1. **What is so confusing about the Confusion Matrix**
2. Resources

How important is Accuracy?

The client asks:

"How do I get the most accurate model?"

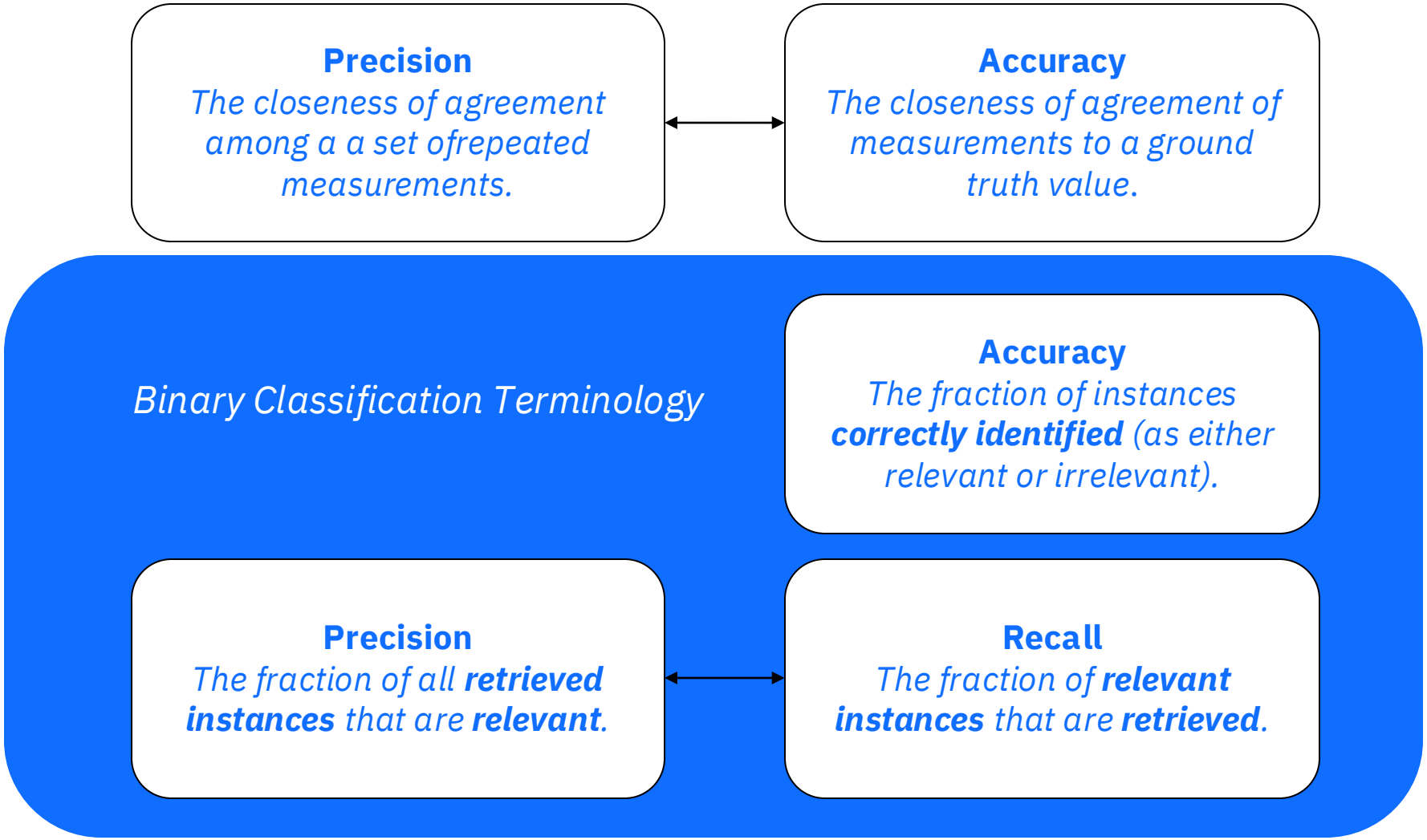
The data scientist responds by asking:

"What business challenge are you trying to solve using the model?"

So Accuracy need not necessarily be the one-and-only model metrics data scientists chase and include simple explanation of other metrics as well.

What about precision and recall?

