



Implementing Artificially Intelligent Employee Retention

**2019 AFCEA Mid-America Cyberspace Symposium
Killian McKee**

My Background

Company

- Booz Allen Hamilton
 - Consulting firm
 - Federal, defense & technology
 - ~26,000 employees



Experience

- Data Scientist at BAH
- Neural machine translation and computer vision projects
- Currently working @ SAFB
- W&M Master's in analytics
- Background in ML/DL



Questions?

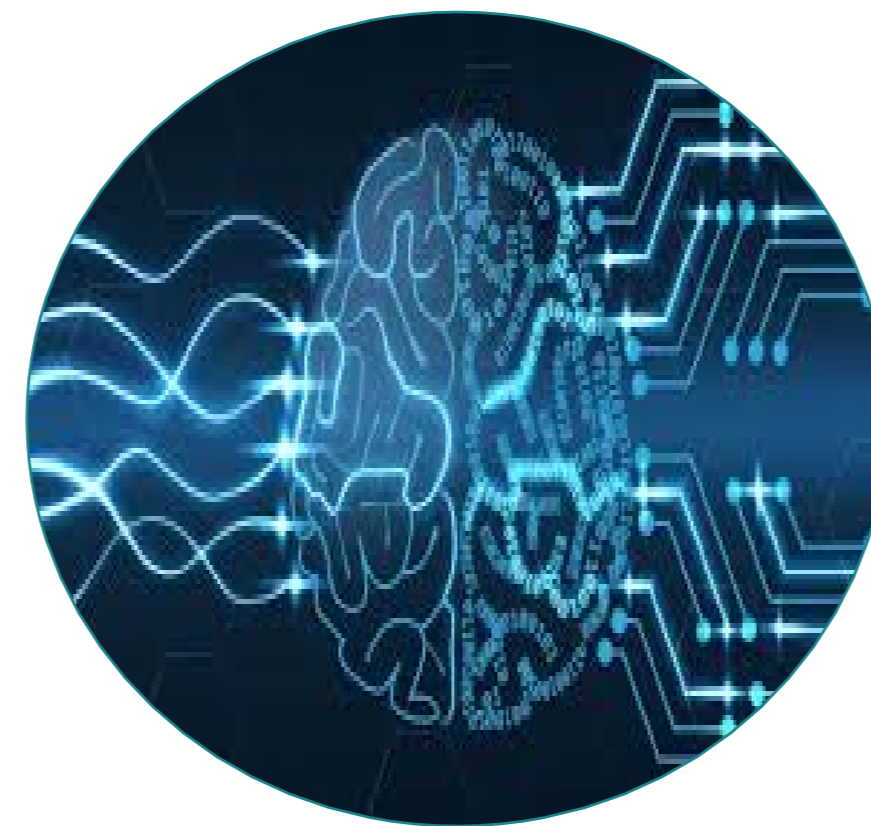
- Feel free to ask questions
- We can also chat after talk
- Other questions?
mckee_killian@bah.com



Our Goal



Understand retention challenges



Understand how AI can reduce attrition



Cover AI implementation challenges + benefits



Cover how to implement an AI retention system

Our Roadmap

Attrition

- What is attrition?
- Why is it important?
- Major causes

AI

- AI level setting
- What is ML?
- How can AI help?

Implementation

- Overcoming obstacles
- DIY
- Process flow
- Example implementation

Conclusion

- Review
- ?'s
- Sources

Attrition

AI

Implementation

Conclusion

What is Employee Attrition?

- Attrition=voluntary staff departures
 - Normal life events
 - Health
 - Retirement
 - Dissatisfaction w/ company
- Does not include employee terminations
- Retention=keeping valued employees @ your company (willingly)



Attrition Significance

- Expensive
 - Center for American Progress estimates:
 - 16% of salary for low pay/high turnover (<\$30k/year)
 - 20% for midrange (\$30-50k)
 - 50-100% for highly educated or technical positions
- High employer priority
 - 87% of employers rate improving retention critical priority (*Forbes* 2017 study)
- Can signal underlying company problems
 - Toxic employees
 - Workplace shortcomings
 - Poor management



