Alípio Jorge
U.PORTO
INESCTEC
TECHNOLOGY & SCIENCE

and what can go wrong…

WHAT IS BOOSTING AI?

WHAT WAS AI THEN?

- Prolog, Lisp
- Planning, A*, Frames, Scripts
- Knowledge
- Expert Systems
- Great hopes
- Disappointment
- Half-time, AI is losing
- Second AI Winter follows

CAN MACHINES THINK?

- 1950
  - Alan Turing asks the question in "Computing Machinery and Intelligence"
  - The imitation game
  - The Turing test

ARTIFICIAL INTELLIGENCE

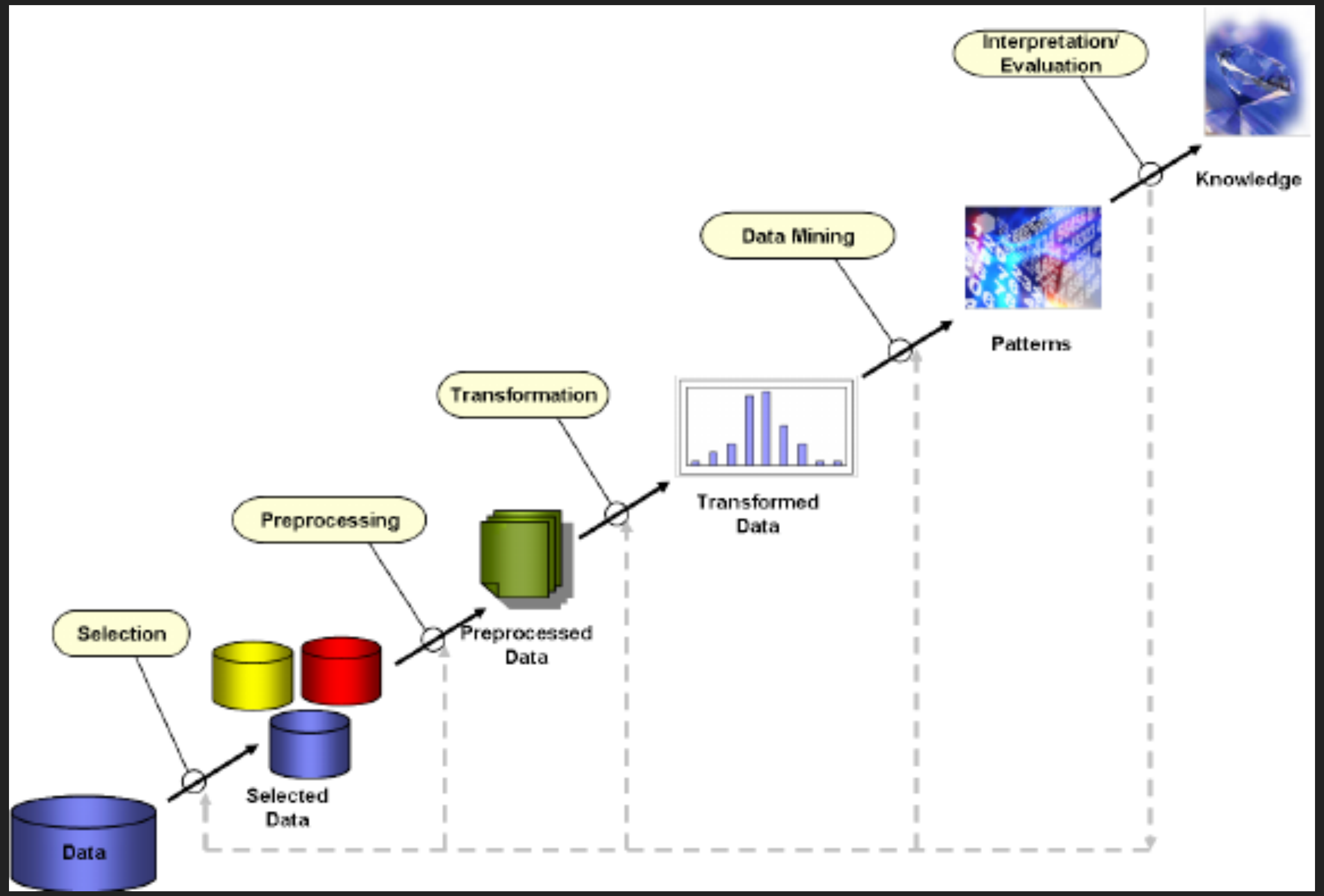
- 1952
  - UNIVAC predicts victory of Eisenhower
  - TV host “talks” to computer
- 1955
  - **McCarthy** and Minsky coin the term

MACHINE LEARNING

- 1959
  - Arthur Samuel coined the term
  - His checkers game program was able to learn how to play
    - machine played against itself to improve board evaluation