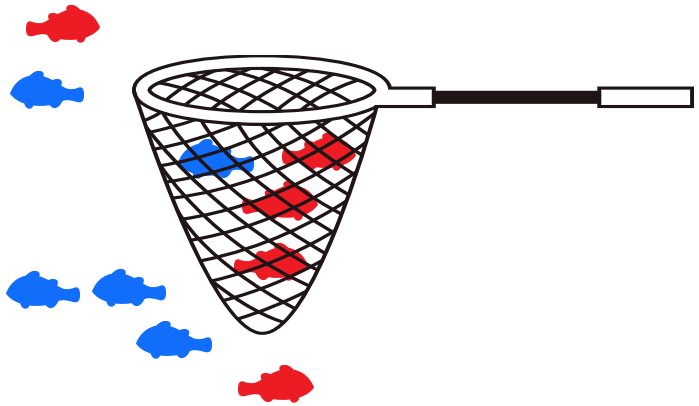
What About Precision and Recall?



When recall is more important than precision

5 Relevant data points (red fish)

5 Irrelevant data points (blue fish)



The search has retrieved 3 relevant documents out of a total of 5 relevant data points from the data set and 1 irrelevant document.

Recall = 3 / 5 = 0.6

Precision = 3 / 4 = 0.75 (the blue fish is not part of the equation at all).

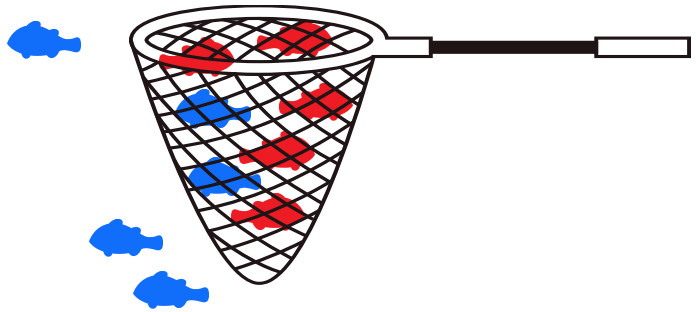
The case of 100% recall and low precision

5 Relevant data points (red fish)

5 Irrelevant data points (blue fish)

In some models, this is the preferred scenario even though there may be some irrelevant data points with a high score.

The algorithm team will then work on increasing the precision of this system.



The case of 100% precision and low recall

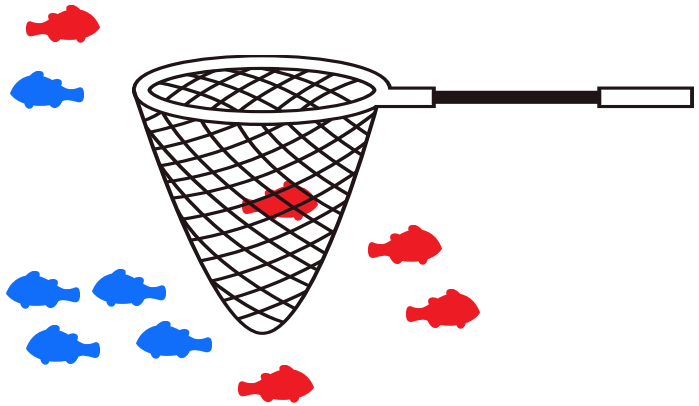
5 Relevant data points (red fish)

5 Irrelevant data points (blue fish)

Zero false positives, 100% precision — no blue fish in the net

But there are many false negatives — many red fish in the sea

There are potentially many data points that we will never consider. Perfect precision with poor recall is of no value to any ML model.



False Positives can be Dangerous

Notice that the accuracy for this model is very high, at 99.9%!!

But what if I mentioned that the positive over here is actually someone who is sick and carrying a virus that can spread very quickly?

Or the positive here represent a fraud case?

Or the positive here represents terrorist that the model says its a non-terrorist?

		Predicted/Classified	
		Negative	**Positive**
Actual	**Negative**	998	0
	Positive	1	1

Precision and Recall

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

		Predicted: Negative	Predicted: Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

Precision

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

		Predicted	
		Negative	**Positive**
Actual	**Negative**	True Negative	False Positive
	Positive	False Negative	True Positive

Recall

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$= \frac{True\ Positive}{Total\ Actual\ Positive}$$

		Predicted	
		Negative	**Positive**
Actual	**Negative**	True Negative	False Positive
	Positive	False Negative	True Positive