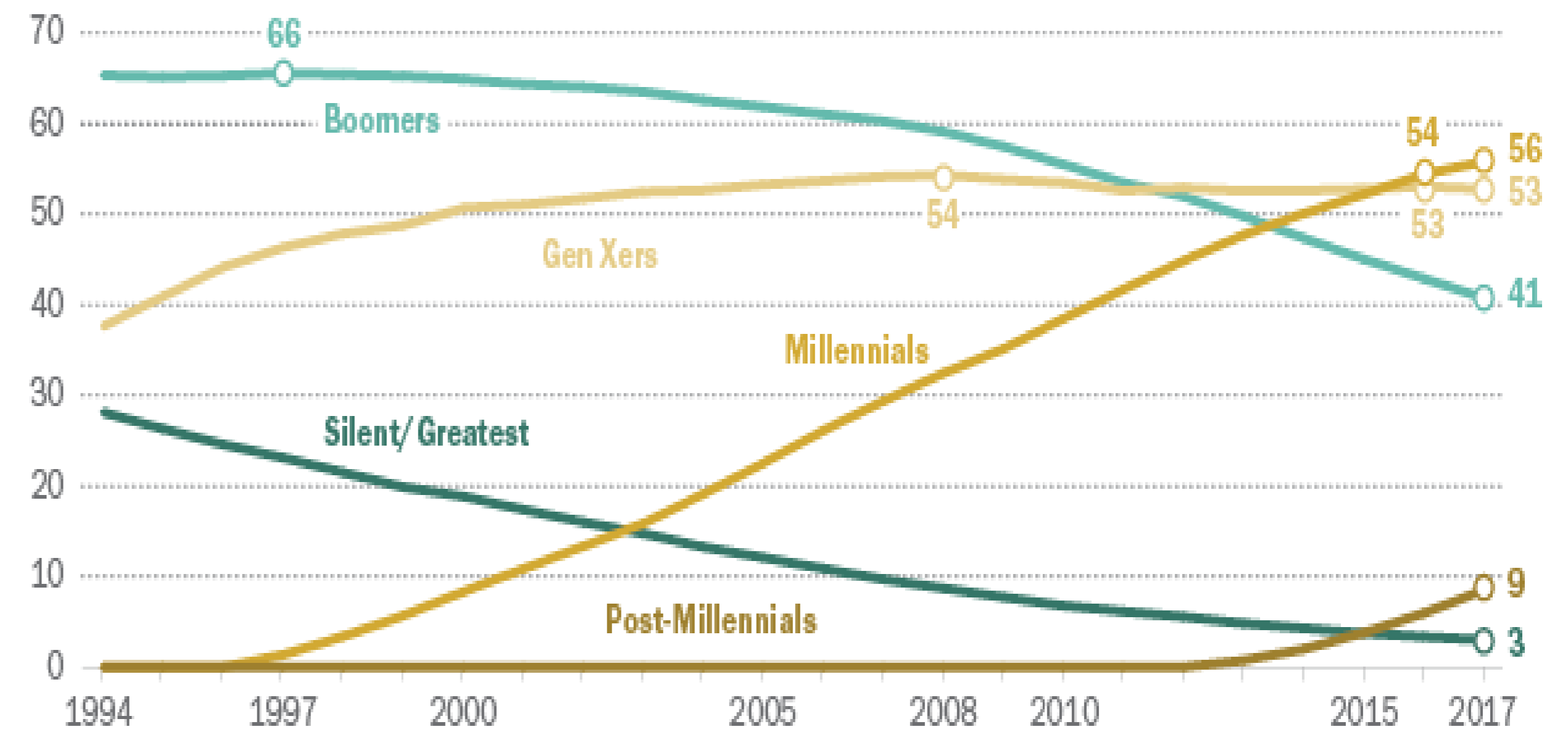
<https://ehorus.com/employee-retention/>

Attrition Causes: Shifting Workforce

- Millennials are now largest U.S. working cohort (35%)
 - Young
 - Willing to uproot
 - Changing values
 - Less loyalty amongst companies + employees
 - More lucrative to switch jobs than ask for a raise
- Post Millennials Increasing
 - Will likely bring similar changes

Millennials became the largest generation in the labor force in 2016

U.S. labor force, in millions



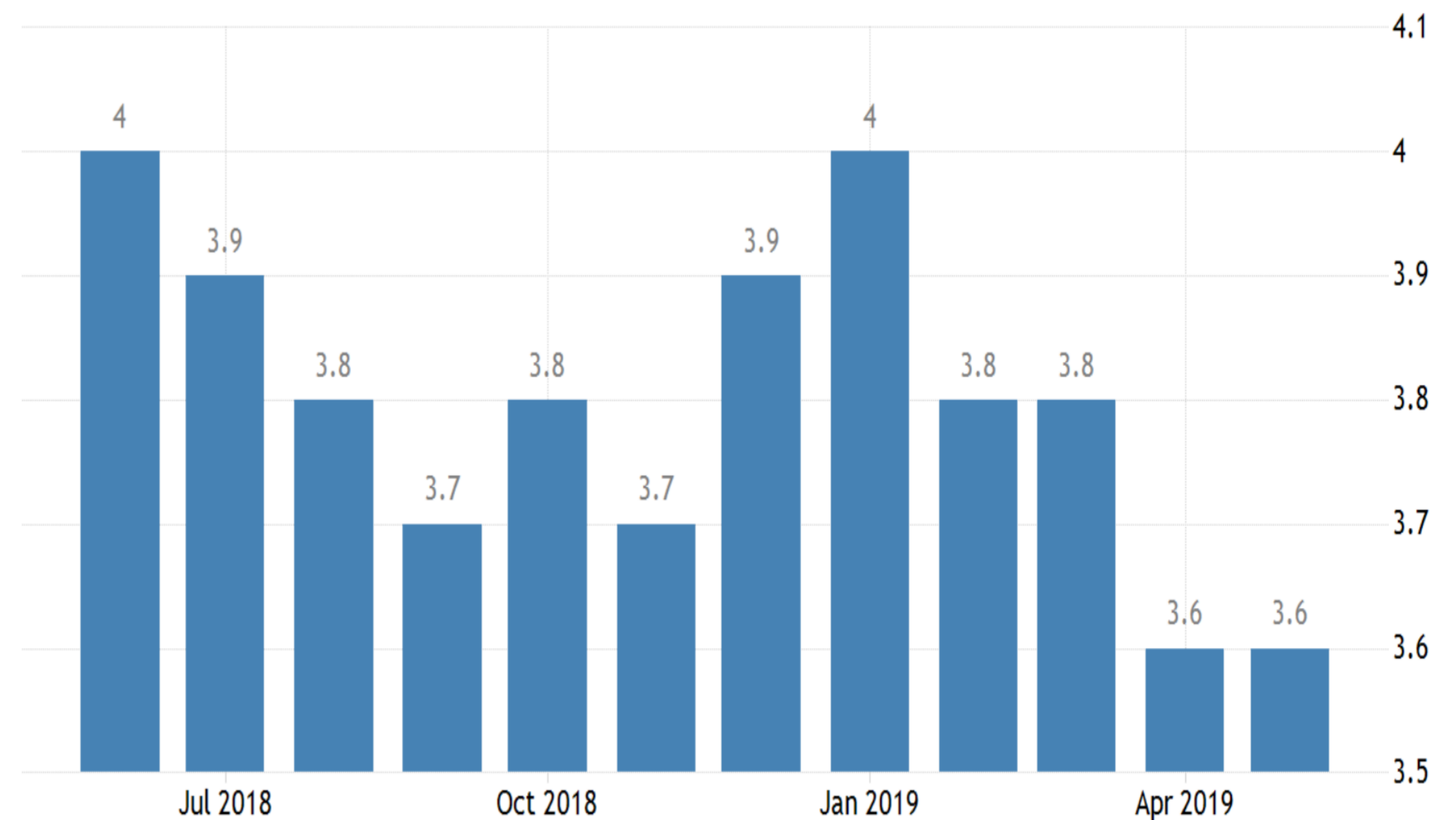
Note: Labor force includes those ages 16 and older who are working or looking for work. Annual averages shown.
Source: Pew Research Center analysis of monthly 1994-2017 Current Population Survey (IPUMS).

