DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

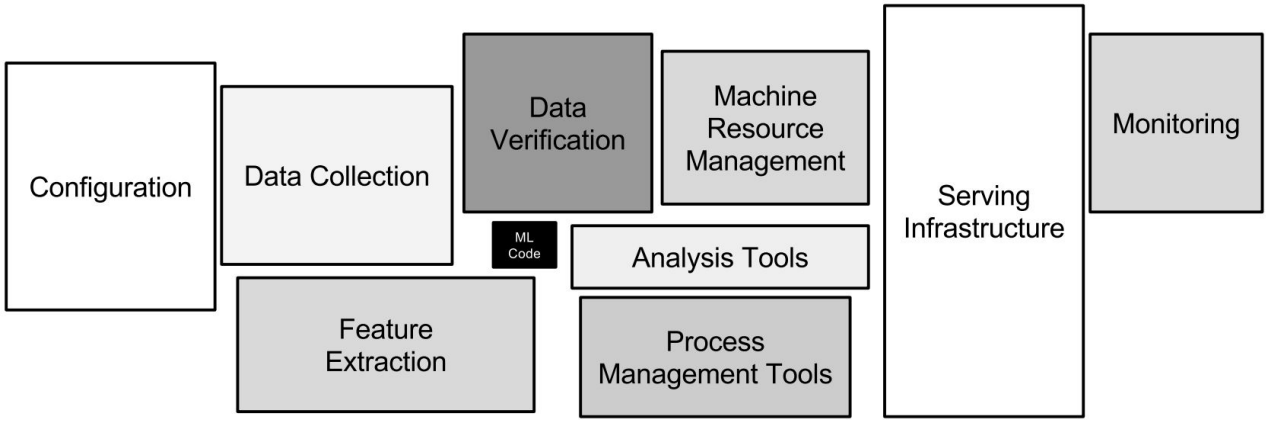
Importance of Testing and Monitoring of AI Models

Woojin Kim, MD

Disclosure

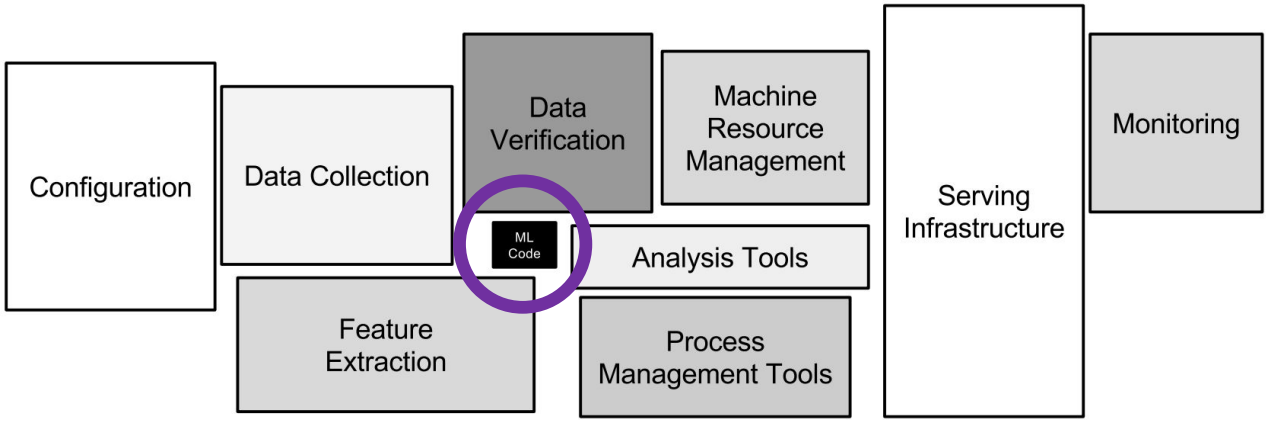
- Independent Contractor Consultant, Nuance Communications
- Advisory Board, Braid Health
- Advisory Board, Inference Analytics
- Advisory Board, Luxonic

DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY



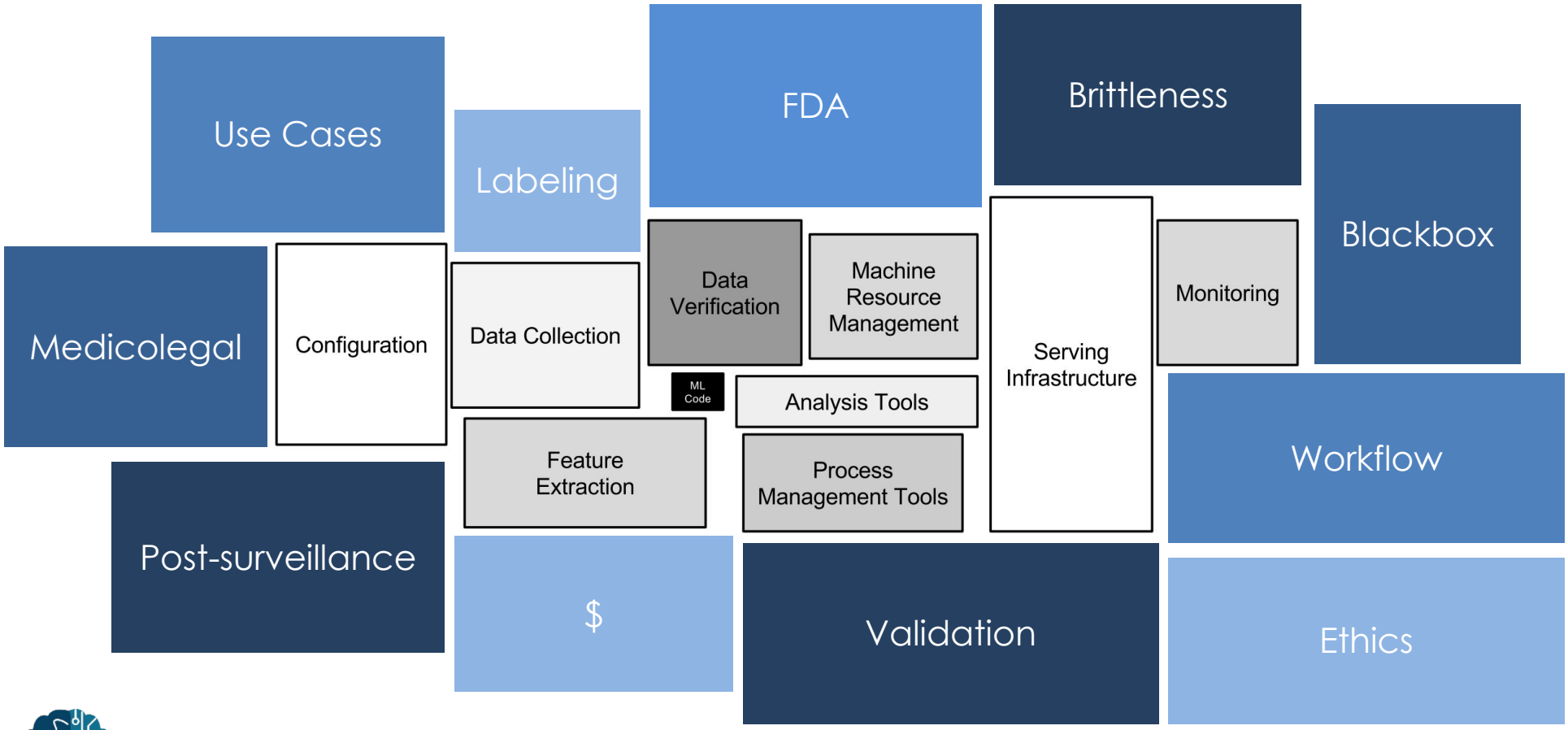
DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

