CAP 6635 – Artificial Intelligence

Lecture 4 : Netflix Prize and beyond (Challenges in AI)

Oge Marques, PhD

Professor

College of Engineering and Computer Science

College of Business

FAU FLORIDA ATLANTIC UNIVERSITY

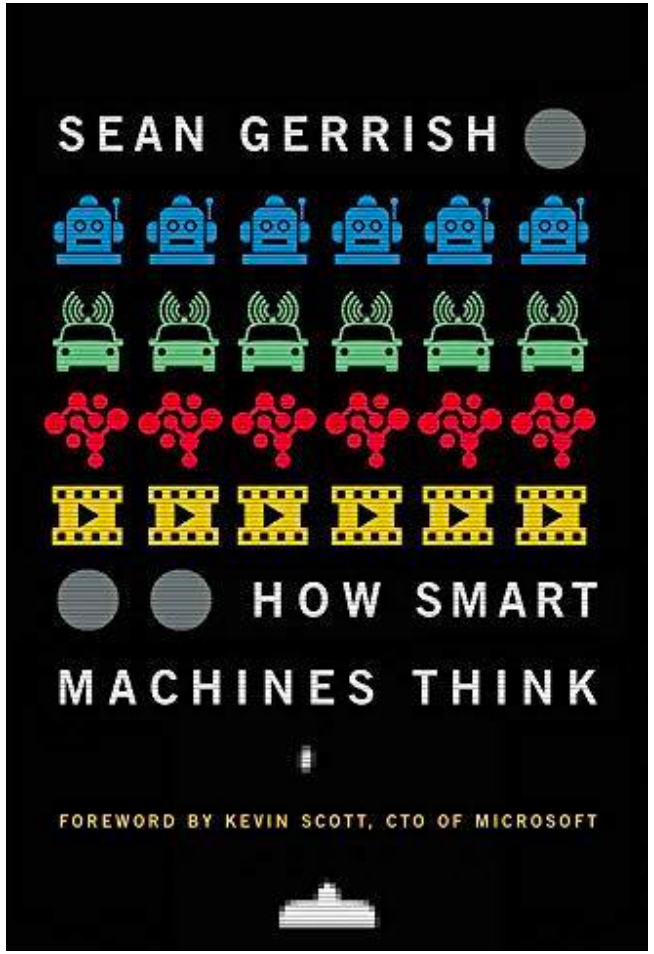
@ProfessorOge

ProfessorOgeMarques

Take-home message

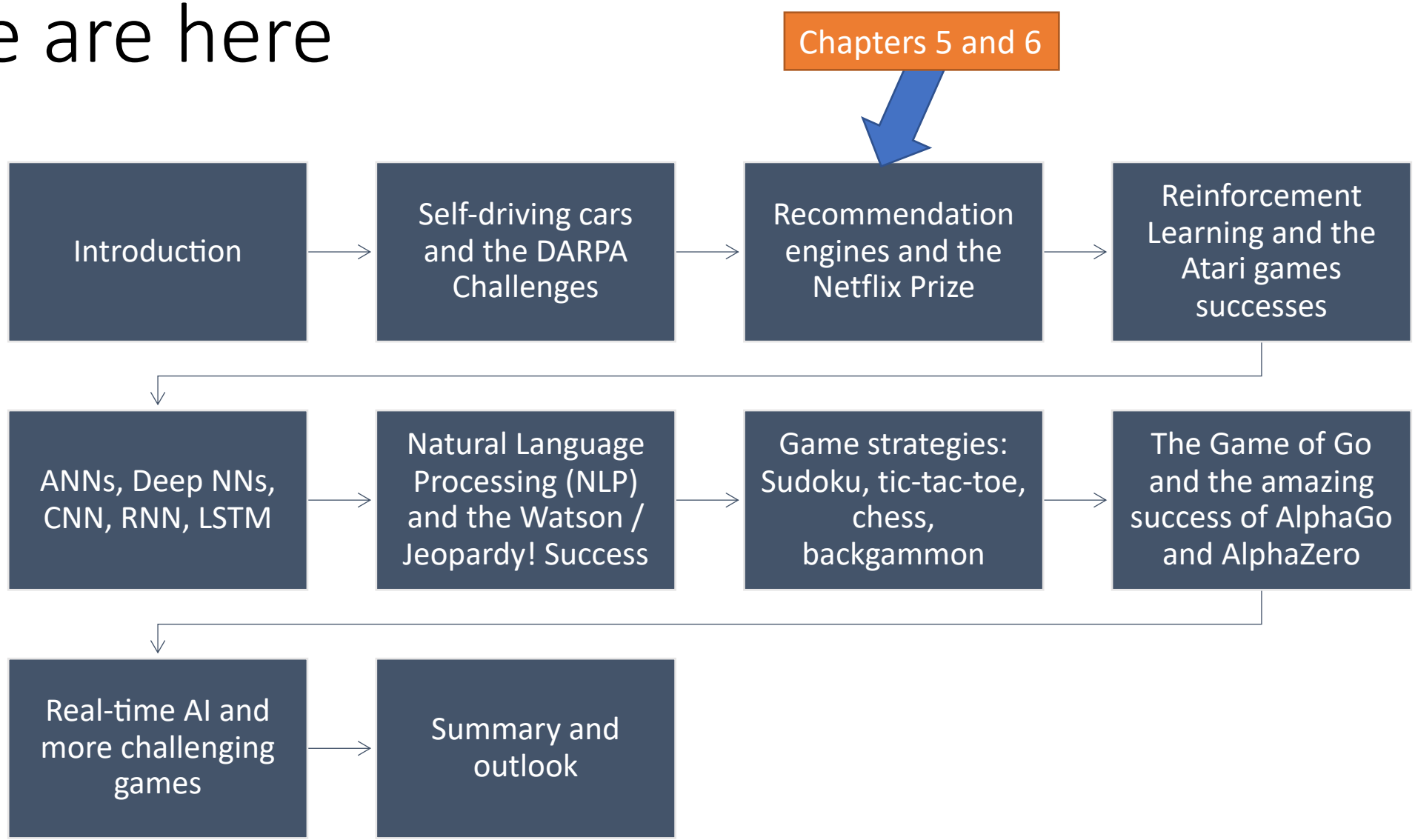
Progress in Artificial Intelligence (AI) (and related areas) **has been accelerated by challenges, prizes, hackathons, etc.**

These are also a **great source of learning opportunities.**



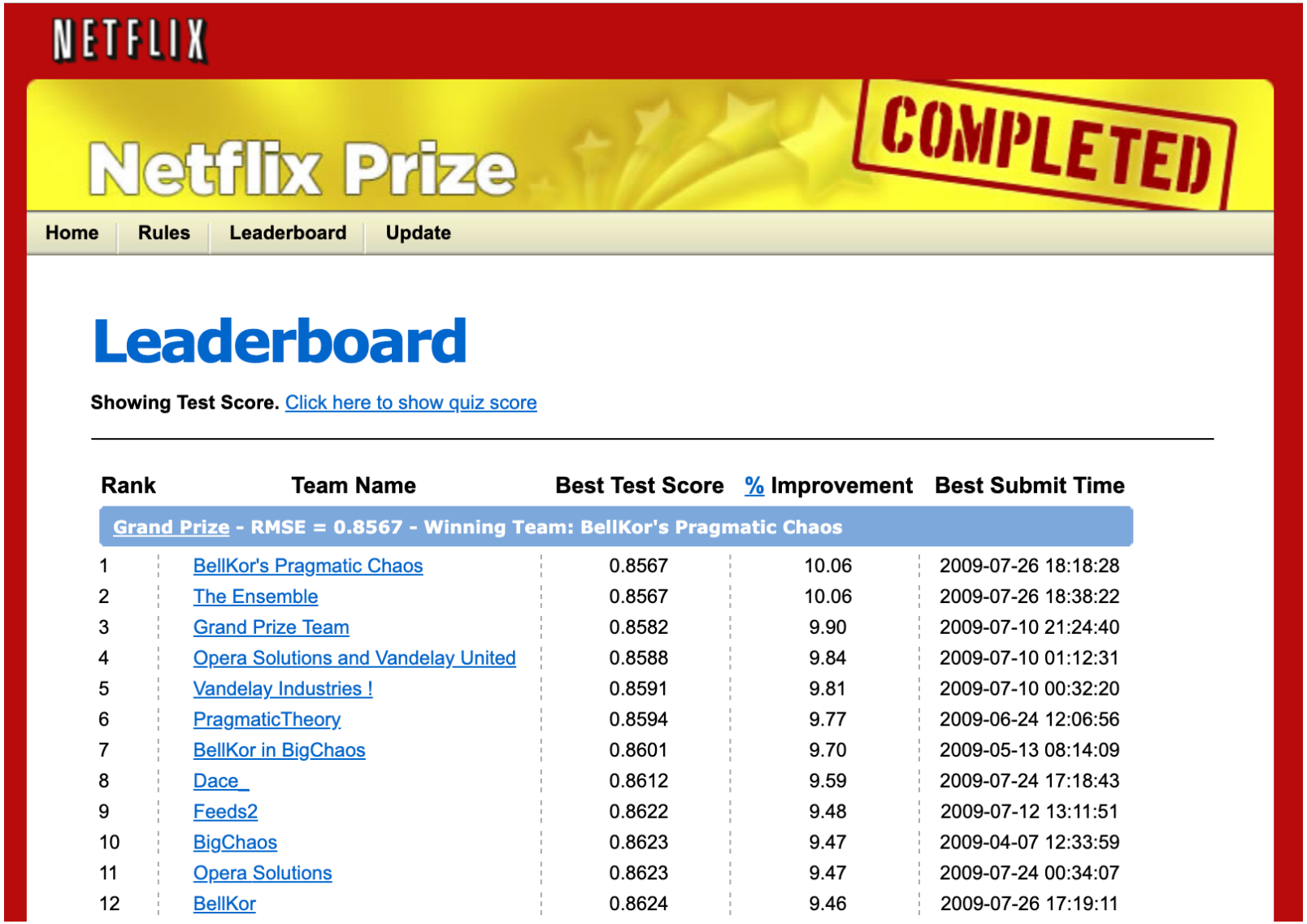
How smart machines think (by Sean Gerrish)

We are here

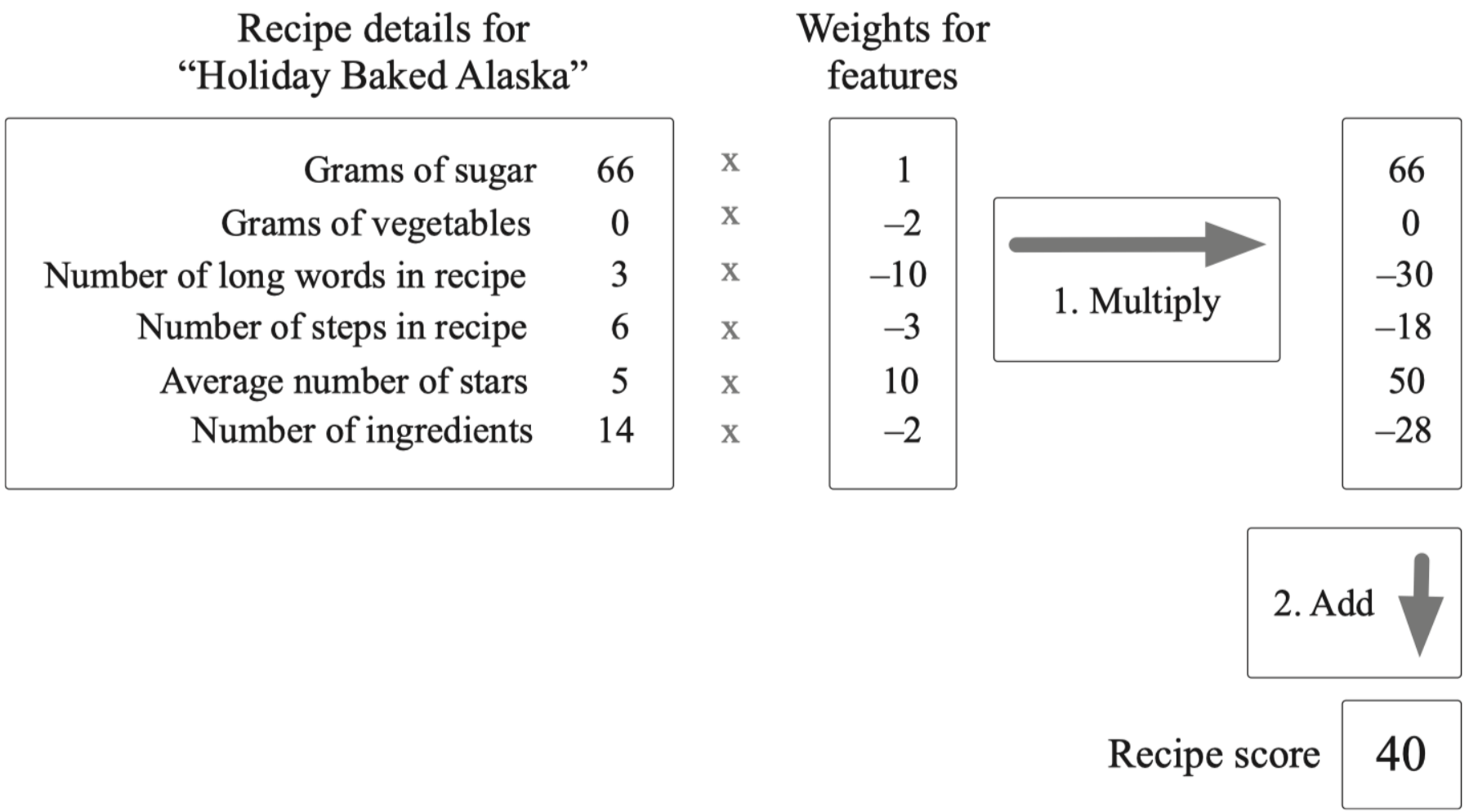

Chapters 5 and 6
Introduction
Self-driving cars and the DARPA Challenges
Recommendation engines and the Netflix Prize
Reinforcement Learning and the Atari games successes
ANNs, Deep NNs, CNN, RNN, LSTM
Natural Language Processing (NLP) and the Watson / Jeopardy! Success
Game strategies: Sudoku, tic-tac-toe, chess, backgammon
The Game of Go and the amazing success of AlphaGo and AlphaZero
Real-time AI and more challenging games
Summary and outlook

Netflix and the recommendation-engine challenge

https://www.netflixprize.com/



How to train a classifier



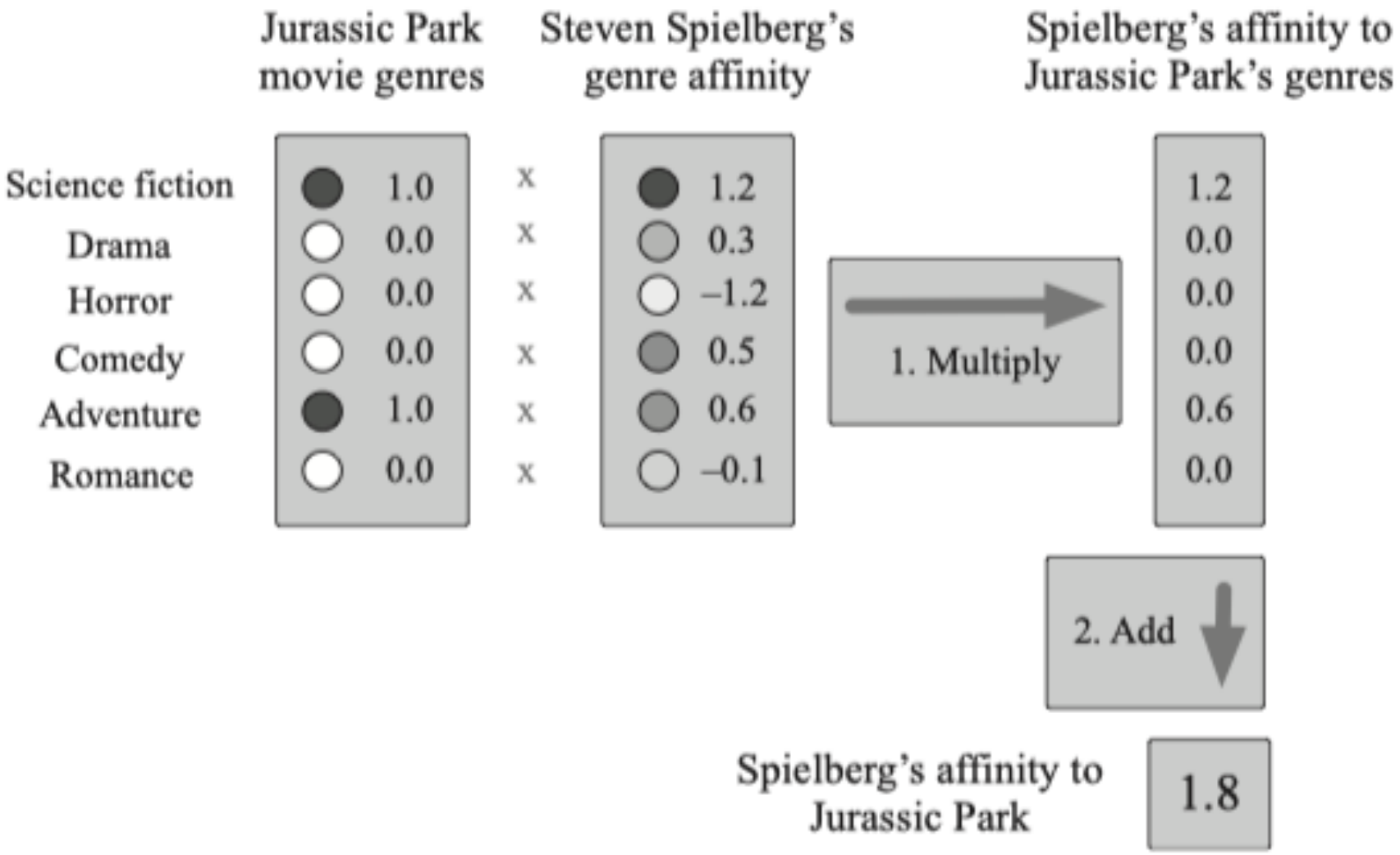
A ratings matrix (17,770 x 480,189)

	User 1	User 2	User 3	User 4	User 5	…	User 480185	User 480186	User 480187	User 480188	User 480189
Terminator 2	5		5		4	…	2	5			5
Gummo	1	1	2	?		…		3	2		?
Clueless		4		?		…	2		4		
Napoleon Dynamite	4		2			…		5	5		
Pan's Labyrinth	4					…		5		5	
…	…	…	…	…	…	…	…	…	…	…	…
The Peanut Butter Solution	3				4	…	?	?			
X-Men	?			4		…	2	4			5
Edward Scissorhands	5			5		…		5			
Short Circuit	4	4				…	1				
Toy Story		?	4		5	…		4			

Matrix factorization

- Matrix factorization relies on the fact that the giant ratings matrix has a lot of redundant information
- It is the most important algorithm for making personalized recommendations
- It assumes that we've summarized each movie with a few numbers and each user with a few numbers, and it provides a way to combine these sets of numbers into a score to describe how much any user will like any movie.
- This is called *matrix factorization* because, the way the math works out, it's equivalent to approximating the original, gigantic ratings matrix as the matrix product of two or more much smaller matrices—its *factors*—which encode exactly the numbers we've used to describe the movies and users.

Matrix factorization: example



Predictions over time

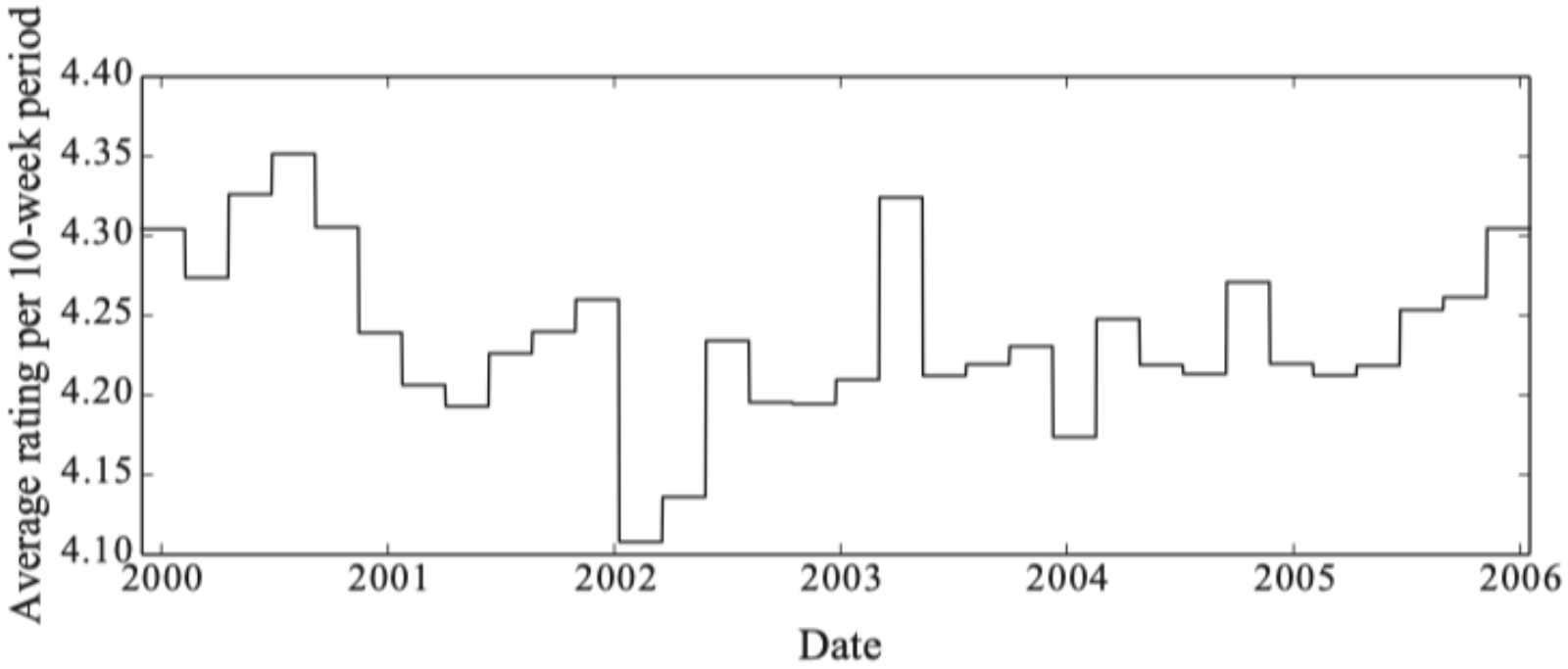
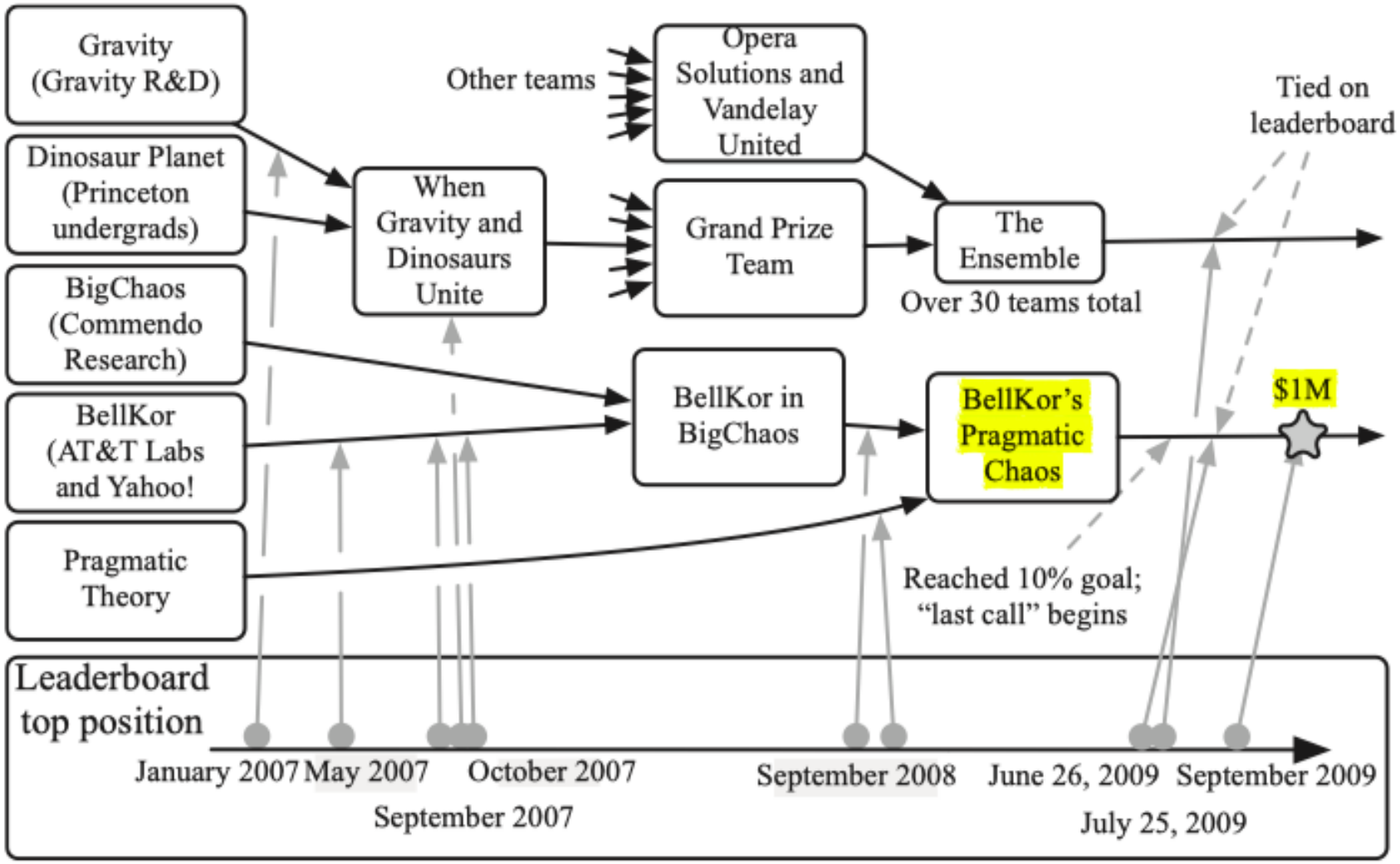


Figure 6.1
Popularity of the movie *The Matrix* over time.

Ensembles of teams: the Netflix prize winners

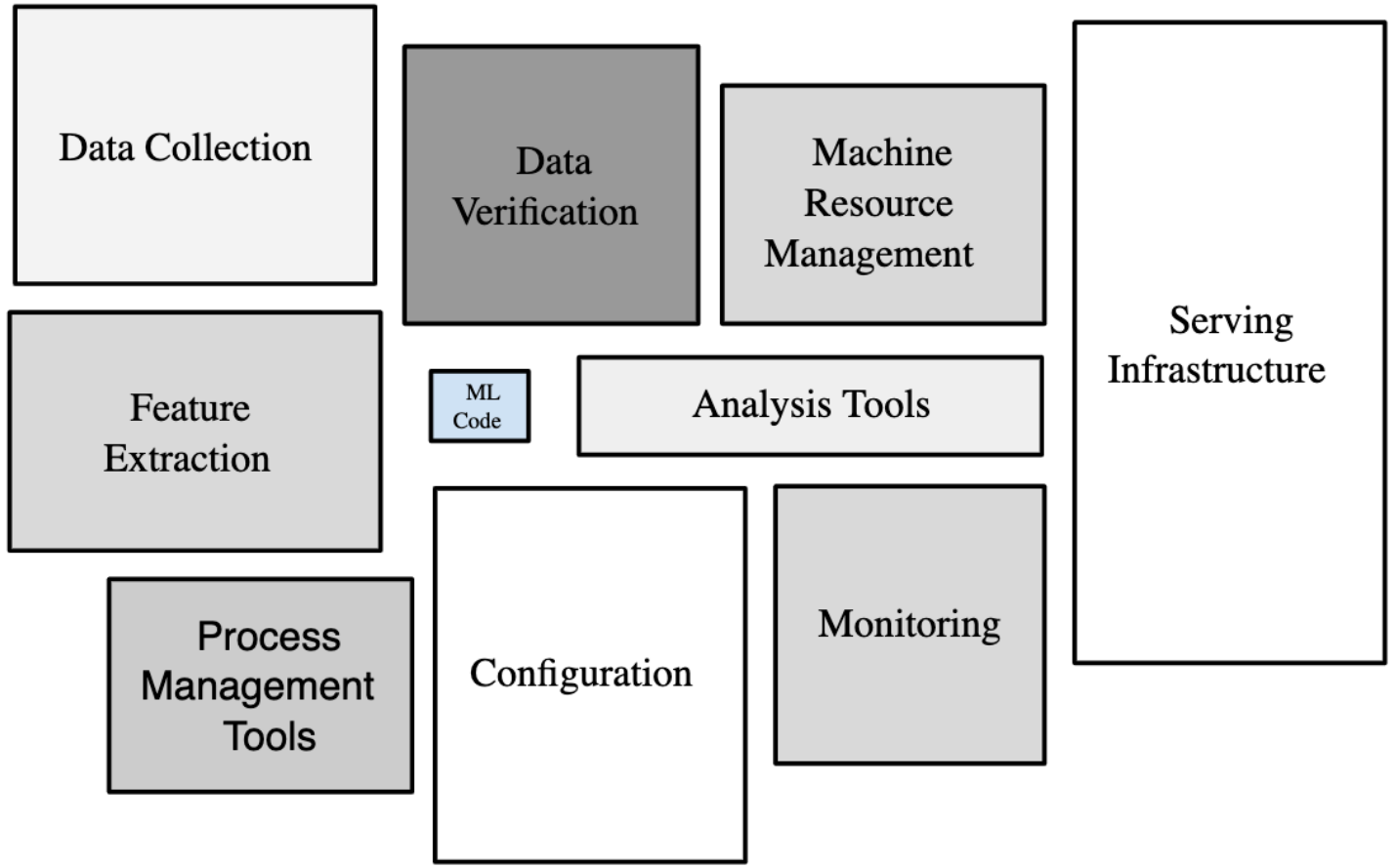


Netflix prize: lessons

- Alliances among teams became essential to success
- It paid off (for Netflix) – many man-hours at reduced cost
- Scientific advancement – and the art of reporting “what didn’t work”
- Timing is everything
- Winning the challenge is just the beginning… and it might not mean much beyond collecting the prize and bragging about it.

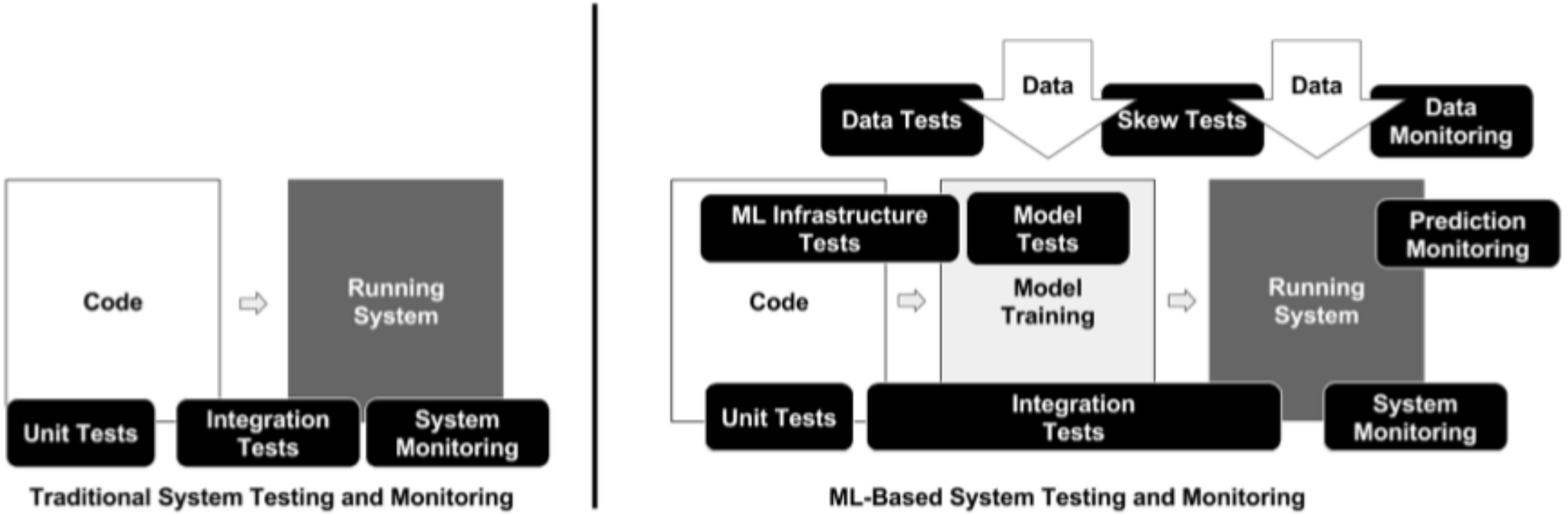
Real-world production ML system

"The ML code is at the heart of a real-world ML production system, but that box often represents only 5% or less of the overall code of that total ML production system."

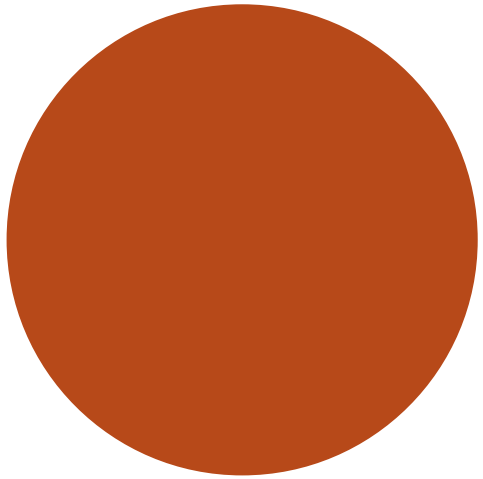
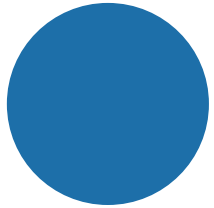


Source: https://developers.google.com/machine-learning/crash-course/production-ml-systems

ML Systems Require Extensive Testing and Monitoring



Source: https://research.google/pubs/pub46555/

Ongoing AI challenges

General AI Challenge by GoodAI

https://www.general-ai-challenge.org/

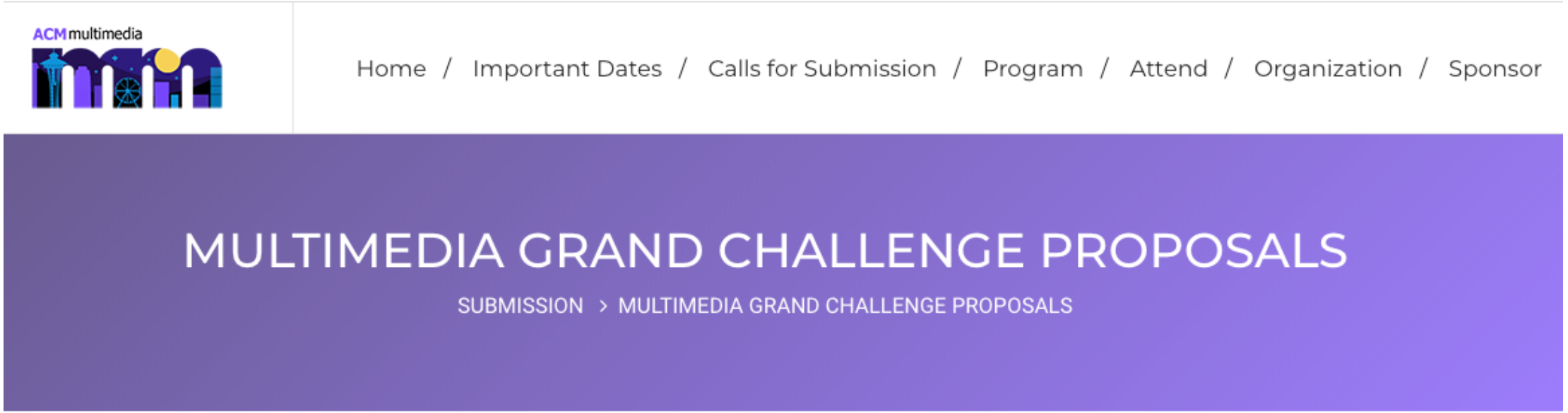
IBM Watson AI XPRIZE

https://www.xprize.org/prizes/artificial-intelligence

Examples of other challenges

Multimedia Community

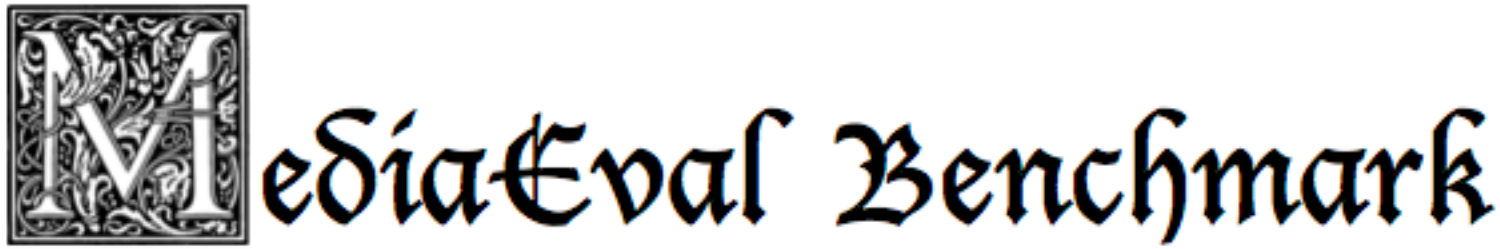
ACM Multimedia Grand Challenges



ACM Multimedia Grand Challenges

- Last year (2019)
  (Complete list at https://2019.acmmm.org/multimedia-grand-challenges/)
  - https://challenge2019.perfectcorp.com/index.html (Beauty)
  - https://github.com/kelkalot/biomedia-2019 (Medical)
  - http://challenge.ai.iqiyi.com/detail?raceId=5c767dc41a6fa0ccf53922e7 (Celebrity)
  - https://www.aitrans.online/MMGC/ (Live video streaming)
  - http://smp-challenge.com/ (Social media)
  - https://github.com/cbvrp-acmmm-2019/cbvrp-acmmm-2019 (Video recommendation – Hulu)

MediaEval



http://www.multimediaeval.org/

TREC Video Retrieval Evaluation: TRECVID

https://trecvid.nist.gov/

Computer Vision Community

Middlebury Computer Vision

http://vision.middlebury.edu/

Pascal VOC

PASCAL2
Pattern Analysis, Statistical Modelling and Computational Learning

http://host.robots.ox.ac.uk/pascal/VOC/

Int J Comput Vis (2015) 111:98–136
DOI 10.1007/s11263-014-0733-5

The PASCAL Visual Object Classes Challenge: A Retrospective

Mark Everingham · S. M. Ali Eslami · Luc Van Gool · Christopher K. I. Williams · John Winn · Andrew Zisserman

Available on Canvas

Microsoft COCO http://cocodataset.org/

What is COCO?

COCO is a large-scale object detection, segmentation, and captioning dataset. COCO has several features:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 330K images (>200K labeled)
- 1.5 million object instances
- 80 object categories
- 91 stuff categories
- 5 captions per image
- 250,000 people with keypoints

Microsoft COCO: Common Objects in Context

Tsung-Yi Lin Michael Maire Serge Belongie Lubomir Bourdev Ross Girshick
James Hays Pietro Perona Deva Ramanan C. Lawrence Zitnick Piotr Dollár

Abstract—We present a new dataset with the goal of advancing the state-of-the-art in object recognition by placing the question of object recognition in the context of the broader question of scene understanding. This is achieved by gathering images of complex everyday scenes containing common objects in their natural context. Objects are labeled using per-instance segmentations to aid in precise object localization. Our dataset contains photos of 91 objects types that would be easily recognizable by a 4 year old. With a total of 2.5 million labeled instances in 328k images, the creation of our dataset drew upon extensive crowd worker involvement via novel user interfaces for category detection, instance spotting and instance segmentation. We present a detailed statistical analysis of the dataset in comparison to PASCAL, ImageNet, and SUN. Finally, we provide baseline performance analysis for bounding box and segmentation detection results using a Deformable Parts Model.

1 Introduction

One of the primary goals of computer vision is the understanding of visual scenes. Scene understanding involves numerous tasks including recognizing what objects are present, localizing the objects in 2D and 3D, determining the objects' and scene's attributes, characterizing relationships between objects and providing a semantic description of the scene. The current object classification and detection datasets [1], [2], [3], [4] help us explore the first challenges related to scene understanding. For instance the ImageNet dataset [1], which contains an unprecedented number of images, has recently enabled breakthroughs in both object classification and detection research [5], [6], [7]. The community has also created datasets containing object attributes [8], scene attributes [9], keypoints [10], and 3D scene information [11]. This leads us to the obvious question: what datasets will best continue our advance towards our ultimate goal of scene understanding?

We introduce a new large-scale dataset that addresses three core research problems in scene understanding: de-

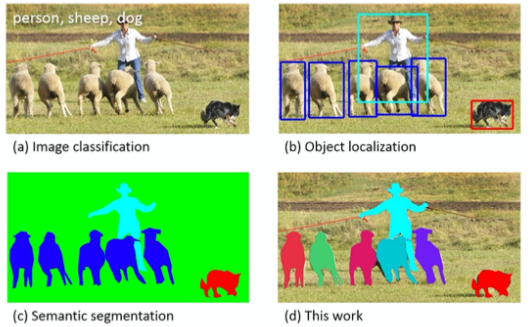
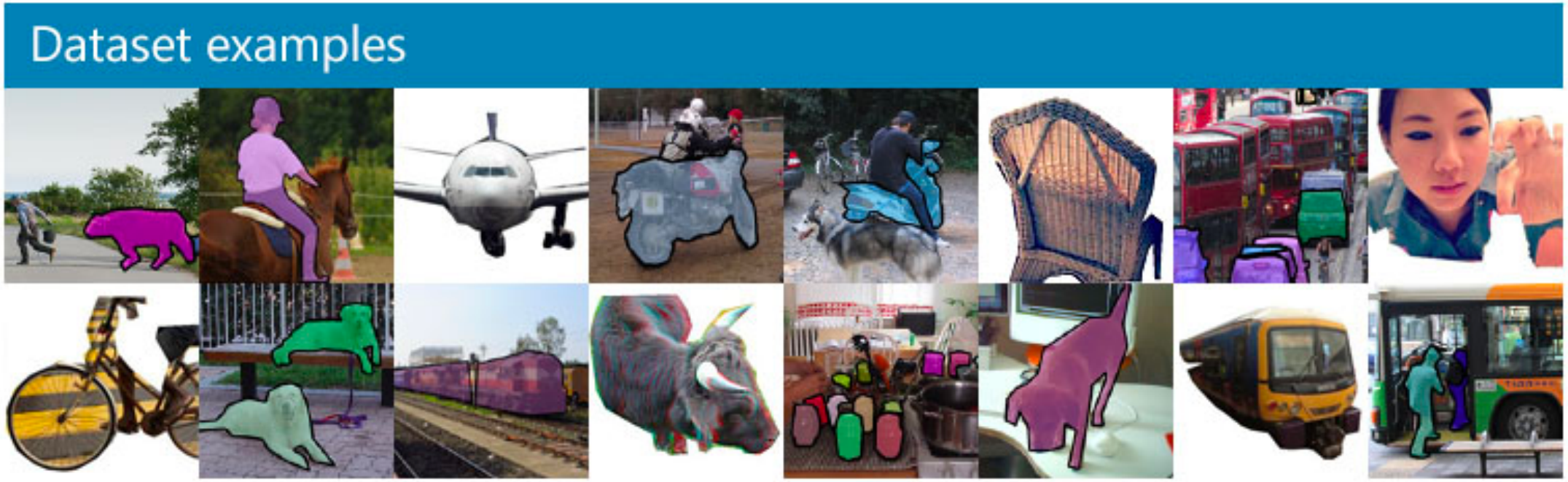


Fig. 1: While previous object recognition datasets have focused on (a) image classification, (b) object bounding box localization or (c) semantic pixel-level segmentation, we focus on (d) segmenting individual object instances. We introduce a large, richly-annotated dataset comprised of images depicting complex everyday scenes of common objects in their natural context.

Available on Canvas

Microsoft COCO http://cocodataset.org/



Caltech 101 and Caltech 256

- http://www.vision.caltech.edu/Image_Datasets/Caltech101/
- http://www.vision.caltech.edu/Image_Datasets/Caltech256/