PEW RESEARCH CENTER

Attrition Causes: Strong Job Market

- U.S. unemployment is low
 - ~3.6%
- Many industries openings>jobs
 - High skilled medical
 - Transportation
 - Engineering
 - IT/comp sci/data science
 - Skilled labor
- Recruiters biggest problem=lack of applicants (Society of Human Resources Management, 2016)

U.S. Unemployment Rate 2018-2019



SOURCE: TRADINGECONOMICS.COM | U.S. BUREAU OF LABOR STATISTICS

Attrition Causes: Technological Advances

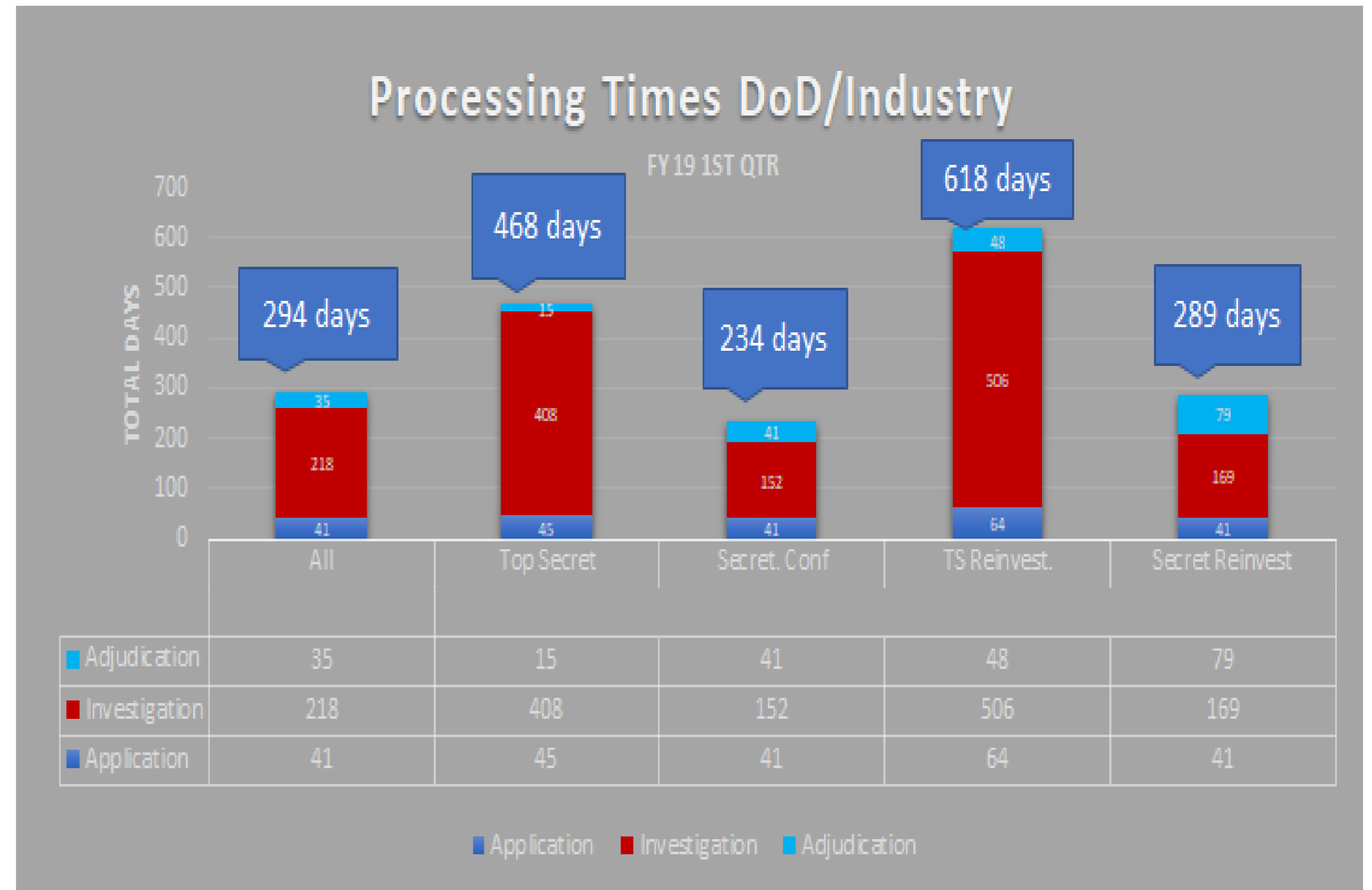
- Technology has increased job + salary info availability
 - LinkedIn
 - Glassdoor
- Easier to discreetly job hunt
- Machine learning helps competing firms target employees precisely
- Easier for recruiters to find + poach skilled employees

Linked 

glassdoor[®]

Cleared Attrition Causes: Clearance Process

- Security clearance process=glacial
 - Avg secret processing time: 234 days
 - Avg top secret: 468 days (DoD/ClearanceJobs.com, 2019)
- Backlog of ~550,000 cases
 - Demand for cleared employees especially high
 - Poaching very common
- Trusted Workforce 2.0 Initiative
 - ODNI effort to reduce wait times + increase quality



<https://news.clearancejobs.com/2019/03/12/how-long-does-it-take-to-process-a-security-clearance-q1-2019-update>