http://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf



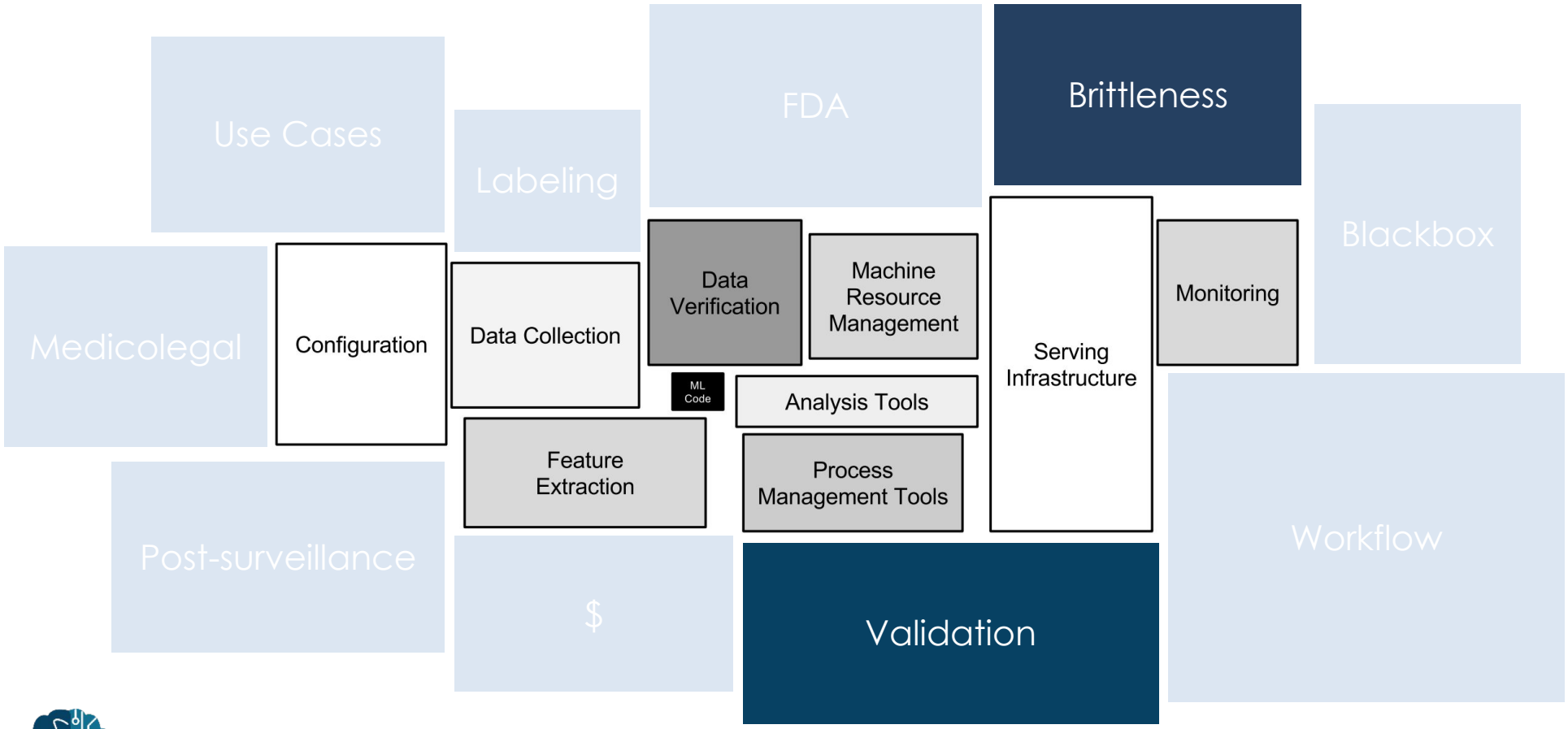
DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

