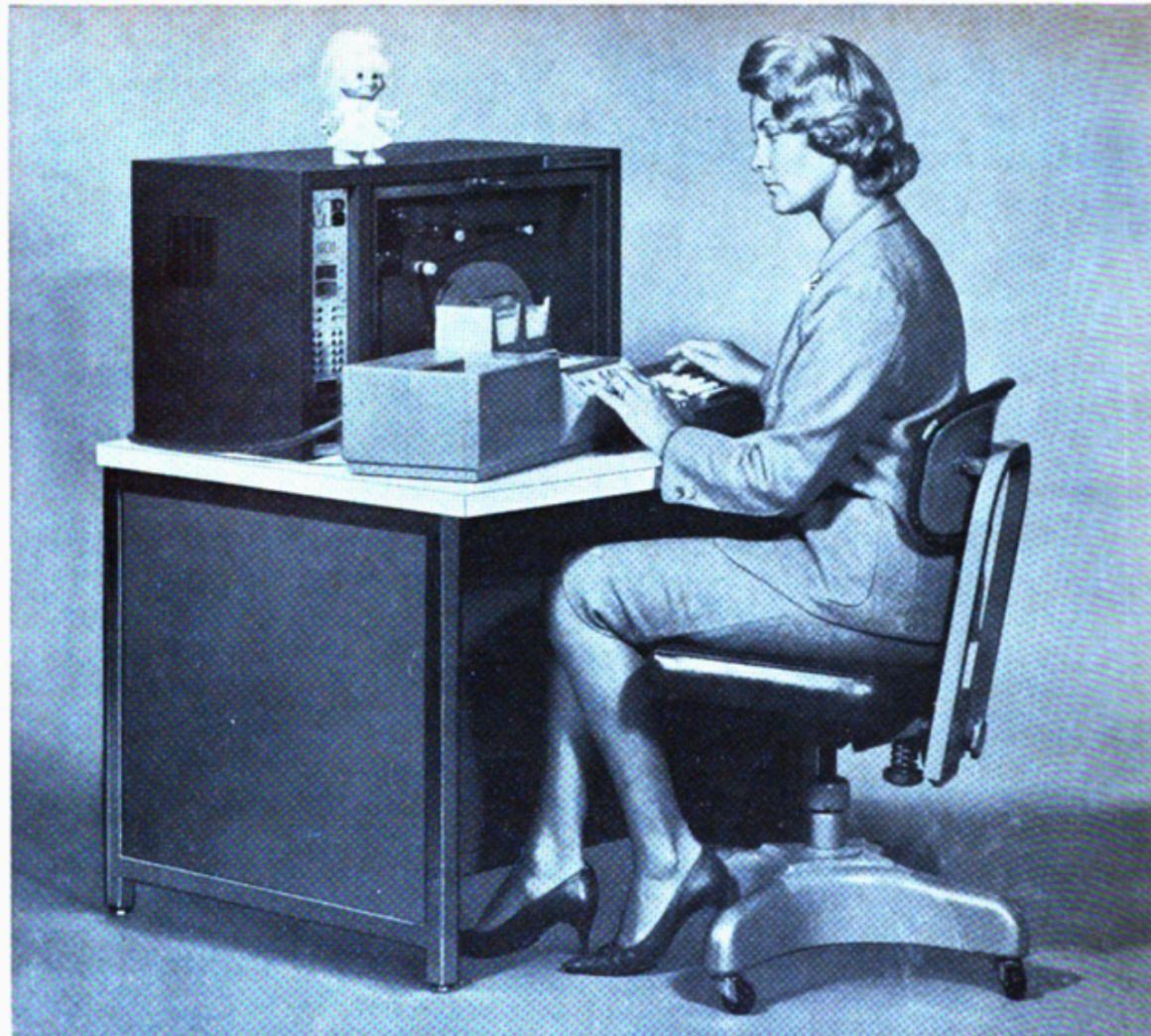
People are part of the information machine

Business loves surveillance (1974)

Modern technology permits credit grantors to respond to consumers efficiently and rapidly partially by virtue of accessing credit information through on-line terminal facilities or alternatively by telephone Inquiries. If the free flow of information is impeded by law, the resulting inefficiencies will necessarily be translated into higher costs to industry and consumer.

Statement Of The National Bank Americard, United States Congress Senate Committee on Government Operations Ad Hoc Subcommittee on Privacy and Information, I: 606

Missing: Labor (1/2)



Missing: Labor (2/2)

**This is for people who go wacky converting
data from cards to magnetic tape**



MDS 3011 Card Reader

The new MDS 3011 Card Reader and MDS 1106 (PCR) DATA-RECORDER simplify the job of adding variable information to magnetic tape along with coded data from punched cards. The 3011 offers either continuous feed or single-card entry . . . translates IBM punched card coding to BCD on magnetic tape in the 1106. On single-card entry, any required variable information can be transcribed directly to the tape in the 1106 by means of the 1106 Keyboard. No separate card punching is required. Verifying of keyed-in data is also performed on the 1106. Errors can be corrected on the tape.

The 3011 automatically reads 75 cards a minute, providing low-cost, card-to-tape conversion. It reads either 80 or 51 column cards . . . stacks them in the same order as the original deck.

The card-reading feature can be made inoperative and the 1106 can then be used for routine transcribing of data from source documents direct to tape, and verifying. The 1106 retains all the time-saving, cost-cutting features of the original MDS 1101 Keyed DATA-RECORDER. Interested? We'll gladly send complete details and arrange a demonstration. Write Dept. D-9:



MOHAWK DATA SCIENCES CORPORATION
Harter Street • Herkimer, New York 13350

SALES-SERVICE OFFICES IN MOST MAJOR MARKET AREAS

CIRCLE 24 ON READER CARD

40

Computer

Original from

DATAMATION

Martha Poon on Labor and FICO credit scores

This was a two-stage process performed by housewives working at piece rates of a few cents per sample application out of their homes. The meticulous work of 'the home- coders' was the backbone of the scorecard since it was their job to interpret the writing on the ledger cards and reliably convert it into the standardized numerical codes demanded by the analytic process. Codes were transferred to paper, reviewed for accuracy by a woman assigned as a 'checker', and subsequently transferred to punch cards so that the data could be read by machine.

Poon on Labor, Cont.

Because of its repetitive and mechanical nature, coding was considered a mundane task in the company. Yet upon scrutiny it is clear that the work that was done involved its own form of skilled decision-making that was far from obvious. A former coder made clear to me that '[t]here was some interpretation on all of this. You couldn't just copy it. That was the hard part, coding it. [. . .] They didn't just say he's been three times thirty days late in nice English' (Senior coder B). Another drew attention to the fact that, "We had to read these logs of payments and every company didn't do the same thing, and we'd get so confused" (Senior coder C).

Poon on Labor, Cont.

The Wells system involved teams of women working with boxes of punch cards and 42 pound Fridan [sic] calculators.

Timeline: 1990s (Lesk '96)

Optimism for WWW

- NB: WWW vs Internet (oh hi! it's our friend ARPA)
- (oh hi! it's our friend AI
 - expert systems
 - speech recognition)
- EFF and 1st amendment values: www will free us from big brother

Optimism for WWW

Next Week: The Dawn of Learning

Try This at Home!

Code

Looking Ahead

Next Week Preview

Lecture 4: "The Dawn of Learning"

- The rebrandings
- From pattern recognition to machine learning
- The data revolution

Questions?