Attrition

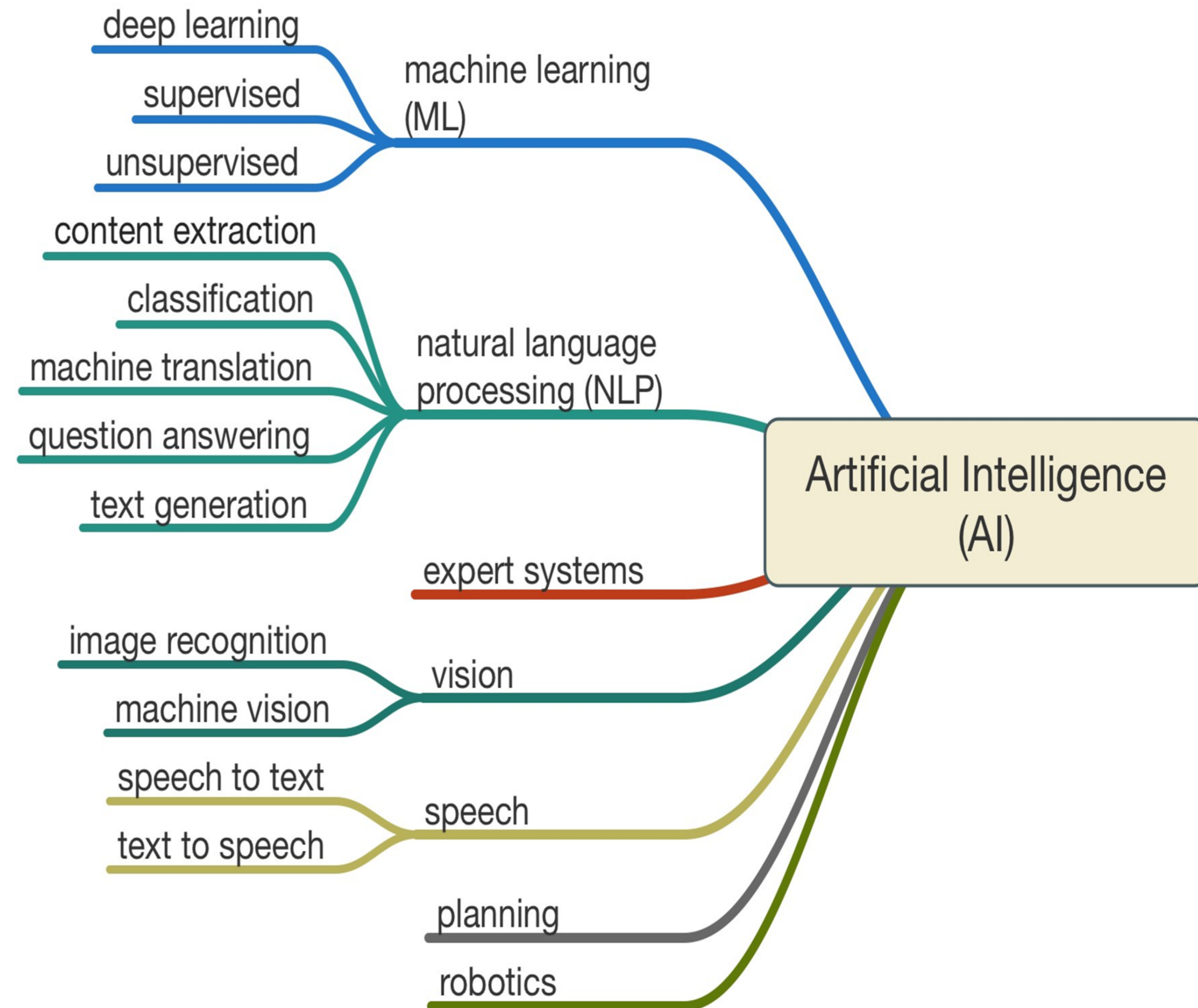
AI

Implementation

Conclusion

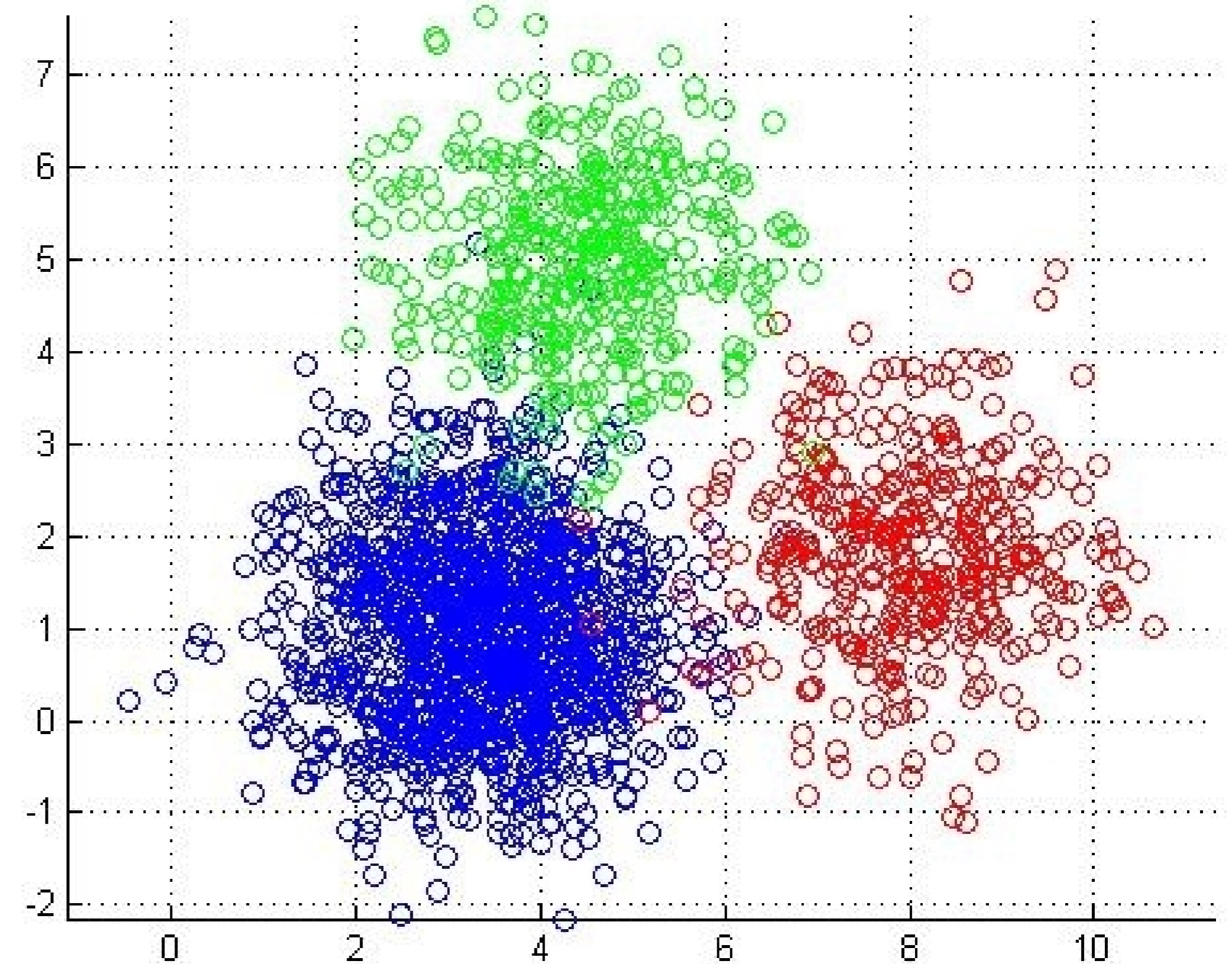
The AI Landscape

- Our AI Definition:
 - Development of computer systems capable of performing tasks requiring human intelligence
 - E.g. visual perception, speech recognition, decision making, decision making, translation, etc.
 - Learning implicitly vs explicitly