http://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf



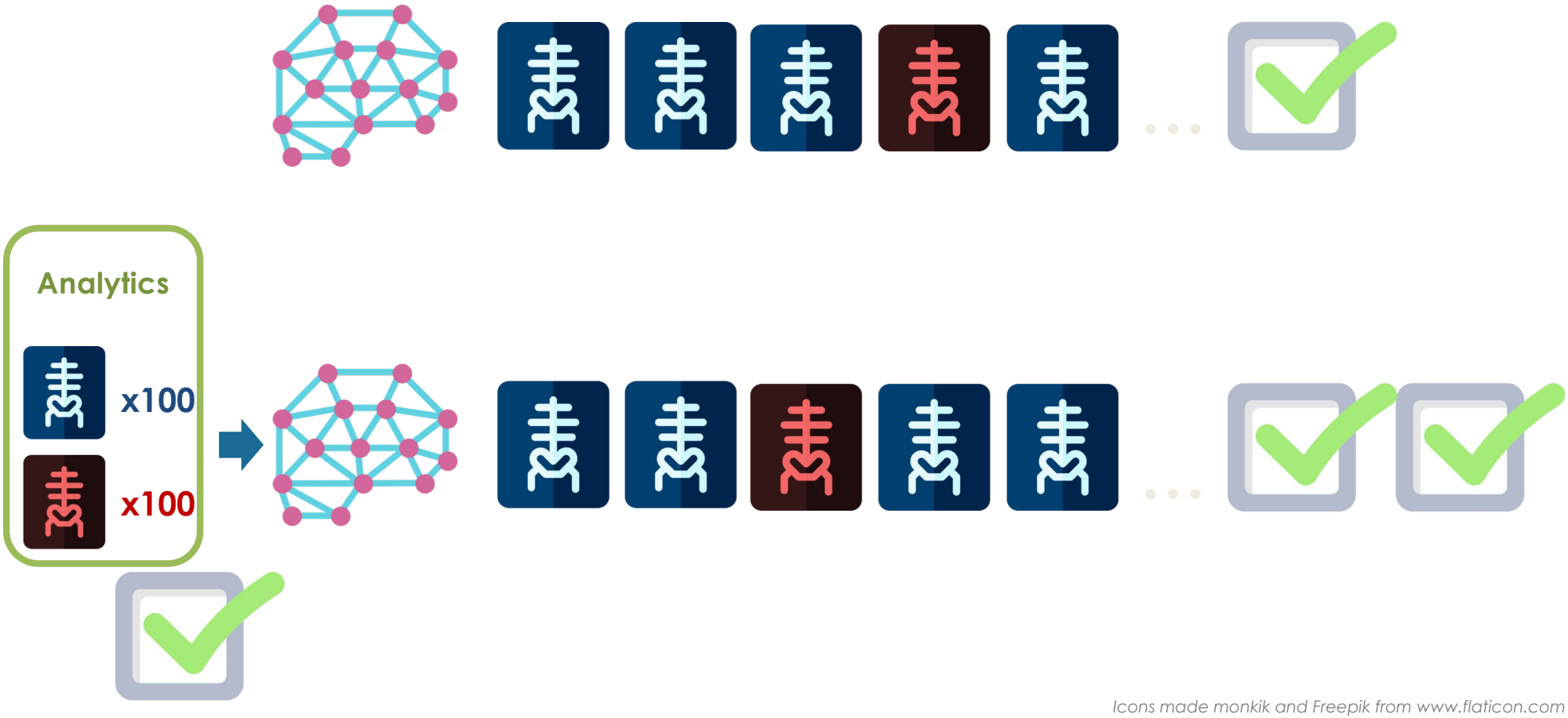
DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