May the Best Model Win

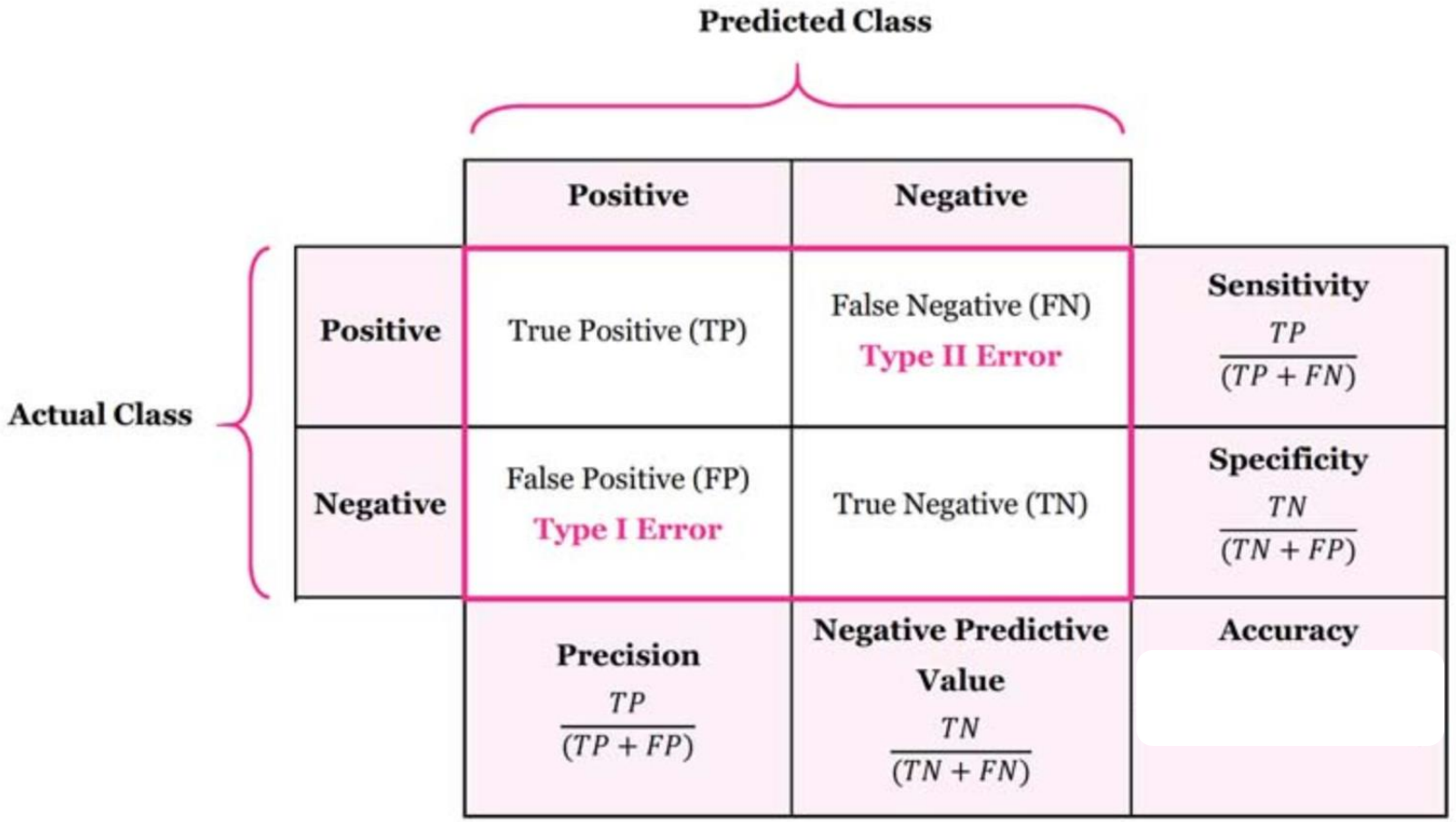
Recall actually calculates how many of the Actual Positives our model captures through labeling it as Positive (True Positive).

Recall shall be the model metric we use to select our best model when there is a high cost associated with False Negative.

In fraud detection or sick patient detection:
If a fraudulent transaction (Actual Positive) is predicted as non-fraudulent (Predicted Negative), the consequence can be very bad for the bank.