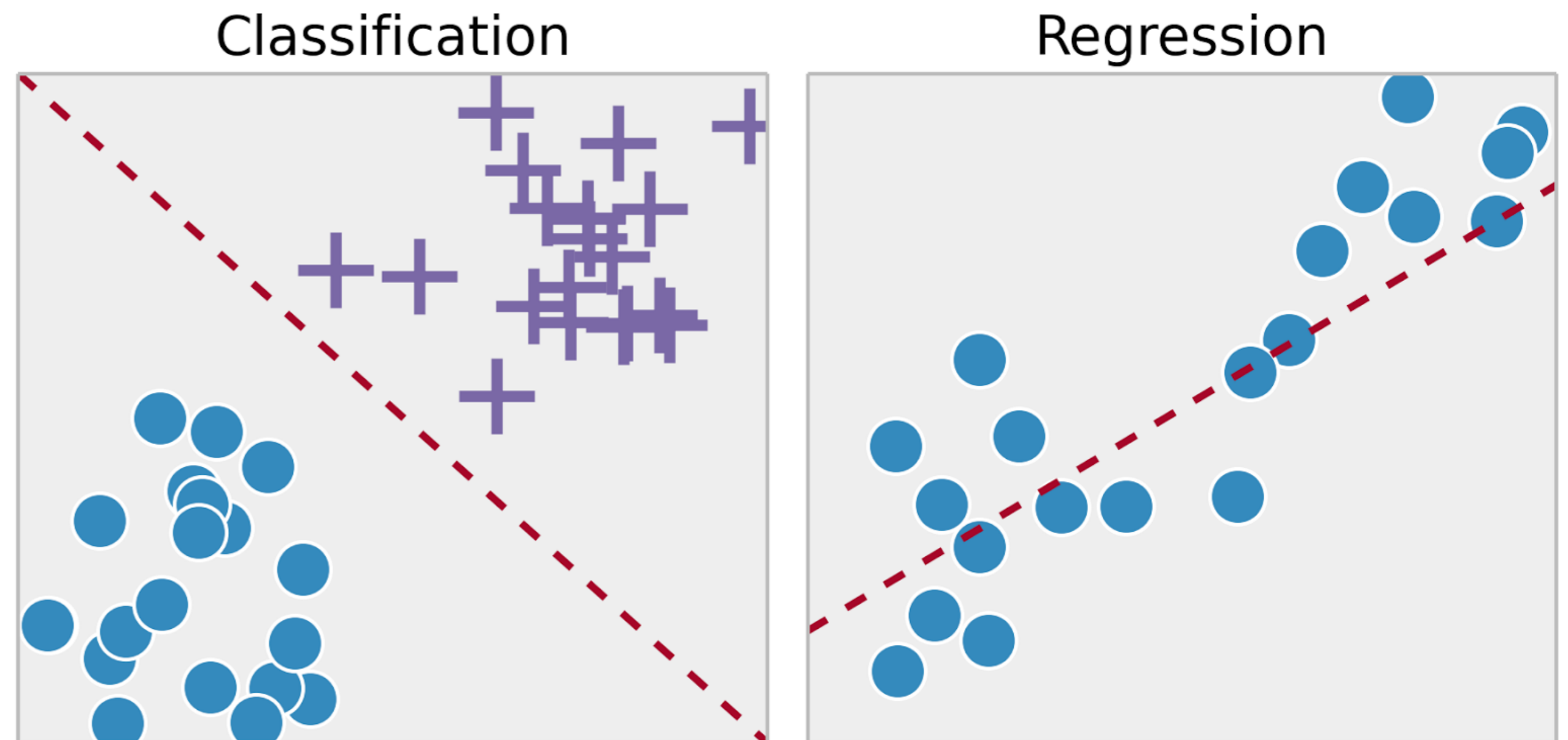
What is Machine Learning?

- Machine learning:
 - Programming a computer to learn implicitly
 - Supervised and unsupervised
- Unsupervised learning
 - Data is passed through an algorithm
 - Algorithm looks for patterns without knowing what it is looking at
 - **Goal:** understand patterns in data
 - Commonly used to group things together (clustering)



What is Supervised Learning?