http://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf

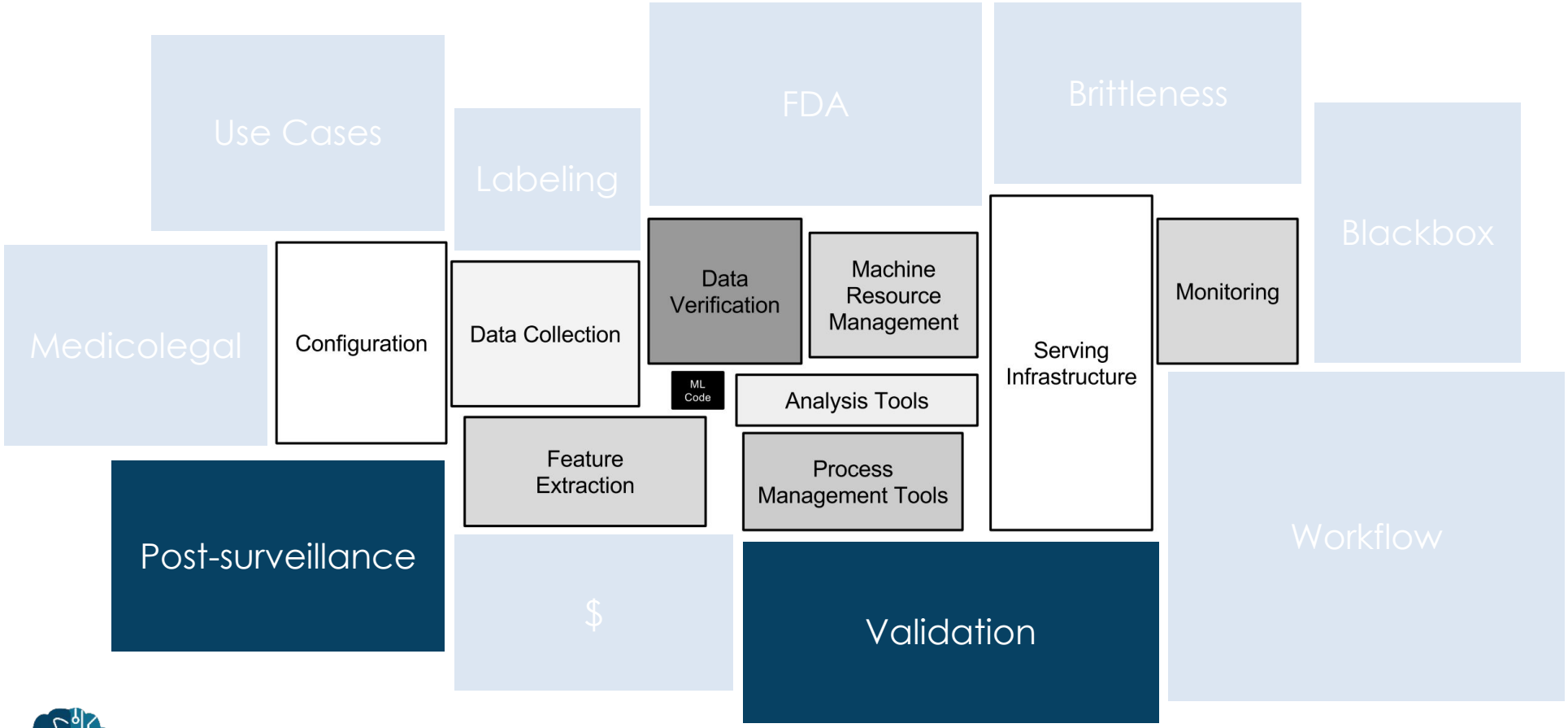


DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

http://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf



Icons made monkik and Freepik from www.flaticon.com

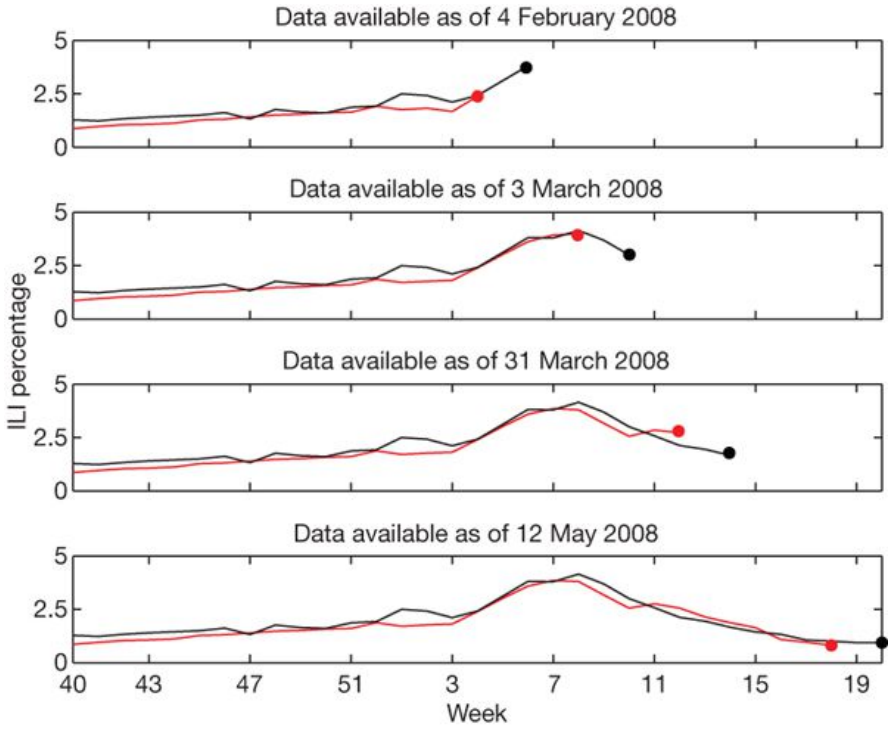


DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