DATA MINING

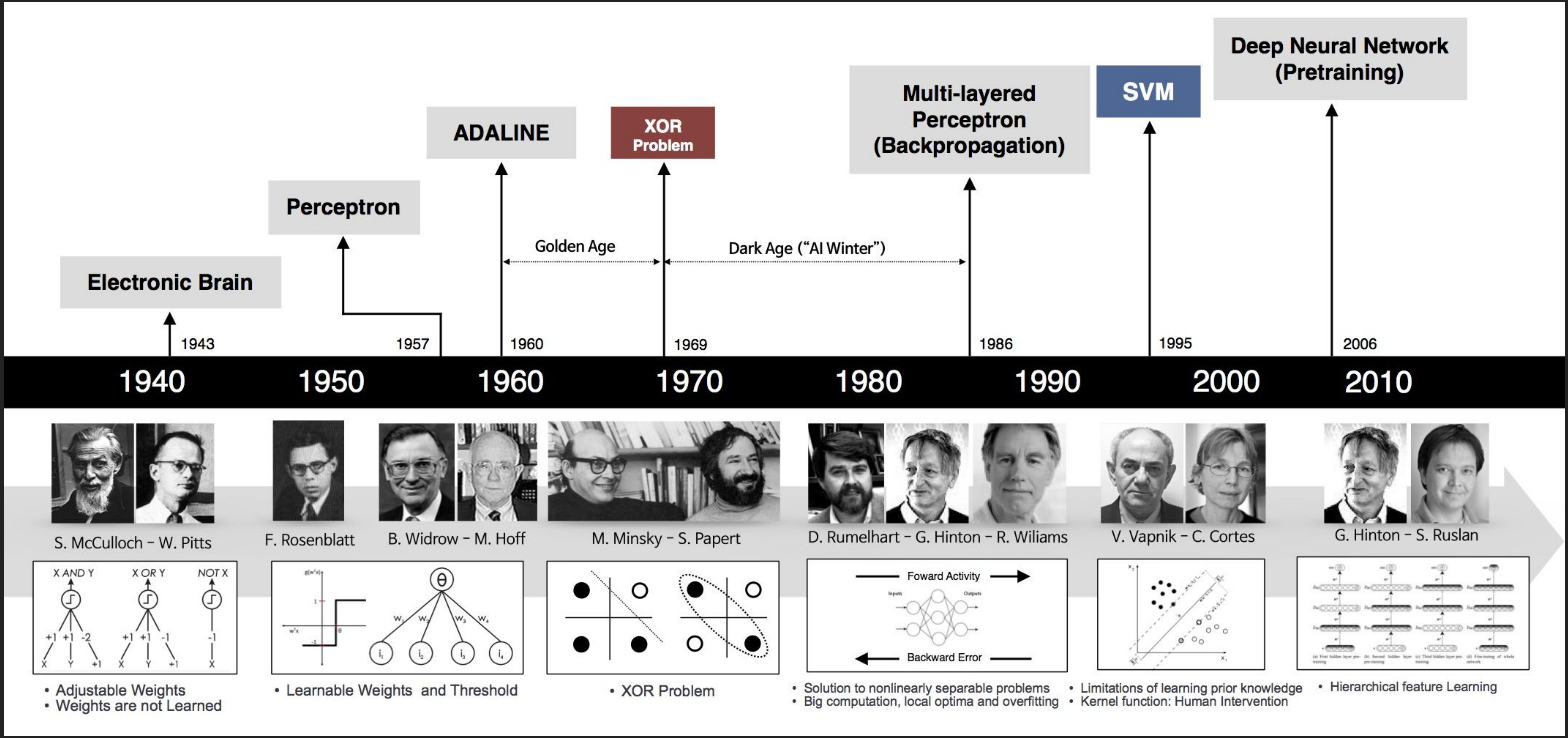
- 1983
  - Data fishing or experimentation
- 1990s
  - KDD - Knowledge Discovery in Databases
  - KDD-95 conference
- Since then
  - A problem-driven discipline
  - Very much human dependent (with exceptions)



DATA SCIENCE

- 1960
  - as “computer science”
- 1996, 1997
  - Statistics=Data Science?
- 2001 - 2003
  - Scientific journals
- 2000 circa
  - **Big Data**: Volume, Variety, Velocity
- 2012
  - **Data Scientist: the sexiest job**

CONNECTIONISM



Forbes

1,051 views | Apr 29, 2019, 05:34pm

So How Goes That AI Spring?

Jim Sinur Contributor
COGNITIVE WORLD Contributor Group
Independent Digital Business Consultant

Copyright Jim Sinur JIM SINUR

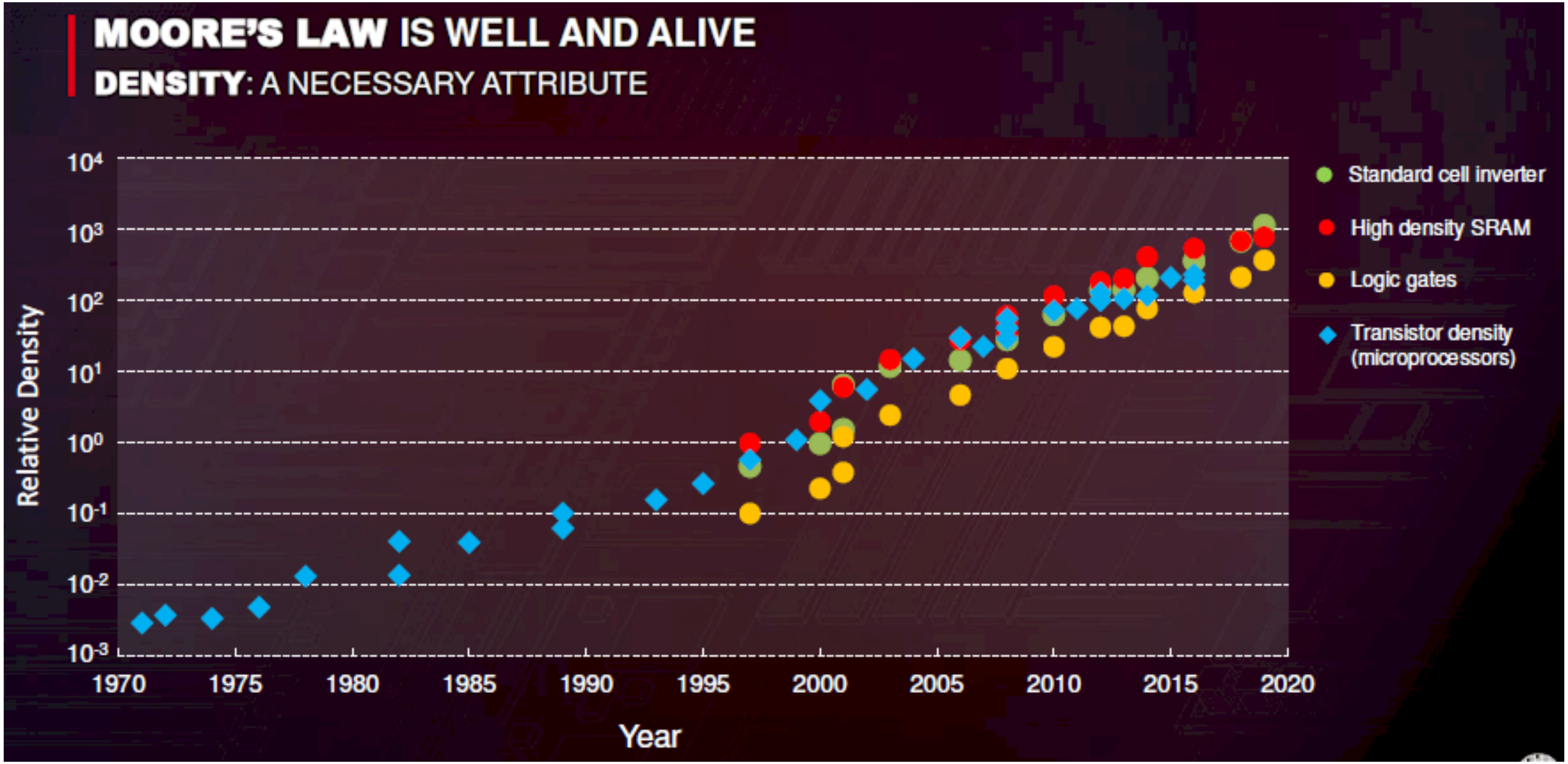
AND NOW?

ALL THAT DATA BOOSTED AI

- Corporate Information Systems
- The Web
- IoT

MODEL BASED AI → DATA DRIVEN AI

AND COMPUTING POWER



Forbes

1,051 views | Apr 29, 2019, 05:34pm

So How Goes That AI Spring?

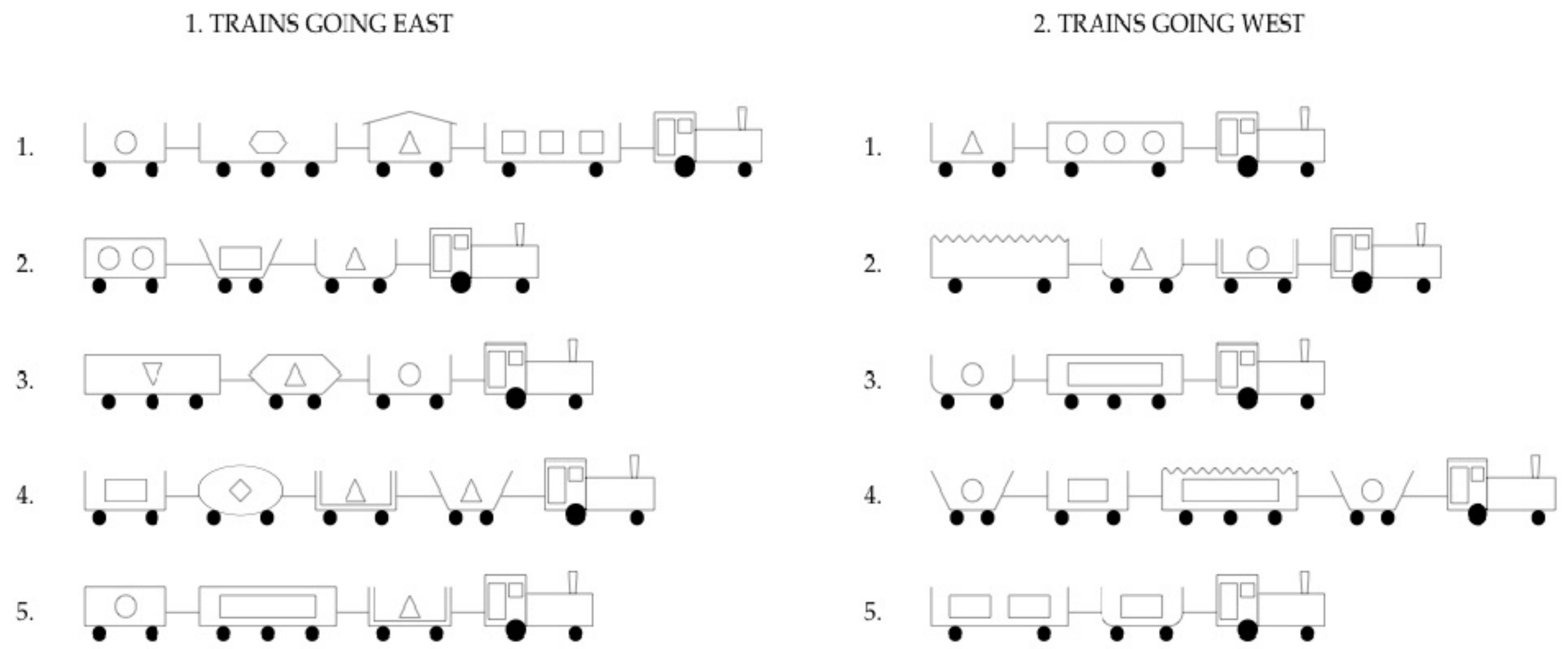
Jim Sinur Contributor
COGNITIVE WORLD Contributor Group
Independent Digital Business Consultant

Copyright Jim Sinur JIM SINUR

WHAT ELSE?