- Supervised learning
 - Pass desired input and output to algorithm so that it can “learn” from ground truth
 - **Goal:** predict a value or class label
- Two sets of problems for supervised
 - Regression
 - Predicting a specific value
 - Ex. House prices
 - Classification
 - Labeling something categorically
 - Ex. Dog, cat, human



<https://towardsdatascience.com/supervised-vs-unsupervised-learning-14f68e32ea8d>

How can we use ML to Reduce Attrition?

- Attrition: classification problem
 - Employee data labeled as either ‘someone who left the company’ or, ‘someone who didn’t leave the company’
 - Doesn’t just have to be binary labels
 - Ex. ‘someone who left’, ‘at risk’, ‘someone who didn’t leave’
- Once model as been trained, it can be used to generate new predictions
- Predictions can be used head off valued employee attrition

Part of an Example Dataset (IBM)

Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField
41	Yes	Travel_Rarely	1102	Sales	1	2	Life Sciences
49	No	Travel_Frequently	279	Research & Development	8	1	Life Sciences
37	Yes	Travel_Rarely	1373	Research & Development	2	2	Other
33	No	Travel_Frequently	1392	Research & Development	3	4	Life Sciences
27	No	Travel_Rarely	591	Research & Development	2	1	Medical
32	No	Travel_Frequently	1005	Research & Development	2	2	Life Sciences
59	No	Travel_Rarely	1324	Research & Development	3	3	Medical
30	No	Travel_Rarely	1358	Research & Development	24	1	Life Sciences
38	No	Travel_Frequently	216	Research & Development	23	3	Life Sciences
36	No	Travel_Rarely	1299	Research & Development	27	3	Medical
35	No	Travel_Rarely	809	Research & Development	16	3	Medical
29	No	Travel_Rarely	153	Research & Development	15	2	Life Sciences
31	No	Travel_Rarely	670	Research & Development	26	1	Life Sciences
34	No	Travel_Rarely	1346	Research & Development	19	2	Medical
28	Yes	Travel_Rarely	103	Research & Development	24	3	Life Sciences
29	No	Travel_Rarely	1389	Research & Development	21	4	Life Sciences
32	No	Travel_Rarely	334	Research & Development	5	2	Life Sciences
22	No	Non-Travel	1123	Research & Development	16	2	Medical
53	No	Travel_Rarely	1219	Sales	2	4	Life Sciences
38	No	Travel_Rarely	371	Research & Development	2	3	Life Sciences
24	No	Non-Travel	673	Research & Development	11	2	Other
36	Yes	Travel_Rarely	1218	Sales	9	4	Life Sciences
34	No	Travel_Rarely	419	Research & Development	7	4	Life Sciences
21	No	Travel_Rarely	391	Research & Development	15	2	Life Sciences
34	Yes	Travel_Rarely	699	Research & Development	6	1	Medical
53	No	Travel_Rarely	1282	Research & Development	5	3	Other


```
graph LR; A[Attrition] --> B[AI]; B --> C[Implementation]; C --> D[Conclusion];
```

Attrition

AI

Implementation

Conclusion

Implementation Benefits



Cost Savings



Straightforward Technical Implementation



Identify Attrition Sources



Happier Employees



Preserve Strongly Performing Teams

Implementation Obstacles & Potential Solution



Privacy Concerns

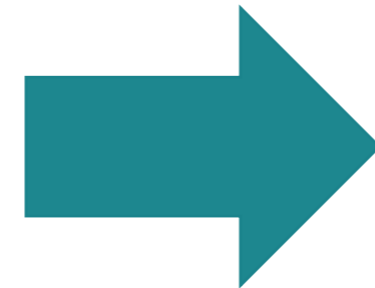
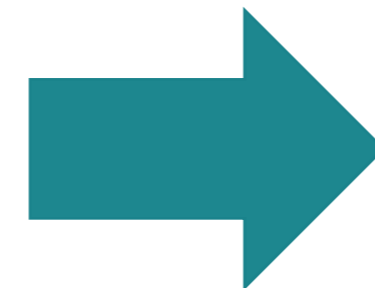
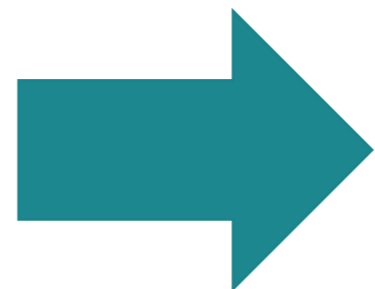
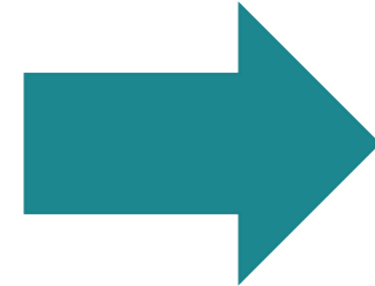
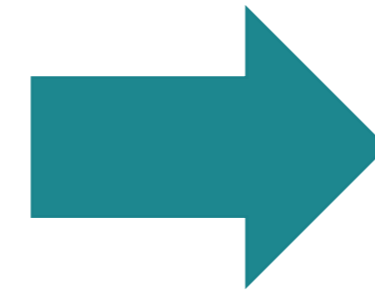
Distrust of AI