http://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf

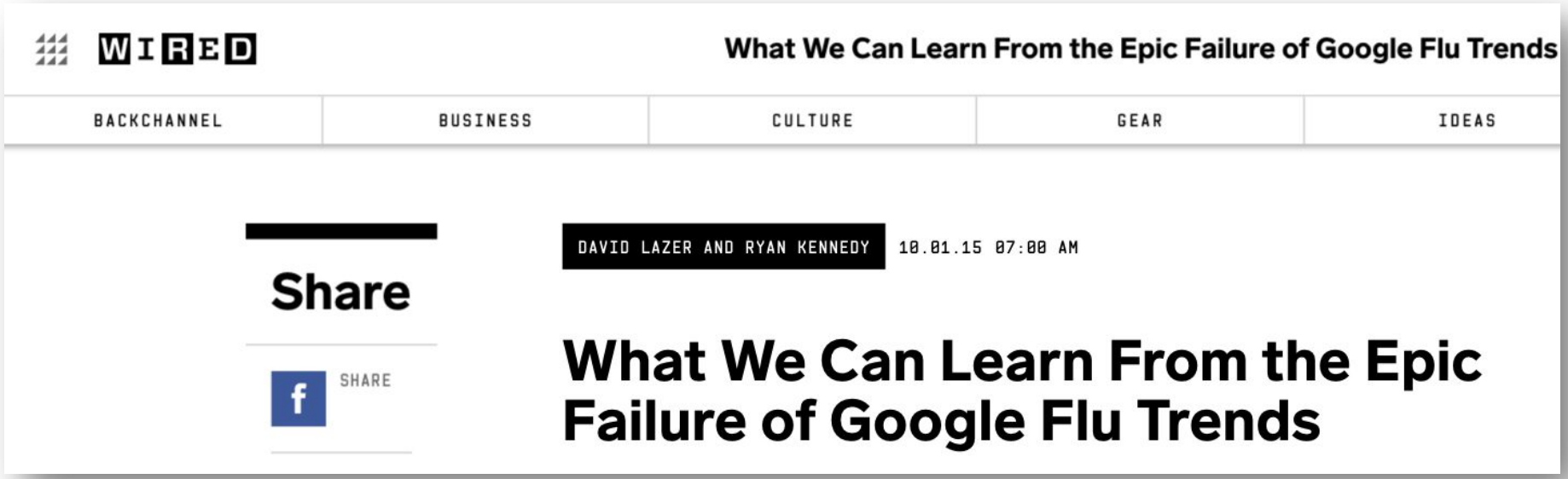
Photo by Jay Heike on Unsplash

ILI percentages estimated by our model (black) and provided by the CDC (red) in the mid-Atlantic region, showing data available at four points in the 2007-2008 influenza season.



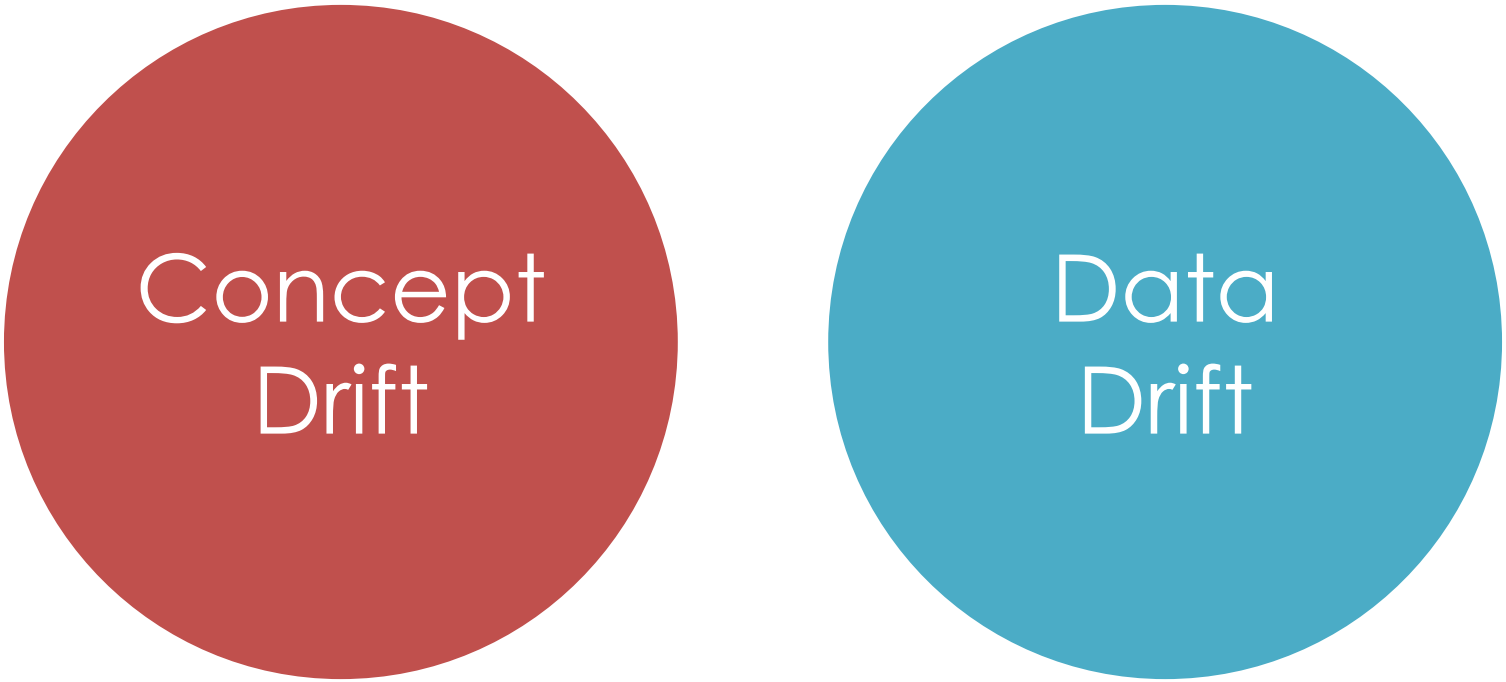
J Ginsberg *et al. Nature* **000**, 1-3 (2008) doi:10.1038/nature07634