Deriving Precision, Recall, and Accuracy from the Confusion Matrix

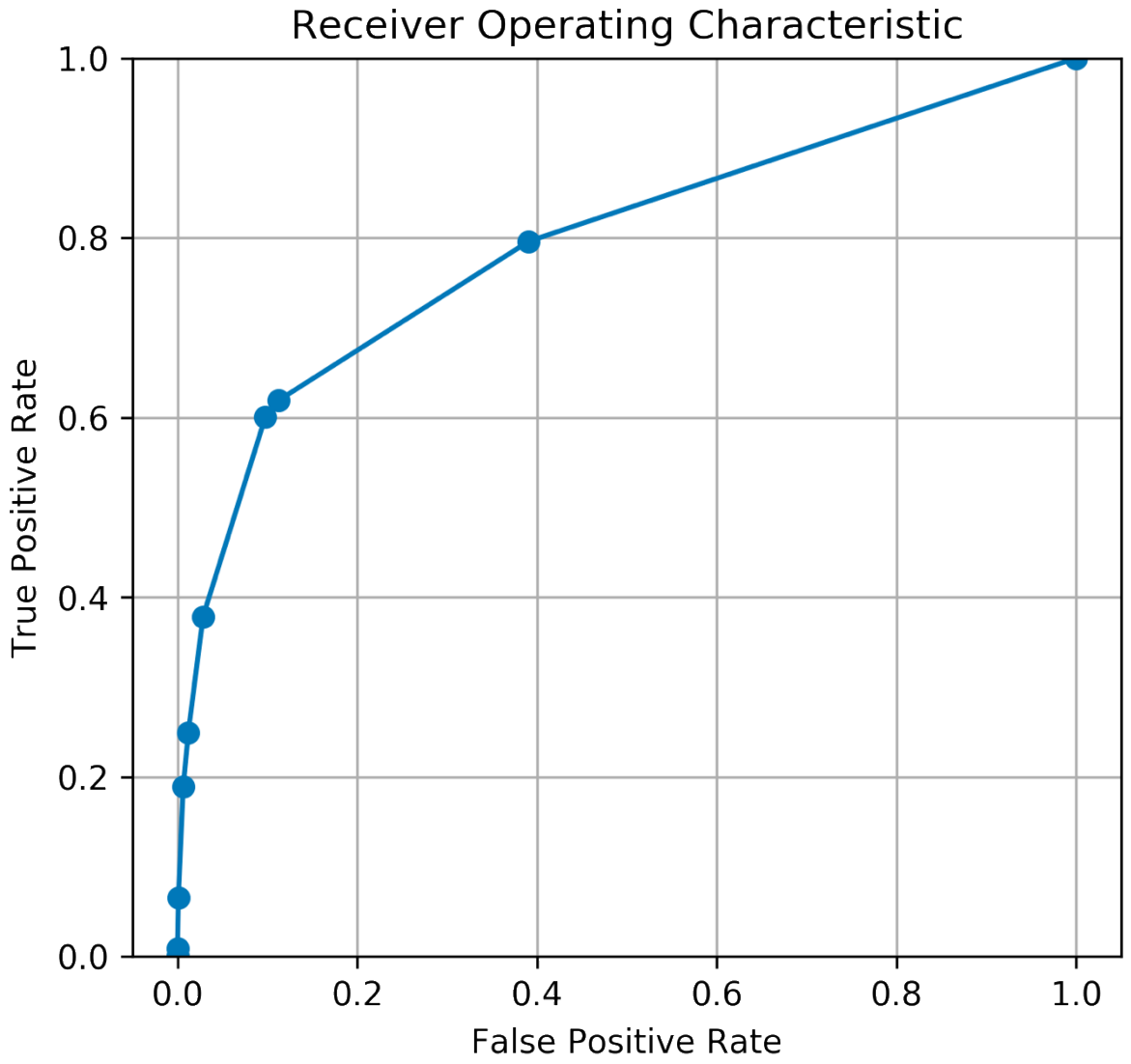
Accuracy = ----------



May the Best Model Win

The ***True Positive Rate*** is sometimes plotted against the ***False Positive Rate*** in order to see the tradeoff between the two as thresholds are adjusted.

This plot is called the ***Receiver Operating Characteristic.***

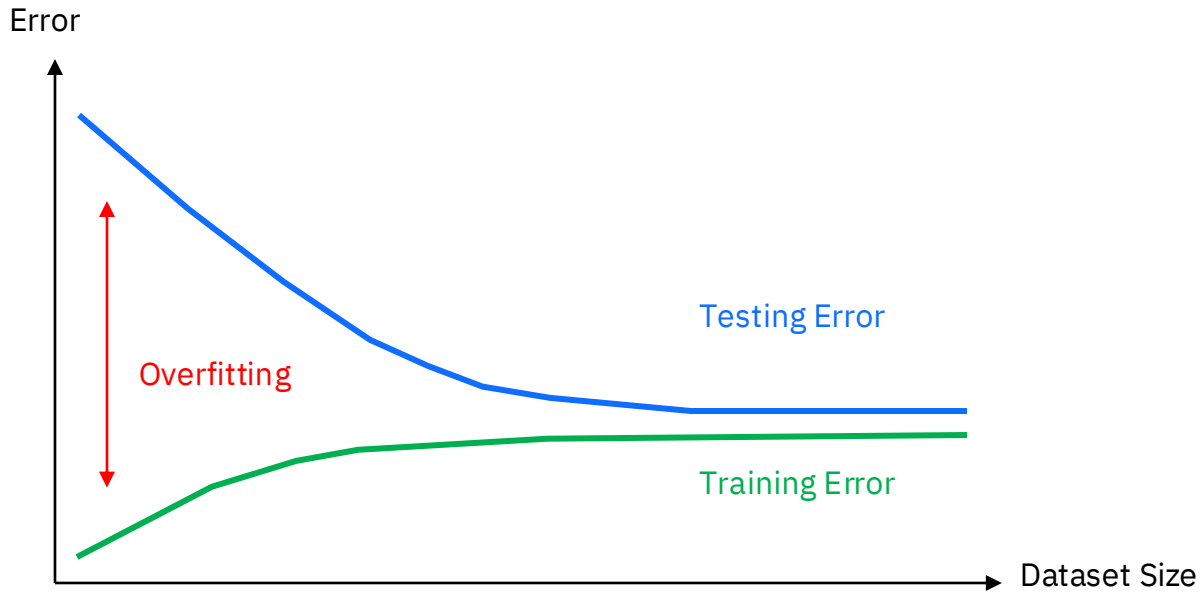


Dangers of Overfitting Data

The biggest risk is overfitting

- A model that works well on training data but performs badly on new data

Trick is to select a model without knowing the data it will be applied to



Example of overfitting

We want to predict if a student with a GPA of 3.3 will be admitted to Wentworth Institute of Technology in Boston

Assume we train a model from a dataset of 5,000 students and their outcomes.

Next, we try the model out on the original dataset, and it predicts outcomes with 99% accuracy... wow!

But now comes the bad news.

When we run the model on a new ("unseen") dataset of student admissions, we only get 50% accuracy... uh-oh!

Our model doesn't *generalize* well from our training data to unseen data.

This is known as overfitting, and it's a common problem in machine learning and data science.

F1 Score

F1 is a function of Precision and Recall

$$F1 = 2 \times \frac{Precision * Recall}{Precision + Recall}$$

F1 Score is needed when you want to seek a balance between Precision and Recall.

1. Natural Language Processing
2. Data Types
3. **Resources**

Resources

[1] Beyond the hype: A guide to understanding and successfully implementing artificial intelligence within your business
https://www.ibm.com/downloads/cas/8ZDXNKQ4

[2] A Practical Guide to Building Enterprise Applications: by Tom Markiewicz and Josh Zheng – Feb 2018 O'Reilly
https://tmarkiewicz.com/getting-started-with-artificial-intelligence/

[3] The New York Times - Nils Nilsson
https://www.nytimes.com/2019/04/25/obituaries/nils-nilssen-dead.html

[4] Why artificial intelligence is enjoying a renaissance
http://www.economist.com/blogs/economist-explains/2016/07/economist-explains-11

[5] How Cognitive Systems Could Redefine The Way Governments Work
http://www.forbes.com/sites/ibm/2016/09/20/how-cognitive-systems-could-redefine-the-way-governments-work/#1e1ed4f52ff1

[6] March of the Machines
http://www.economist.com/news/leaders/21701119-what-history-tells-us-about-future-artificial-intelligenceand-how-society-should

[7] We have to upgrade our skills to match intelligent machines
http://www.businessinsider.com/how-labor-can-keep-up-with-artificial-intelligence-2016-10?IR=T

[8] Forbes & IBM - Intelligent Automation: How AI and Automation are Changing the Way Work Gets Done
https://www.ibm.com/downloads/cas/RE2XMOLR