Resistance to Change

Overestimating Implementation
Difficulty

Lack of Data

Poor Data Governance



Can use common HR info

Transparent process, allow opt out

Clarify benefits

Open source options available

Can implement upsampling

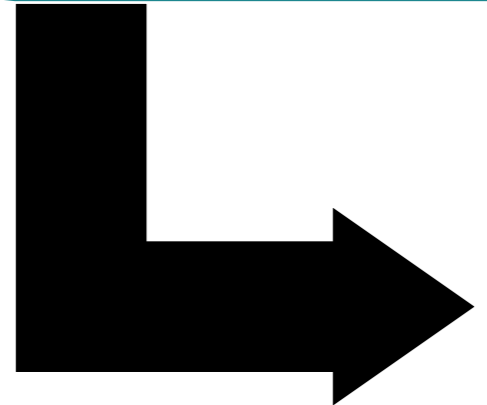
Consolidating data sources, DMBOK

Attrition Classifier: DIY



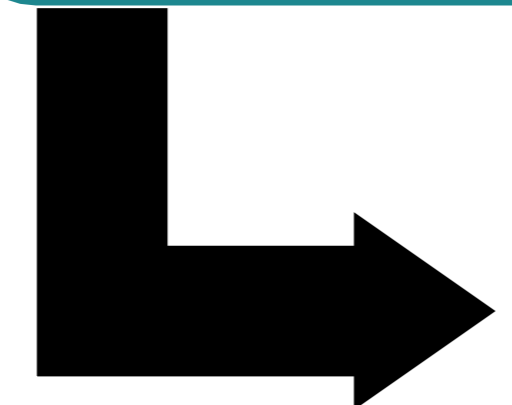
Get Data

- Get employee data like # years at company, salary, etc.
- HR + finance info great starting point
- Label training data



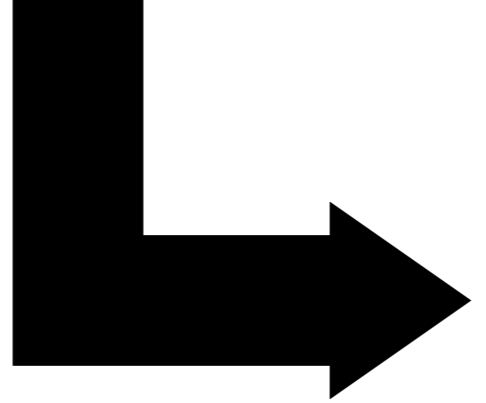
Build Model

- Python (or similar language)
- Tensorflow (for deep learning)
- AWS Sagemaker or Lambda for Cloud



Train Model

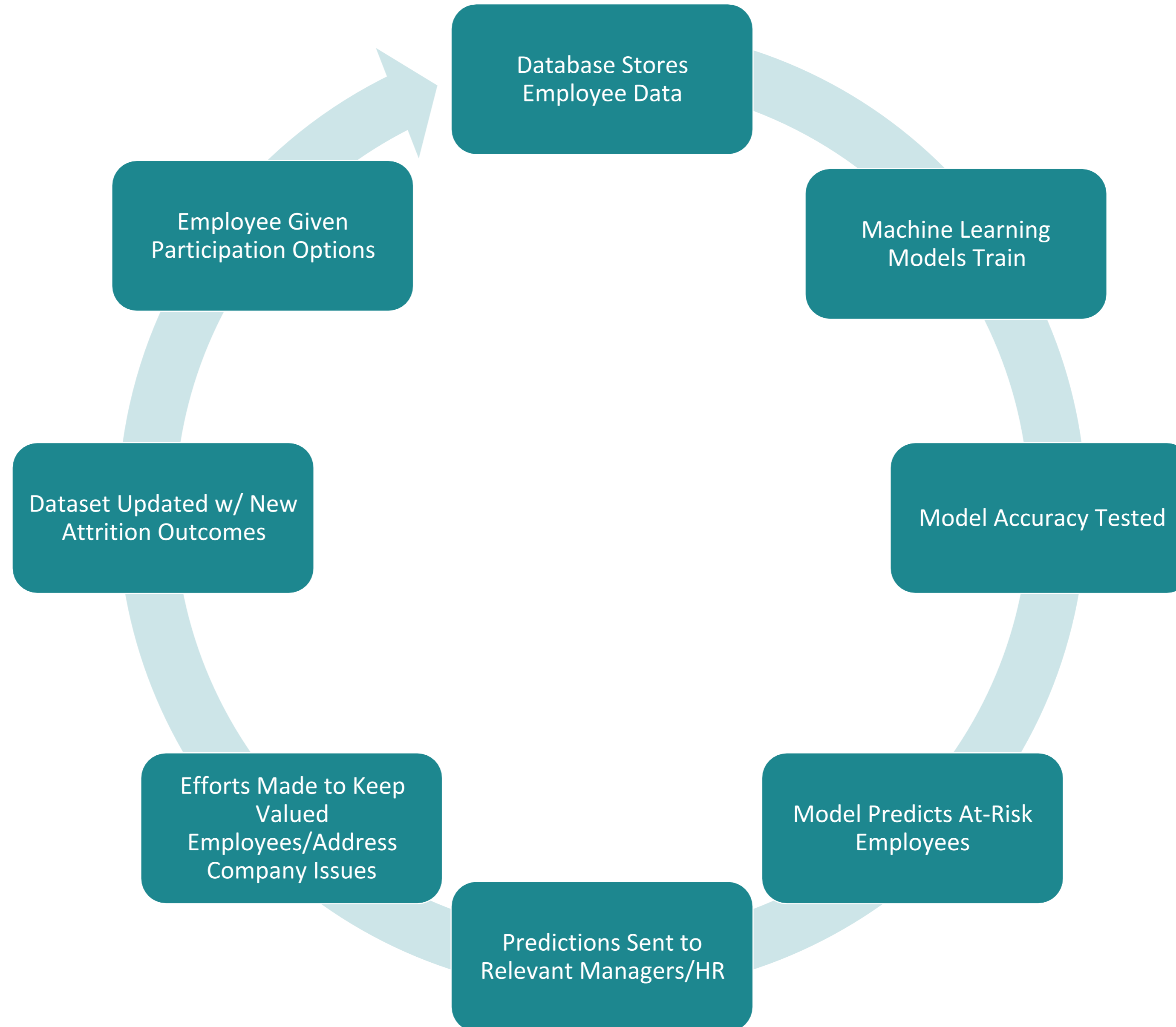
- Train locally on CPU or GPU, or in cloud with AWS instance