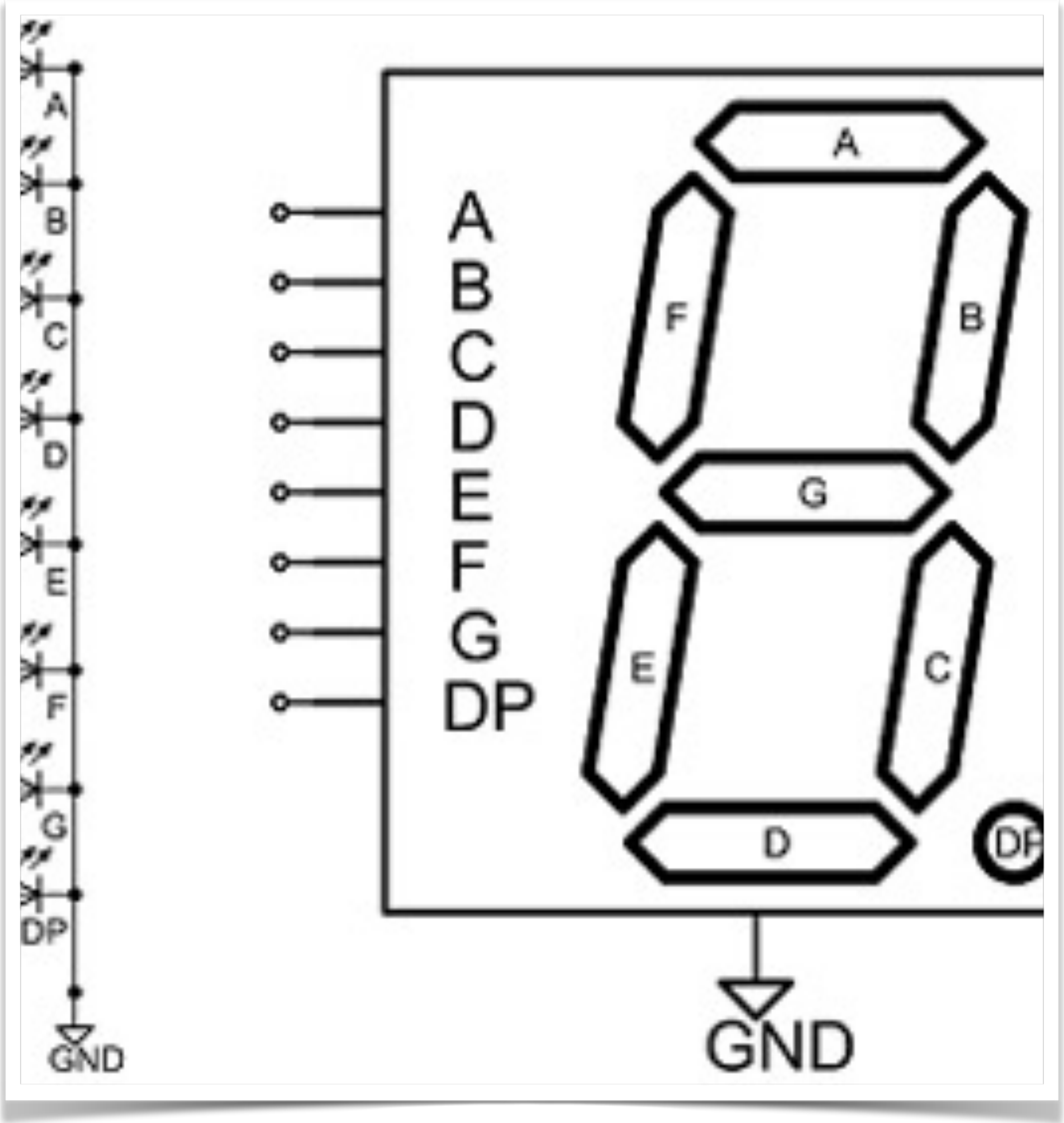
HOW TO BUILD A CLASSIFIER

- Michalsky trains
  - train goes east or west?
  - features
    - human palatable
    - language bias
    - the space of solutions is discrete
      - lacks smoothness
  - Useful, but…



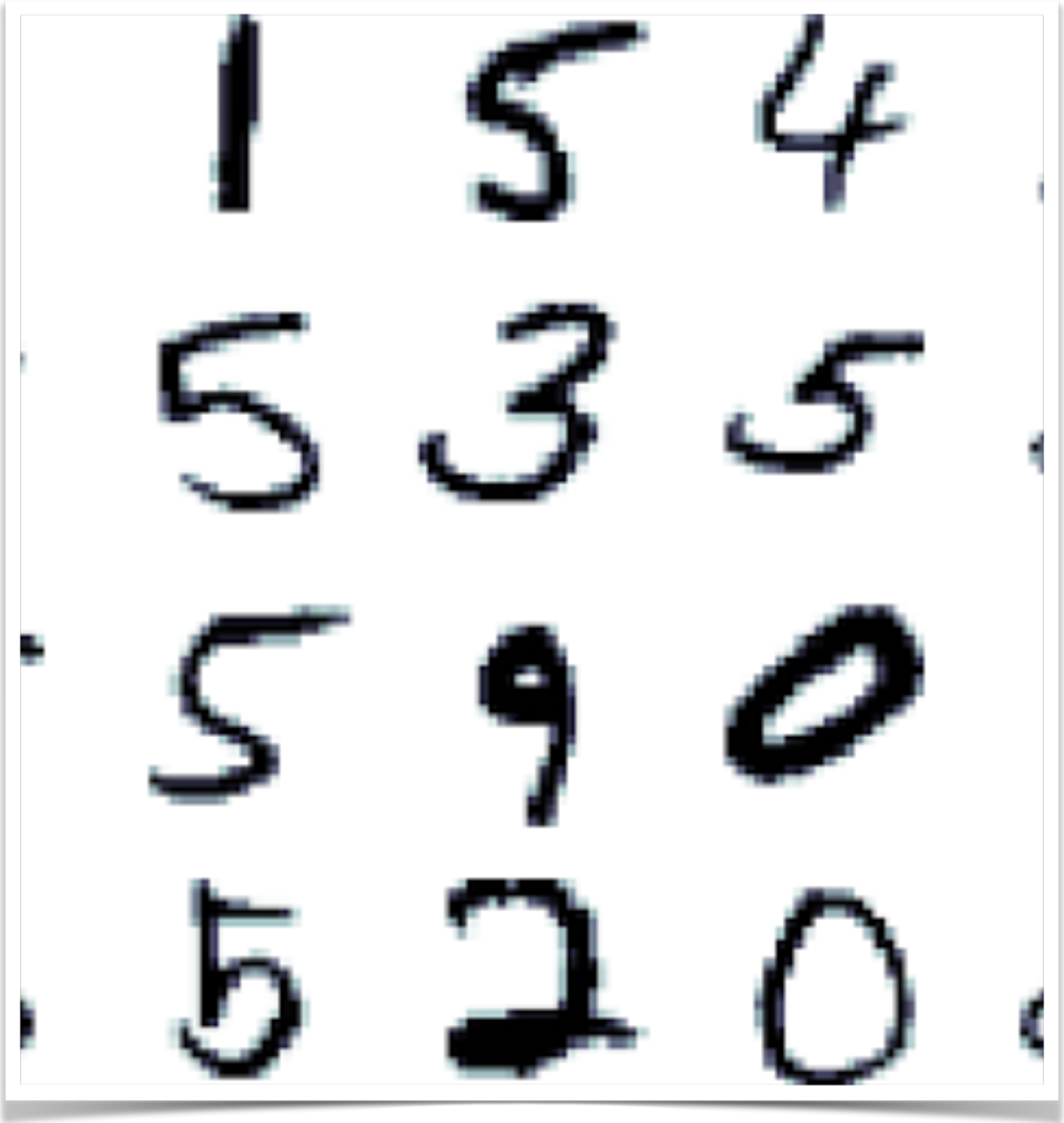
CLASSIFYING DIGITS

- Seven segments
  - Decision tree, SVM, MLP
- Feature Engineering
  - Human bottleneck
- Knowledge
  - Injected by humans
  - Hard to convey



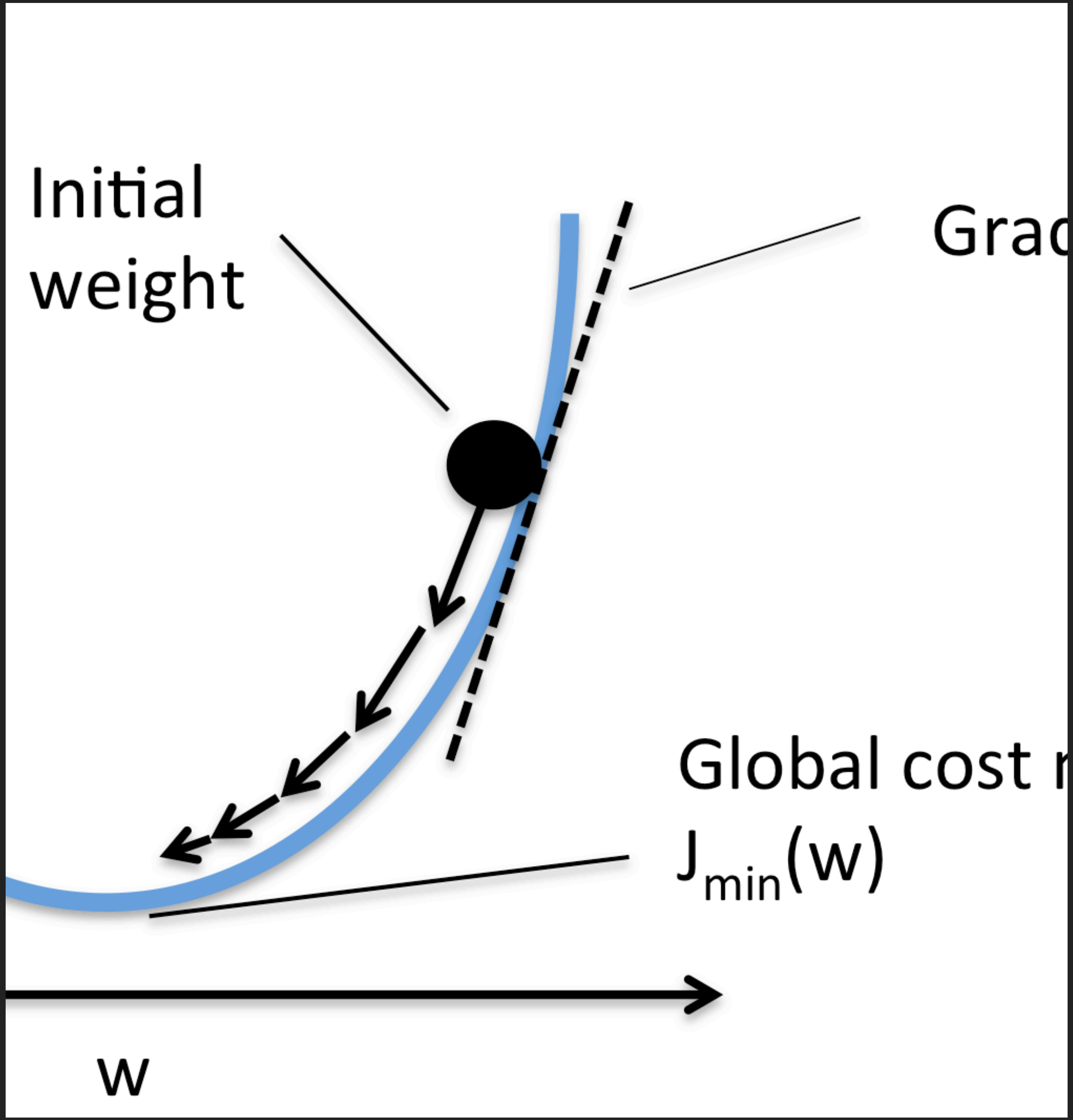
CLASSIFYING DIGITS

- Images
  - presented raw
- "Deep" Neural Network
  - Convolutional Neural Net
  - Learns representation
- No engineering?
- Artificial perception



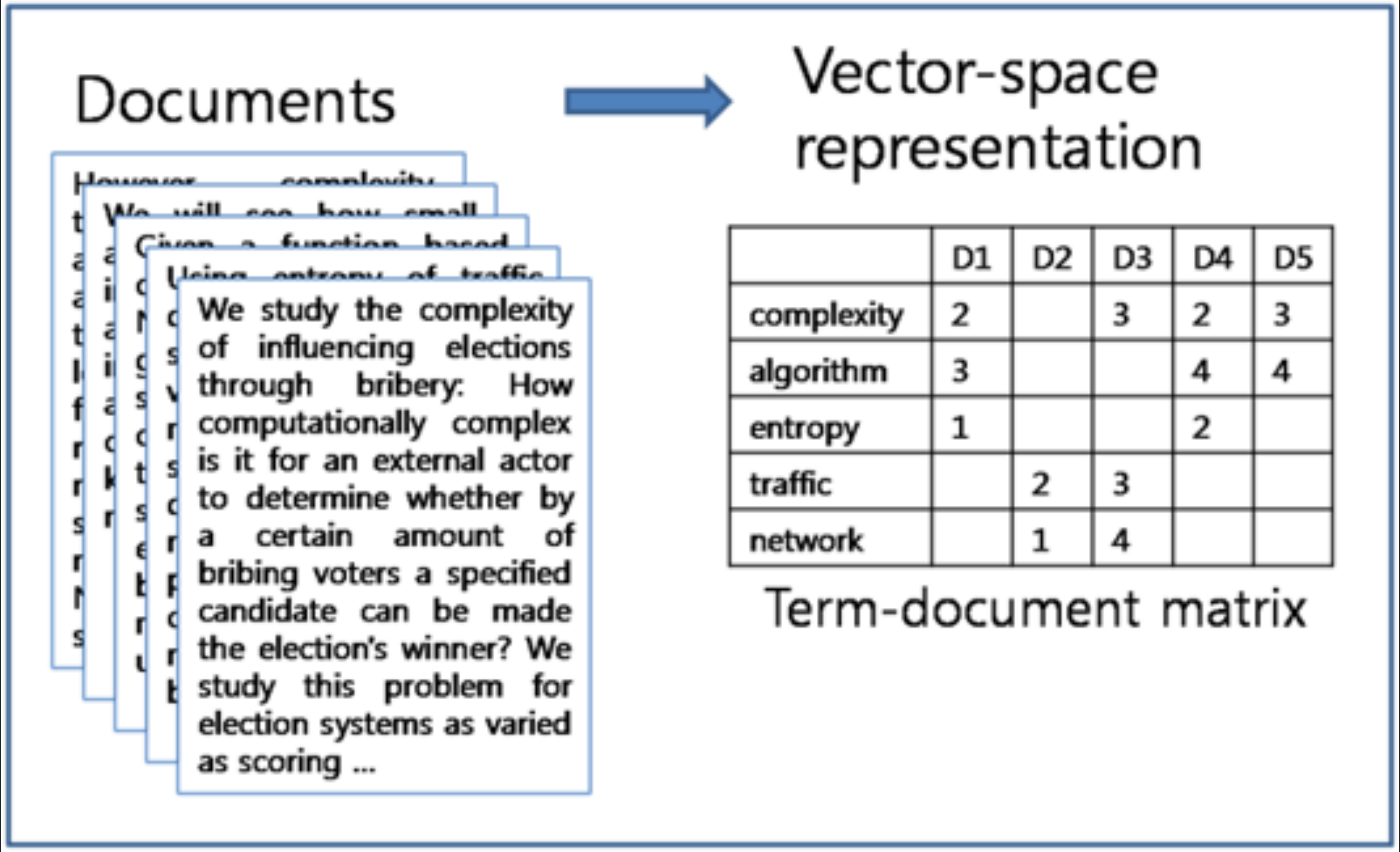
UNBLOCKING AI (ACTUALLY, MACHINE LEARNING)

- Search space
  - discrete vs continuous
- Features
  - engineered vs raw
  - approximate vs "optimal"
  - learned to the task



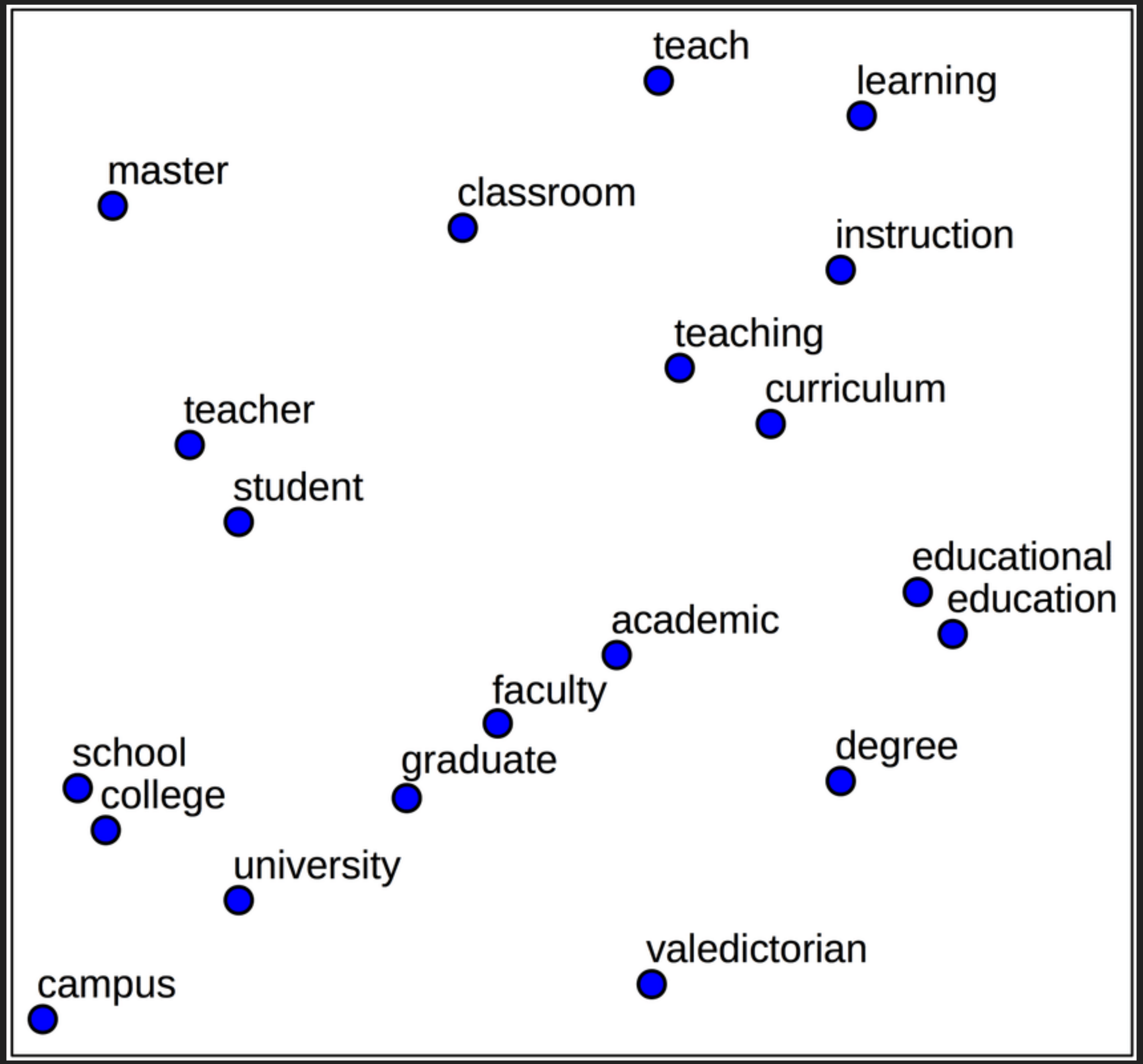
NLP / INFORMATION RETRIEVAL

- Vector representation
  - coded from text
  - TF-IDF
  - tens of thousands of dimensions
  - sparse, stiff, not smooth



NLP / INFORMATION RETRIEVAL

- Word embeddings / Neural language models
  - are **learned** not just coded
    - predict the next word in a ______
  - do not require (much) engineering
  - tens of dimensions do a good job
  - dense
  - continuous
  - enable semantic inference



NLP / INFORMATION RETRIEVAL

- BERT
  - Bidirectional Encoder Representations from Transformers
    - predicts “masked” words from context
  - learns a language model
  - can be used for multiple NLP tasks
    - NER, sentiment identification, translation
  - Applicable to sequences

RECOMMENDATION

- Natural representation
  - A user is a vector of size #items
  - Sparse, hard to relate, …

USER	ITEM
1	A
1	B
1	G
2	A
2	C
3	B
3	G
3	F
3	I
4	B
4	C
5	G
5	F
5	I
5	J
6	A
6	C

	A	B	C	F	G	I	J
1	1	1	0	0	1	0	0
2	1	0	1	0	0	0	0
3	0	1	0	1	1	1	0
4	0	1	1	0	0	0	0
5	0	0	0	1	1	1	1
6	1	0	1	0	0	0	0

RECOMMENDATION

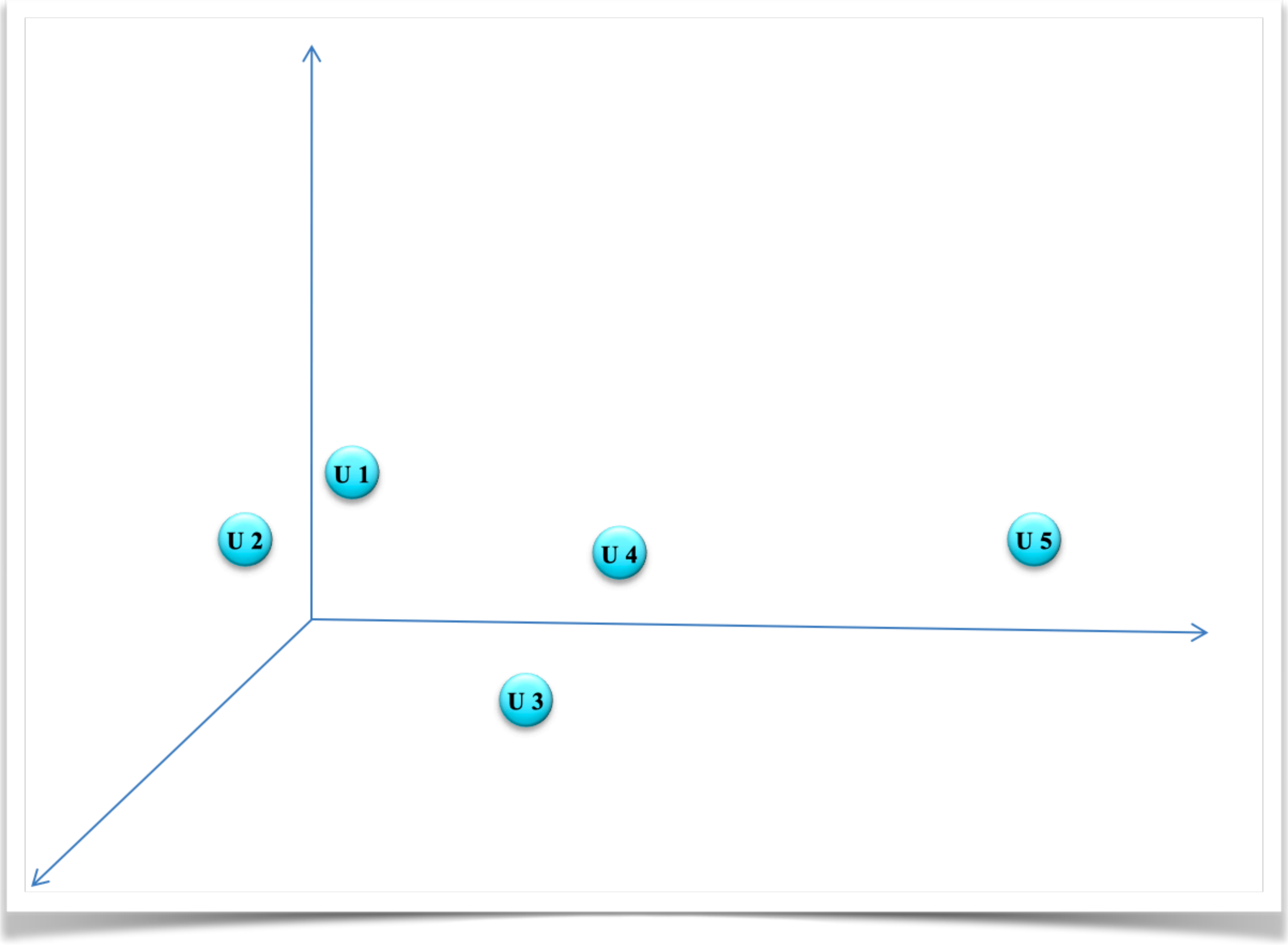
- Using **association rules**
  - recommendations very easy to understand
  - fixed representation
  - uniqueness and unrelatedness of items and users
  - low support
  - hard to see a bigger picture

	A	B	C	F	G	I	J
1	1	1	0	0	1	0	0
2	1	0	1	0	0	0	0
3	0	1	0	1	1	1	0
4	0	1	1	0	0	0	0
5	0	0	0	1	1	1	1
6	1	0	1	0	0	0	0

A & B —> G

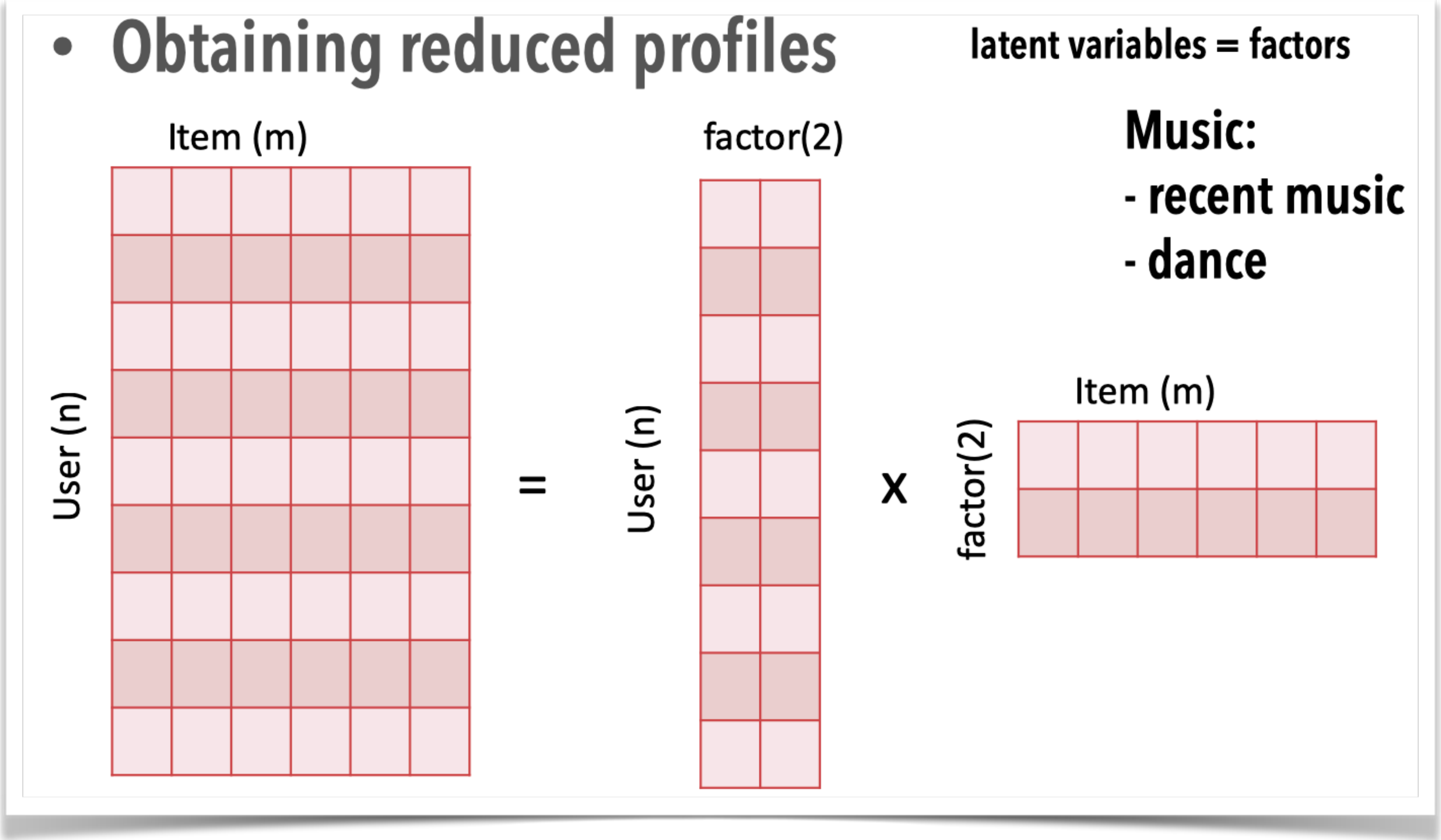
RECOMMENDATION

- Using **similarity based approaches** (aka collaborative filtering)
  - solves some problems of AR
  - still very sparse
  - fixed (no learning)
  - harder to read



RECOMMENDATION

- Dense, learned representations
  - Matrix factorization
  - "Optimal" features
  - Low-dimensional
  - Dense
  - Capture semantics
  - … unintelligible



WHAT MAKES THE DIFFERENCE?

- Learning from raw data
- Smooth representation space
  - similar objects have similar representation
  - objects can be arbitrarily similar
- Dense representations
- Hierarchical structure of features
- Learning as optimisation in a continuous space
- Dealing with incomplete information in a efficient way

LEARNED REPRESENTATIONS AND TRANSPARENCY

- Learned representations are for machines, not humans
- Can we get away with it?
  - the importance of algorithmic transparency
  - explainability (XAI)
  - human in the loop
- What can we do?
  - combining symbolic and sub-symbolic ?
  - bayes networks and causality

WHAT NEXT?

WHAT NEXT IN AI?

- human in the loop / human-centric AI
- autonomy
- grounding
- causality
- green AI
    - neuromorphic, Cerebras AI chip
- edge AI
    - federated learning

REFERENCES

- Representation Learning: A Review and New Perspectives, Yoshua Bengio, Aaron Courville, and Pascal Vincent , arXiv, 2014
- An overview on the exploitation of time in collaborative filtering. João Vinagre, Alípio Mário Jorge, João Gama, **Wiley Interdisc. Review.: Data Mining and Knowledge Discovery** 2015
- The Night A Computer Predicted The Next President: www.npr.org.
- A proposal for the dartmouth summer research project on artificial intelligence: McCarthy, Minsky, Rochester, Shannon
- Computing Machinery and Intelligence: Alan Turing
- Moore’s Law is Alive and Well, Eric Martin, medium.com, Dec 22, 2018
- So How Goes That AI Spring?, Jim Sinu, Forbes, 2019