nature



DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends/


Concept
Drift
Data
Drift


Concept
Drift

Data
Drift


Concept
Drift
Data
Drift

Forbes Billionaires Innovation Leadership Money Business Small Business Lifestyle Lis

19,668 views | Apr 3, 2019, 05:00am EDT

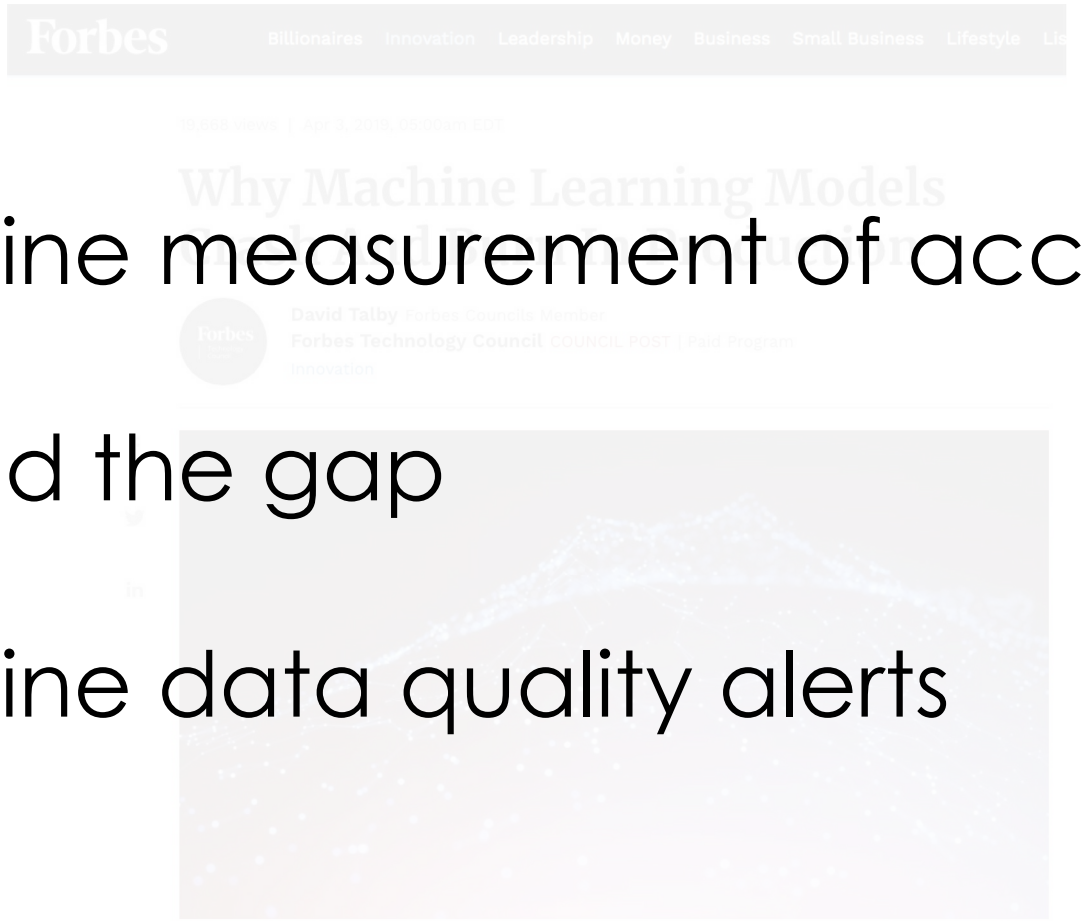
Why Machine Learning Models Crash And Burn In Production