Deploy

- Local application/script
- AWS

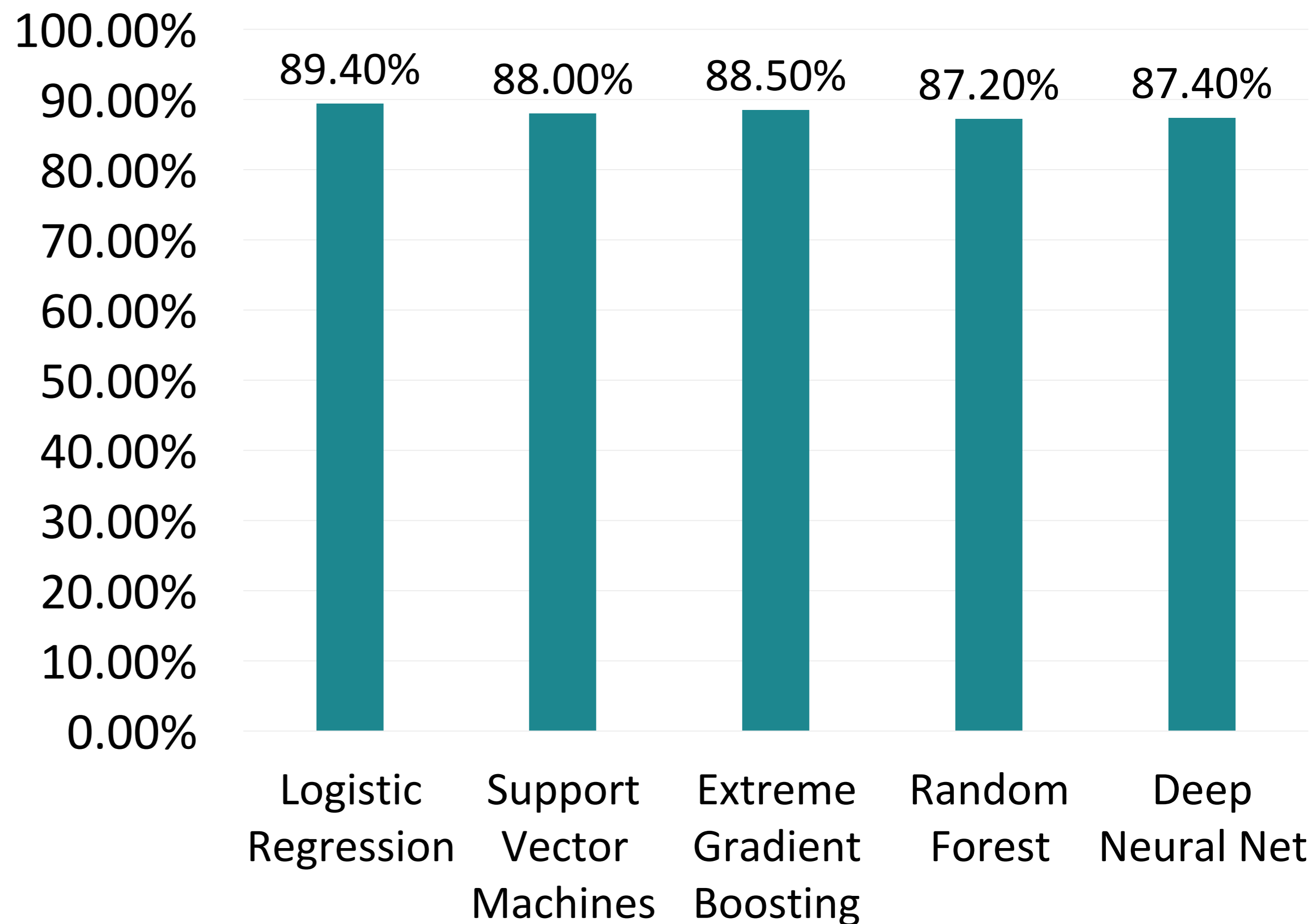
An Example Process Flow



An Example Implementation

- IBM Attrition Dataset
 - ~1500 employees
 - ~80% did not commit attrition
 - ~20% left company
- Tested 5 different algorithms
- Compared Model Accuracies
 - Log Reg highest at 89.4%
 - 9.4% more accurate than guessing
- You can download my code here::
 - https://github.com/kcmckee/AFCEA_2019

Example Classifier Model Accuracies



Attrition

AI

Implementation

Conclusion



Review

- **Attrition**
 - Expensive
 - Caused by external and internal factors
- **AI**
 - General landscape
 - What is ML?
 - Attrition can be treated as classification problem
- **Implementation**
 - Benefits
 - Obstacles/Solutions
 - Process Tools & Steps
 - Example Model



Connect

- Email: mckee_killian@bah.com
- LinkedIn: Killian McKee

Sources

- <https://www.americanprogress.org/wp-content/uploads/2012/11/CostofTurnover.pdf>
- <https://www.pewresearch.org/fact-tank/2018/04/11/millennials-largest-generation-us-labor-force/>
- <https://www.monster.com/career-advice/article/shrm-industries-need-more-workers-1216>
- <https://www.linkedin.com/feed/update/urn:li:activity:6524020164739022848>
- <https://blog.gethyphen.com/blog/people-analytics-reveals-top-reasons-for-attrition-and-it-isnt-compensation>
- <https://www.bls.gov/>
- <https://channels.theinnovationenterprise.com/articles/ai-s-role-in-improving-employee-retention>
- <https://www.digitalistmag.com/future-of-work/2018/04/10/ai-cracks-employee-retention-code-for-small-midsize-businesses-06064626>
- <https://www.hrci.org/community/blogs-and-announcements/hr-leads-business-blog/hr-leads-business/2018/12/17/how-ai-is-helping-improve-employee-retention>
- <https://gallery.azure.ai/Experiment/HR-Analytics-Predict-Employee-Attrition>