David Talby Forbes Councils Member
Forbes Technology Council COUNCIL POST | Paid Program
Innovation

DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

https://www.forbes.com/sites/forbestechcouncil/2019/04/03/why-machine-learning-models-crash-and-burn-in-production/#716e5ac92f43

1. Online measurement of accuracy
2. Mind the gap
3. Online data quality alerts



DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

https://www.forbes.com/sites/forbestechcouncil/2019/04/03/why-machine-learning-models-crash-and-burn-in-production/#716e5ac92f43

ORIGINAL ARTICLE

Artificial Intelligence in Imaging: The Radiologist's Role

Daniel L. Rubin, MD, MS[a]

Abstract

Rapid technological advancements in artificial intelligence (AI) methods have fueled explosive growth in decision tools being marketed by a rapidly growing number of companies. AI developments are being driven largely by computer scientists, informaticians, engineers, and businesspeople, with much less direct participation by radiologists. Participation by radiologists in AI is largely restricted to educational efforts to familiarize them with the tools and promising results, but techniques to help them decide which AI tools should be used in their practices and to how to quantify their value are not being addressed. This article focuses on the role of radiologists in imaging AI and suggests specific ways they can be engaged by (1) considering the clinical need for AI tools in specific clinical use cases, (2) undertaking formal evaluation of AI tools they are considering adopting in their practices, and (3) maintaining their expertise and guarding against the pitfalls of overreliance on technology.

Key Words: Artificial intelligence, radiology, imaging, evaluation

J Am Coll Radiol 2019;16:1309-1317.

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

ORIGINAL ARTICLE

Check for updates

Artificial Intelligence in Imaging: The Radiologist's Role

Daniel L. Rubin, MD, MS[a]

> The ability to quickly produce AI tools is outpacing the thoroughness with which they are being evaluated or validated on independent data in new settings.

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

ORIGINAL ARTICLE

Check for updates

Artificial Intelligence in Imaging: The Radiologist's Role

Daniel L. Rubin, MD, MS[a]

Some radiologists may be tempted to assume that such data collection and evaluation of metrics are not needed, presuming that the AI will work well for their patients if the algorithm has been FDA cleared. **This is not necessarily a good assumption** because the data sets used for FDA clearance may not be sufficiently representative of every radiology practice.

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

Steps for radiologists in evaluating AI algorithms

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

Steps for radiologists in evaluating AI algorithms

1. Understand AI outputs and clinical relevancy

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

Steps for radiologists in evaluating AI algorithms
1. Understand AI outputs and clinical relevancy
2. Collect test cases

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

Steps for radiologists in evaluating AI algorithms
1. Understand AI outputs and clinical relevancy
2. Collect test cases
3. Establish ground truth

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

Steps for radiologists in evaluating AI algorithms
1. Understand AI outputs and clinical relevancy
2. Collect test cases
3. Establish ground truth
4. Choose appropriate evaluation metrics, like sensitivity, specificity, PPV, etc.

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

Steps for radiologists in evaluating AI algorithms
1. Understand AI outputs and clinical relevancy
2. Collect test cases
3. Establish ground truth
4. Choose appropriate evaluation metrics, like sensitivity, specificity, PPV, etc.
5. Define performance threshold

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

Steps for radiologists in evaluating AI algorithms
1. Understand AI outputs and clinical relevancy
2. Collect test cases
3. Establish ground truth
4. Choose appropriate evaluation metrics, like sensitivity, specificity, PPV, etc.
5. Define performance threshold
6. Evaluate test cases

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

Steps for radiologists in evaluating AI algorithms
1. Understand AI outputs and clinical relevancy
2. Collect test cases
3. Establish ground truth
4. Choose appropriate evaluation metrics, like sensitivity, specificity, PPV, etc.
5. Define performance threshold
6. Evaluate test cases
7. Have a monitoring strategy

https://www.jacr.org/article/S1546-1440(19)30636-2/pdf

Conclusion

Banner vector created by starline - www.freepik.com

Thank you

woojinrad@gmail.com

@woojinrad

www.linkedin.com/woojinkim

DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY