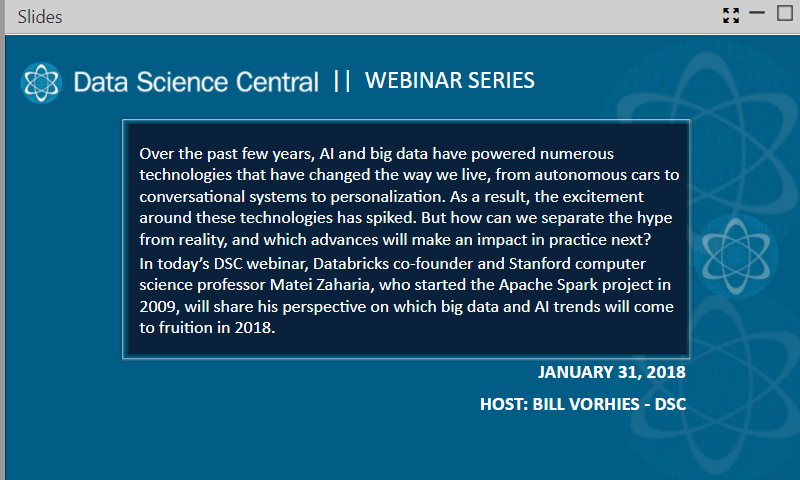
Exercise 3

[Take pictures and make notes about the predictions. 1](#_Toc505287973)

[My own predictions 20](#_Toc505287974)

## Take pictures and make notes about the predictions.

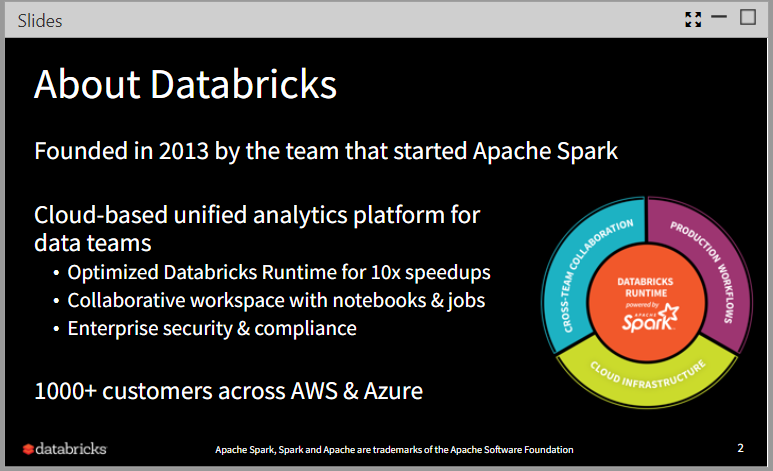


How can we find the reality of AI and big data, and which advances will have an impact in practice?

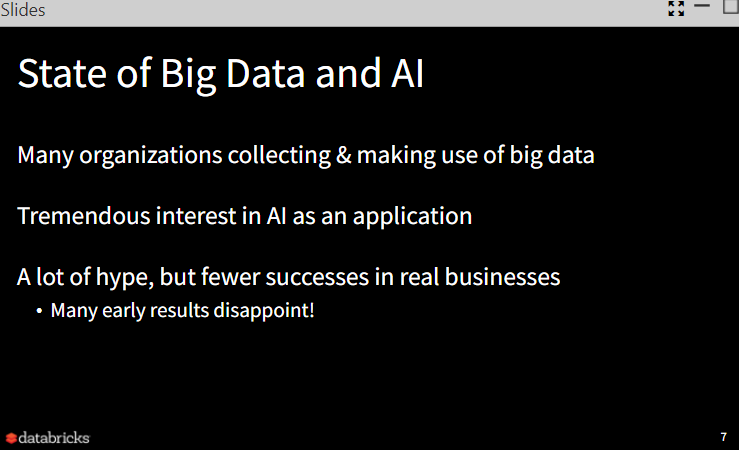


1. The current state of big data and AI
2. Some of the new innovations taking place in research
3. Key challenges that companies face in getting value from data and AI
4. Mater’s predictions for 2018(how companies and the technology industry will overcome these challenges)

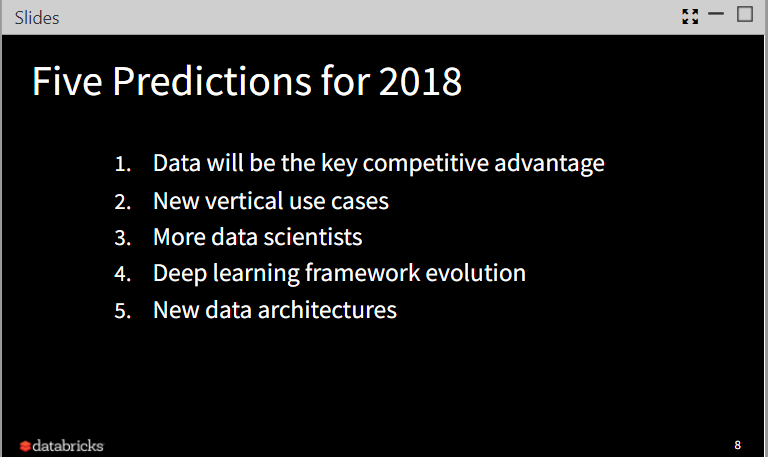




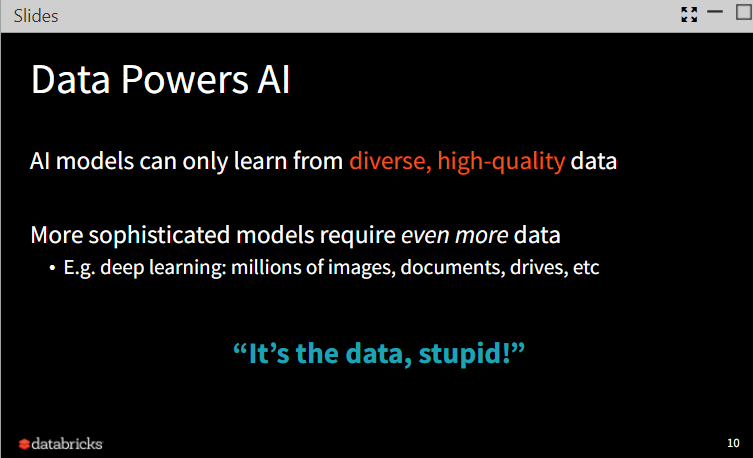
(introduction of Databricks-data teams analytics platform)



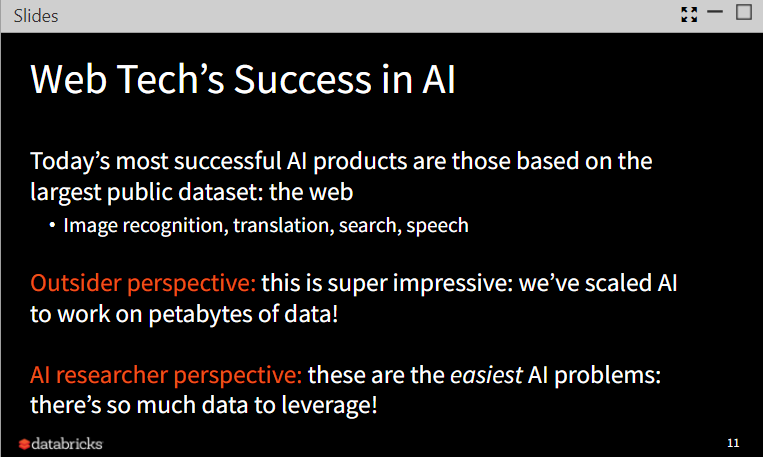
Big data is being used in many organizations, but fewer successes happened in real business



This is the predictions made by Matei about AI and big data

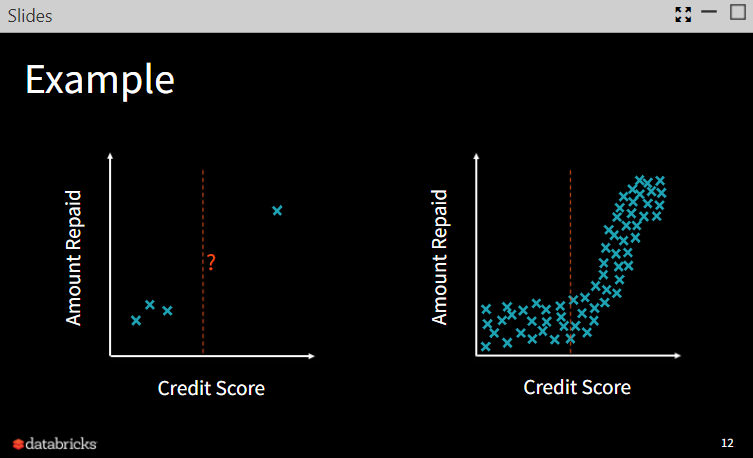


The requests from AI model to data is huge, which means data powers AI, it also shows the importance of data in AI.

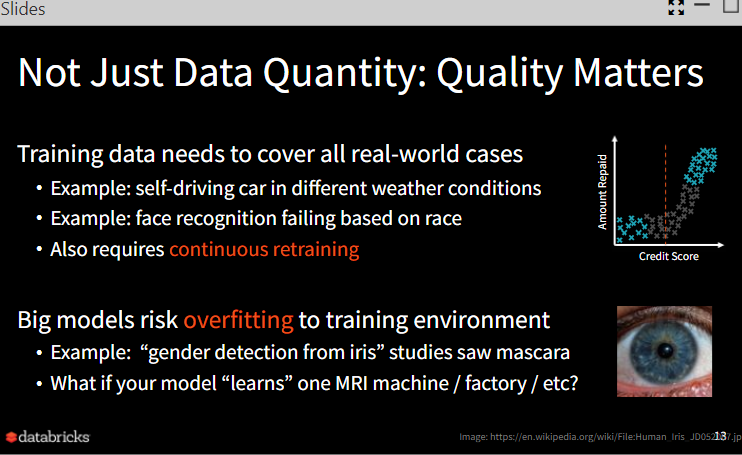


We can find that many web tech companies have made many successful products in AI, because they have the largest public dataset: the web, such as Image recognition (e.g. Microsoft cognitive service)

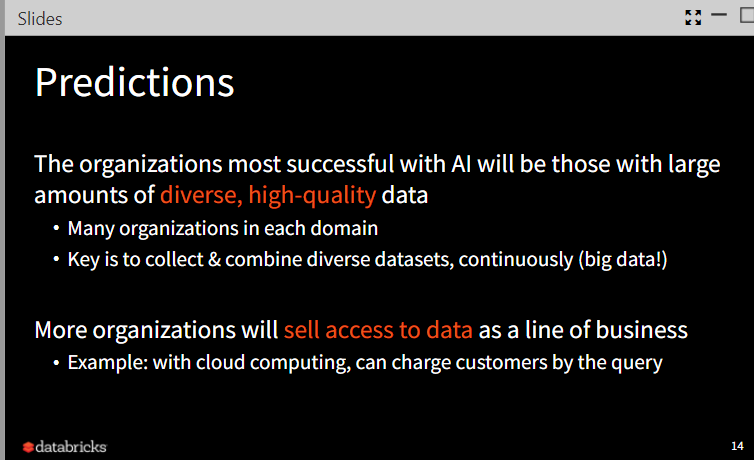
For AI researcher: there is so much data to leverage, the time and energy spent on dealing with data is taking significant percentage among all..



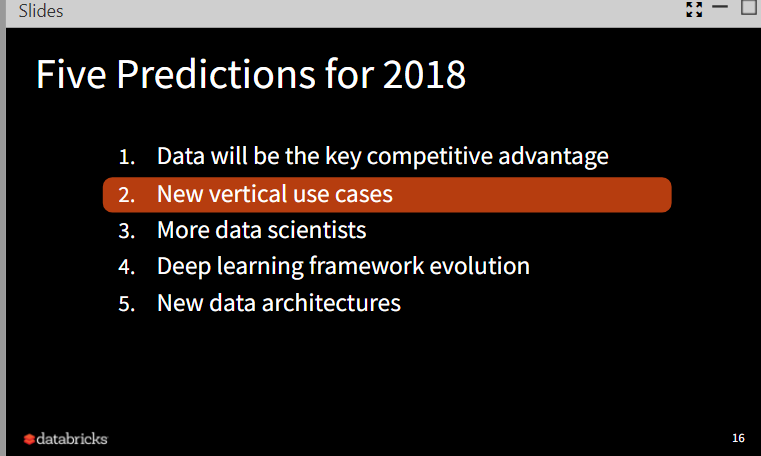
We can analysis little things from little data, but with amount of data, the conclusion is easier to get. Data quantity is important.

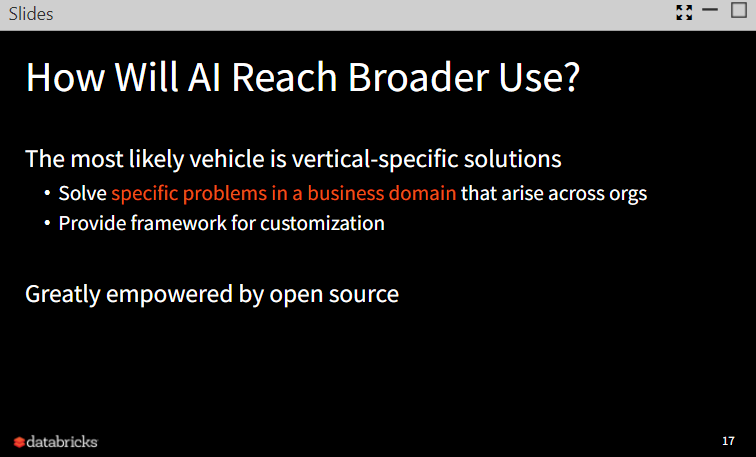


Quality of data is also important, even though the data is with large amount, if the quality can’t satisfy the request, we can’t get good model.

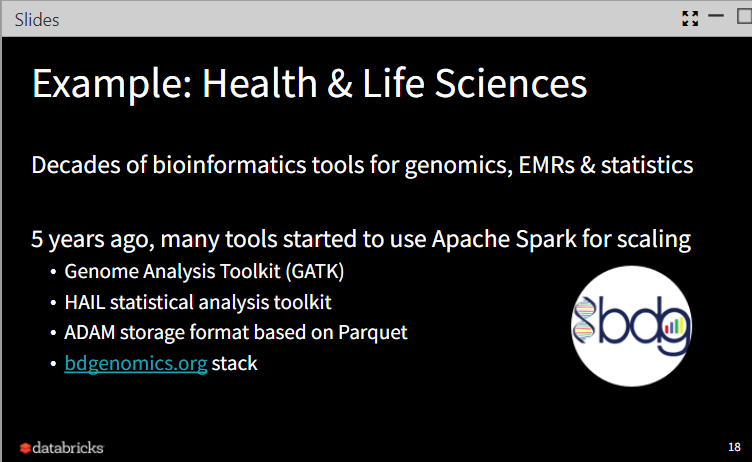


Data with diversity and high quality will help the success of AI product, also, there is a demanding of data, so more organizations will sell the access to data as a business.

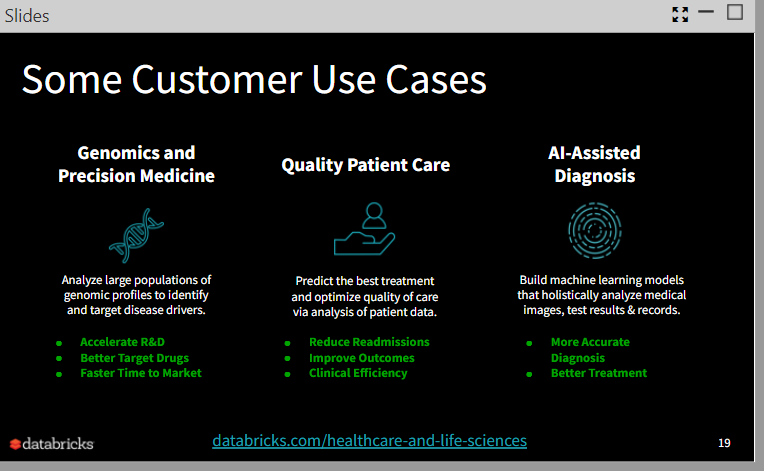




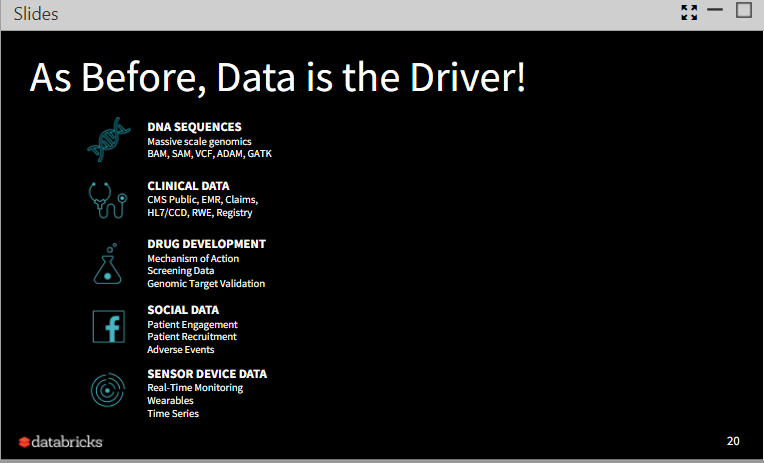
AI reaches broader use such as vertical-specific solutions, provide framework for customization



Such as bioinformatics tools for genomics, EMRs and statistics



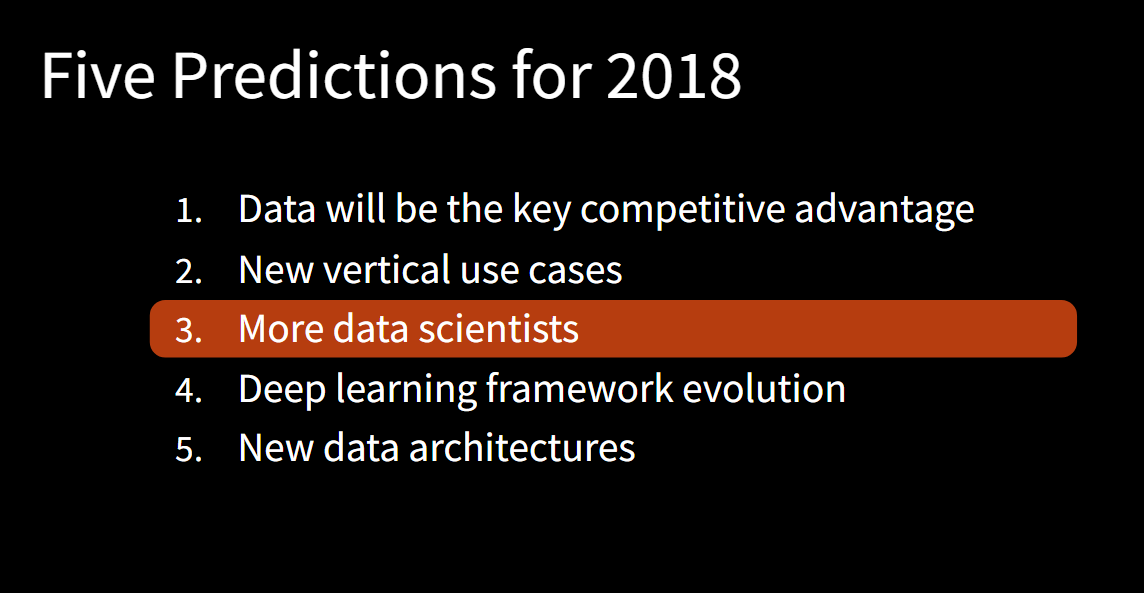
There are some customer user cases: genomics and precision medicine; quality patient care; AI-Assisted Diagnosis

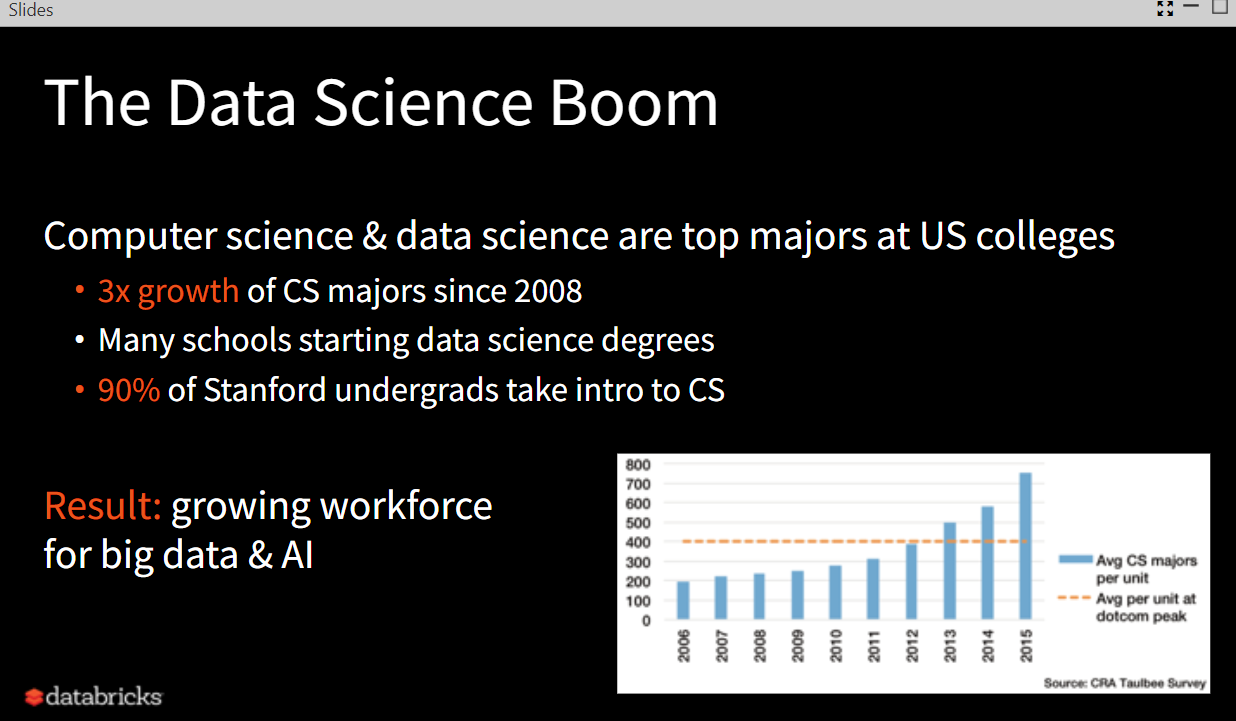


For DNA Sequences, clinical data, drug development, social data, sensor device data and so on, data is the driver



Here introduces other domain-specific libraries that are related to AI and big data

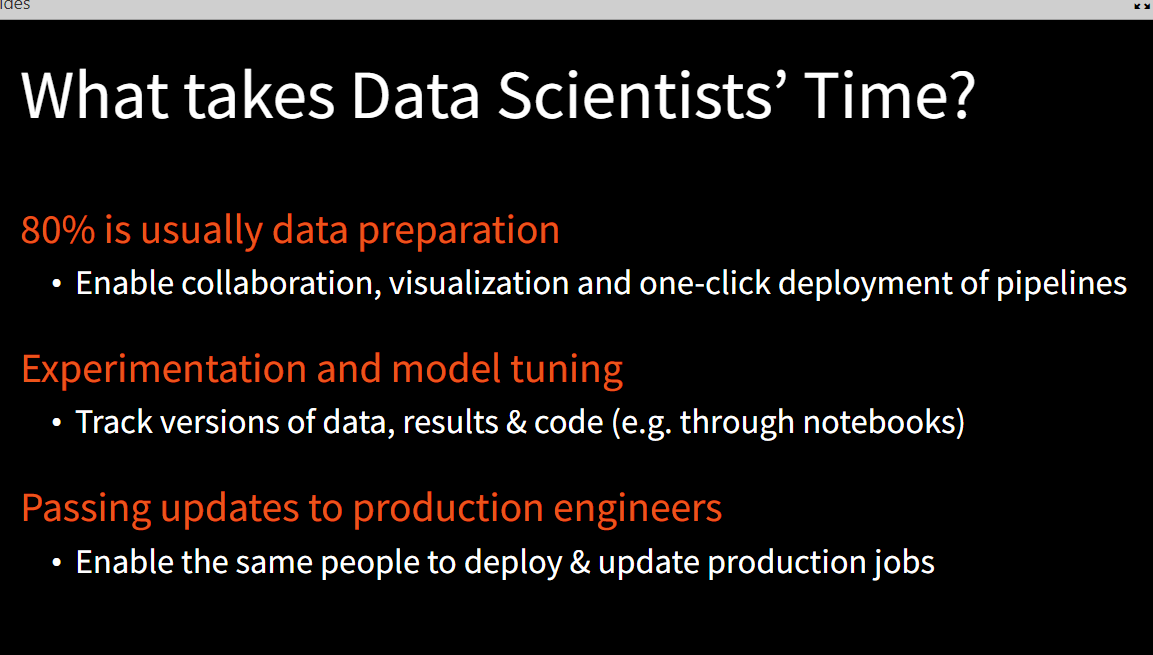




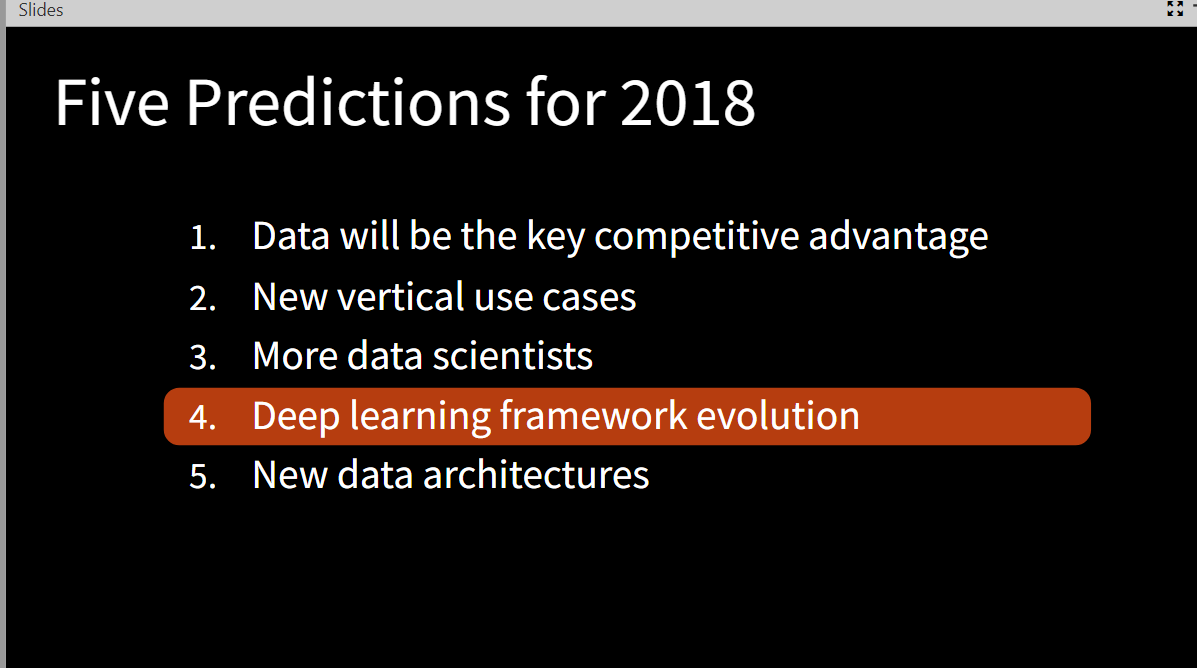
More and more people pay attention to data science and computer science

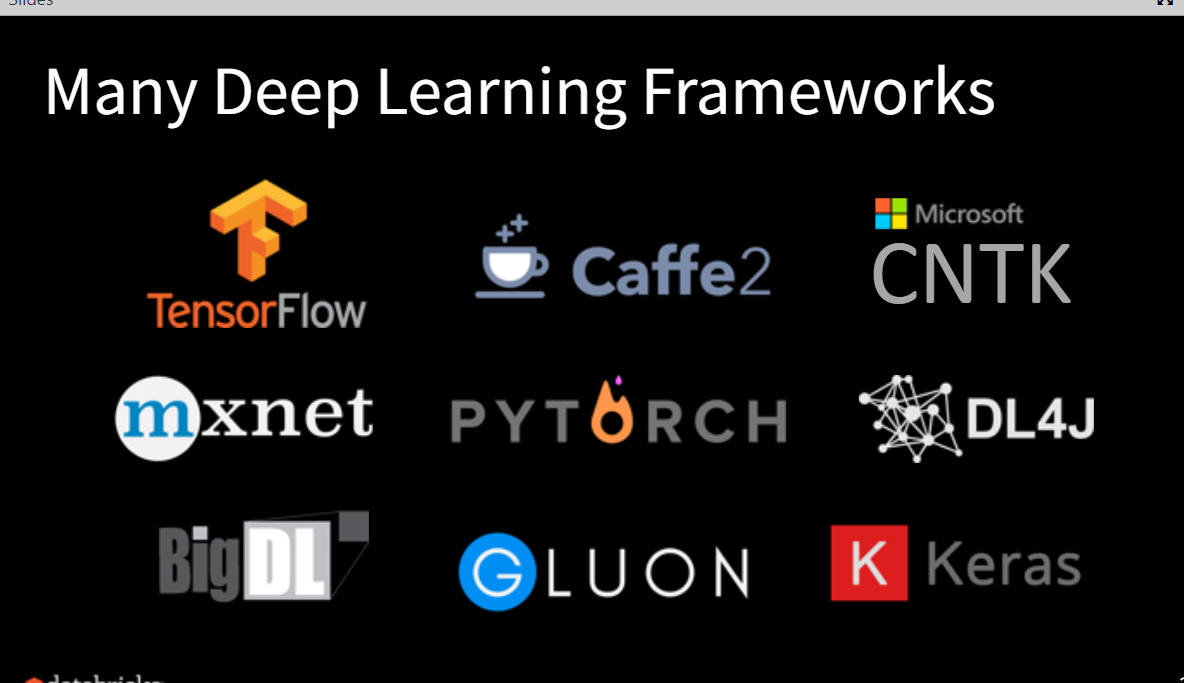


Although more and more people pay attention to the data science, the demanding for data scientist is still huge.

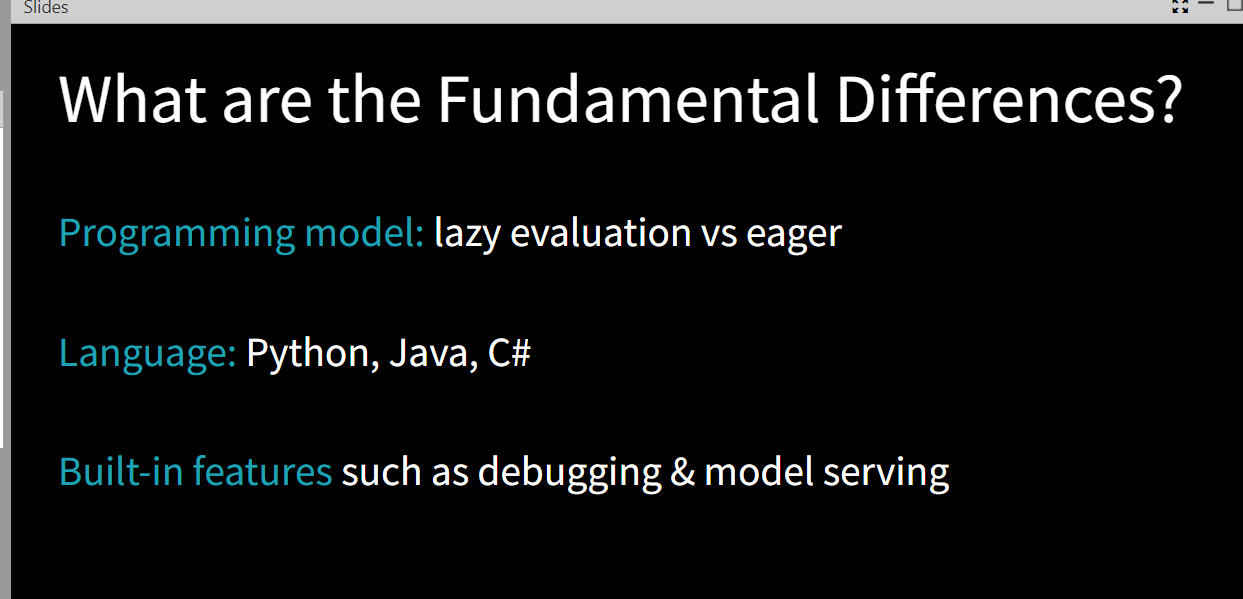


Data preparation is the work that take most of the time for data scientist

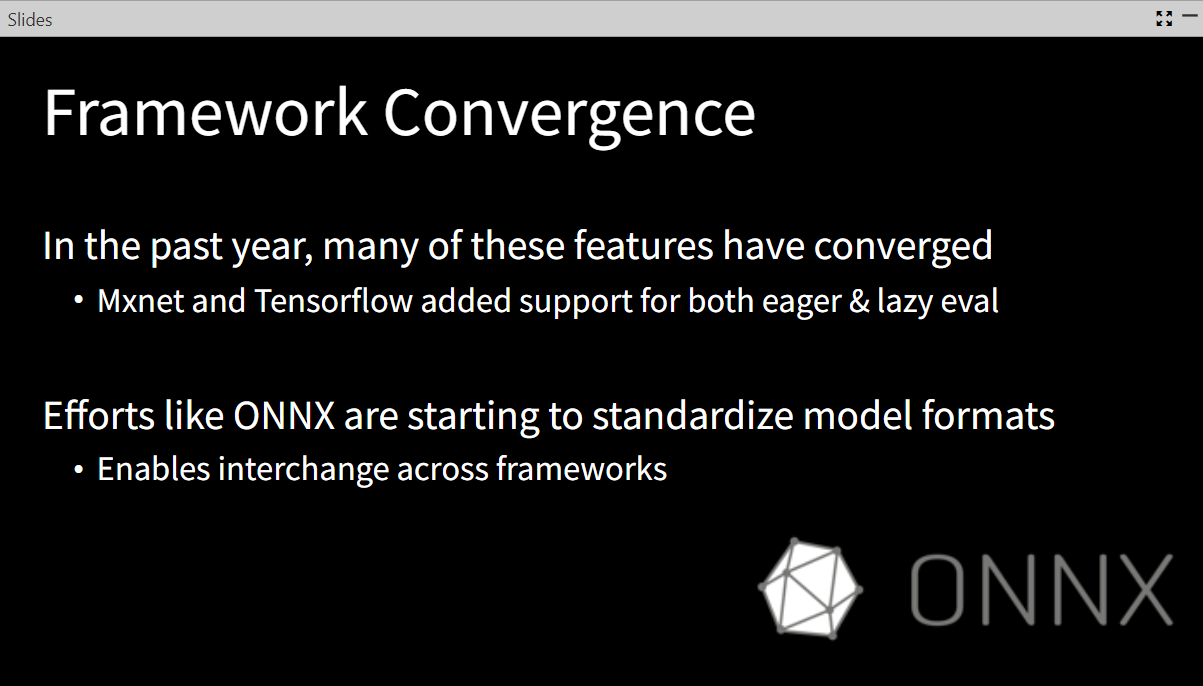




There are some deep learning frameworks that are famous



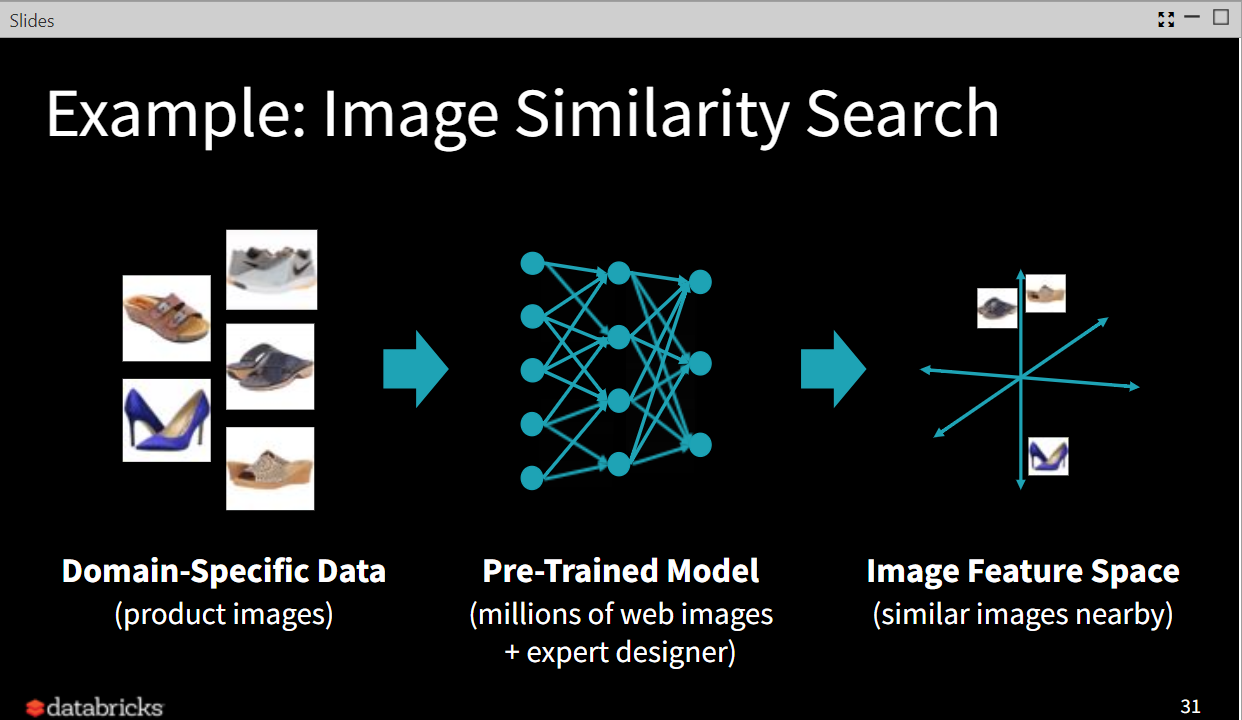
The differences among frameworks: programming model, language, built-in features



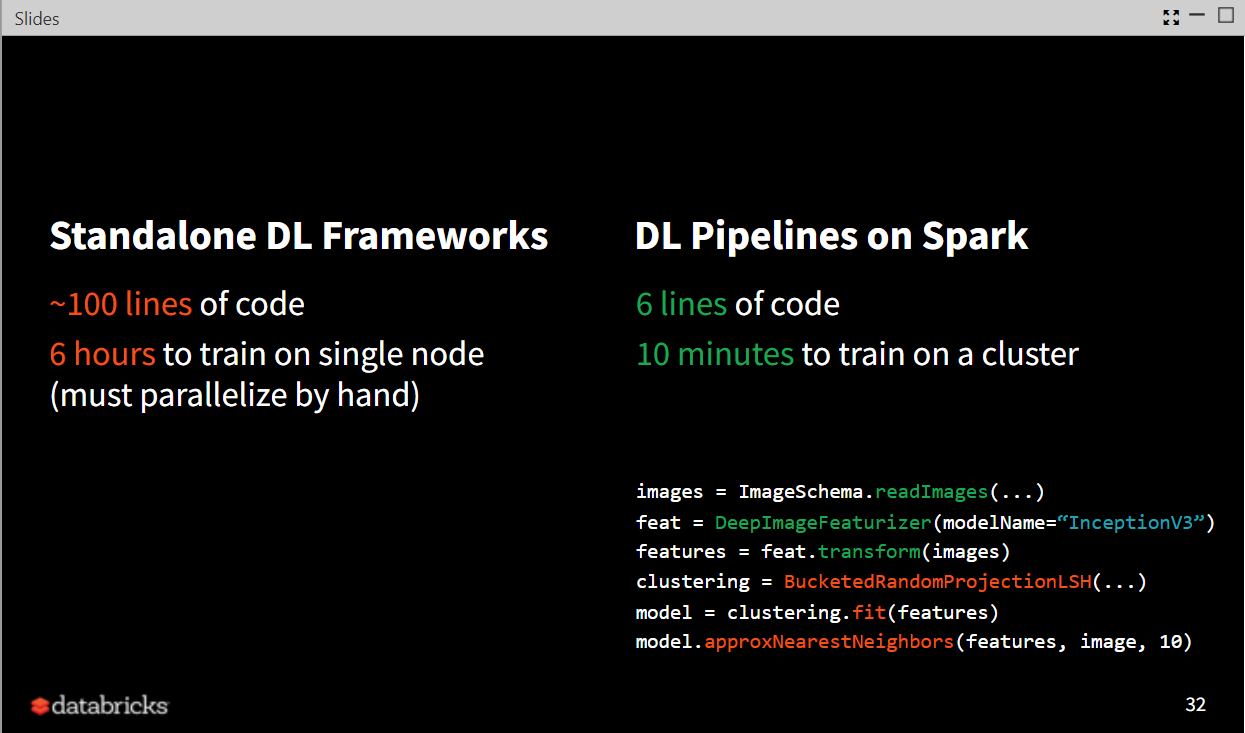
The standardization and convergence have been started



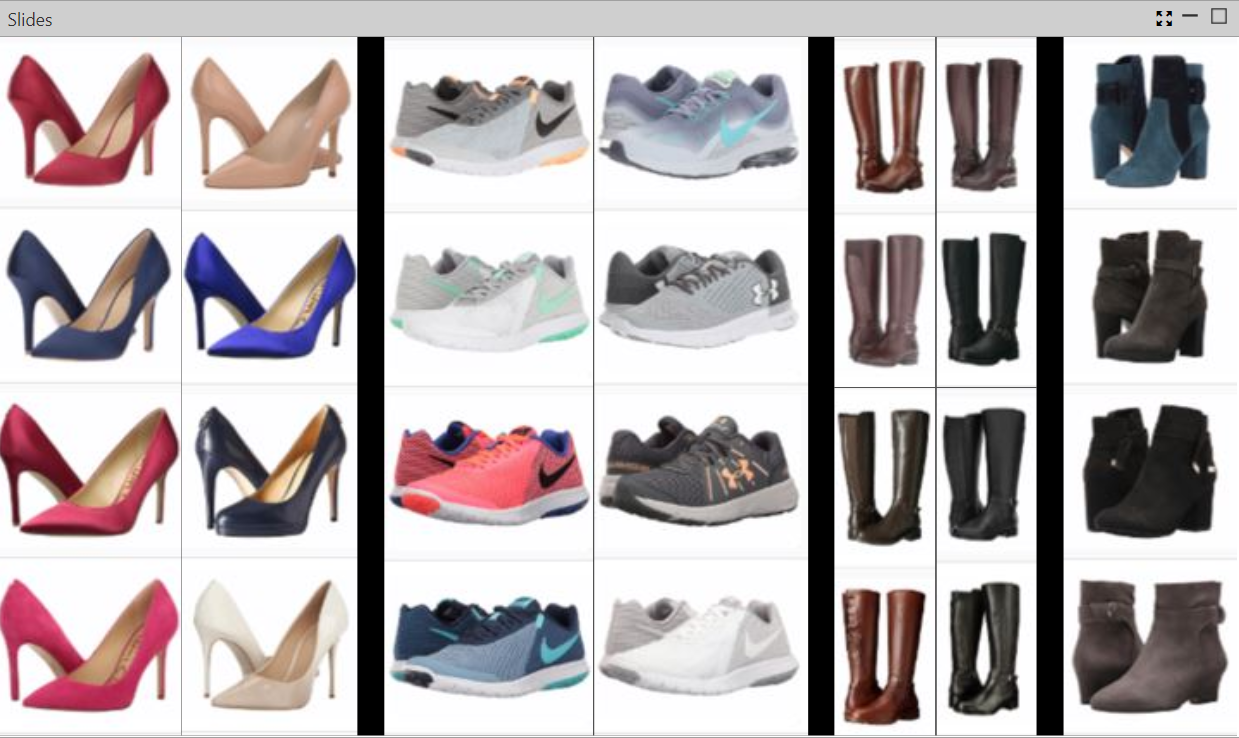
Current frameworks mostly focus on model developers; various projects focus on end-users



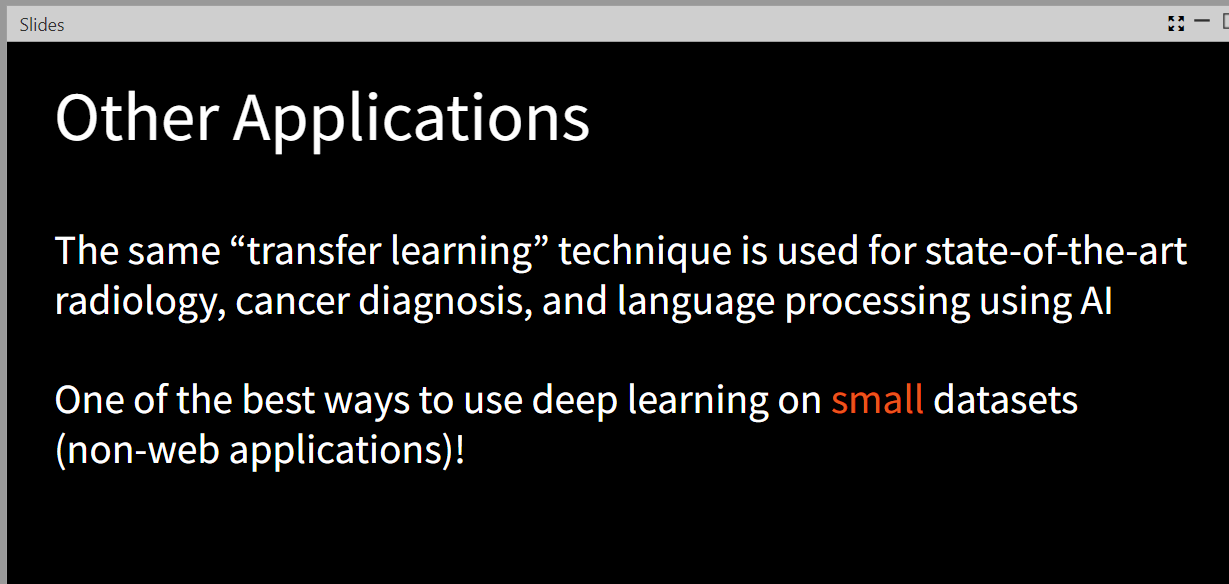
When applying domain-specific data, we can get pre-trained model, than generate an image feature space



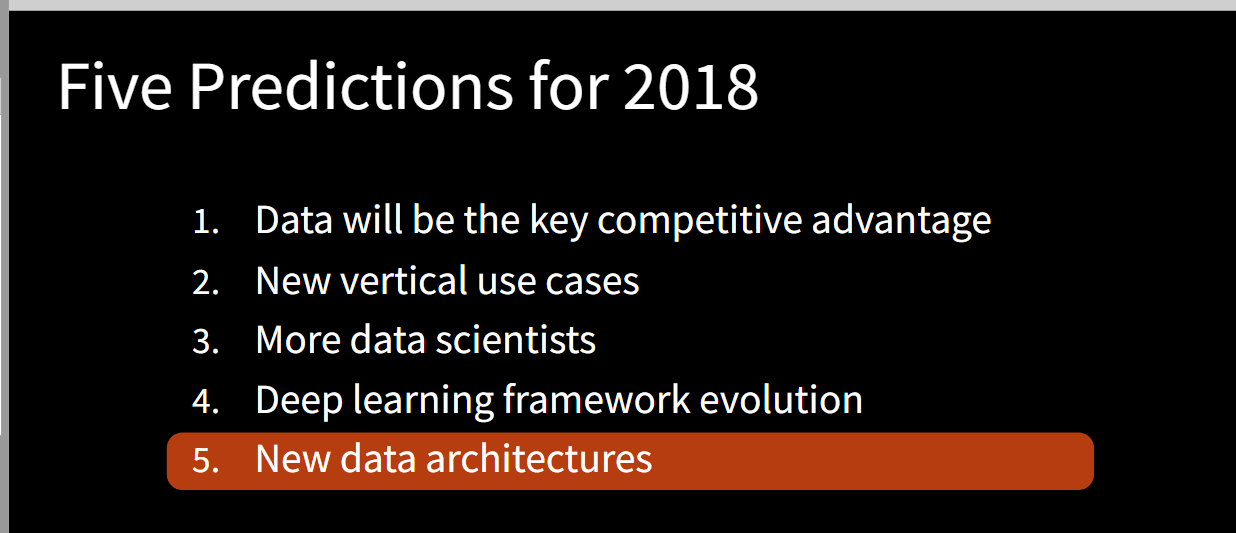
(example of frameworks)

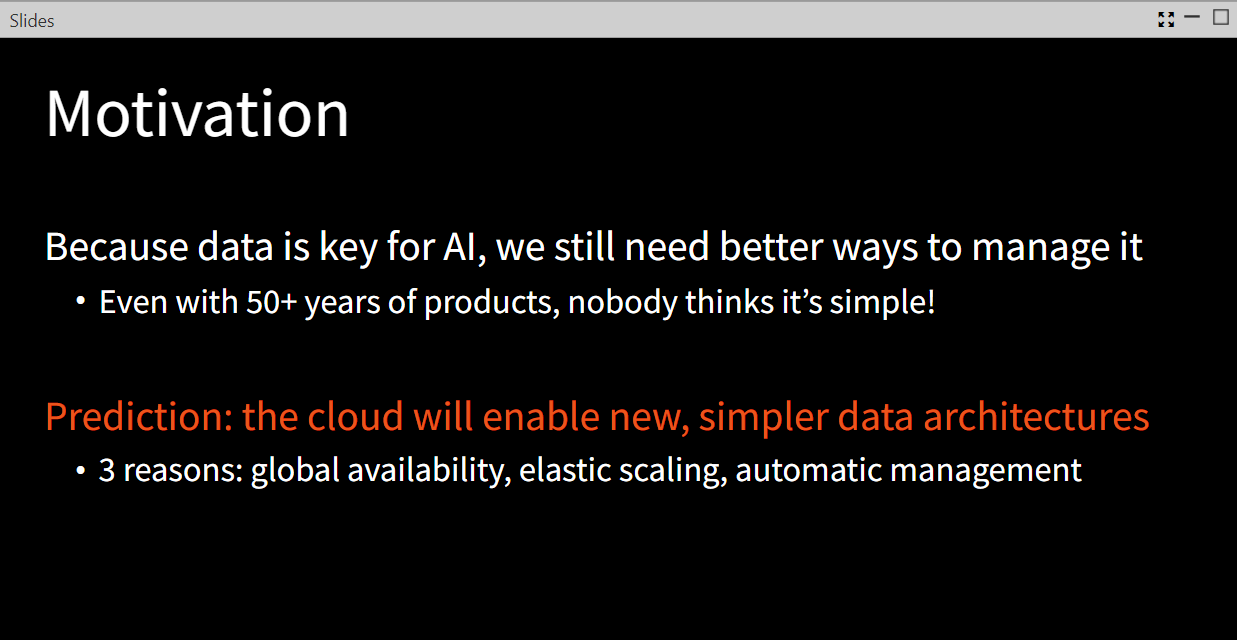


Shoes with different similarity

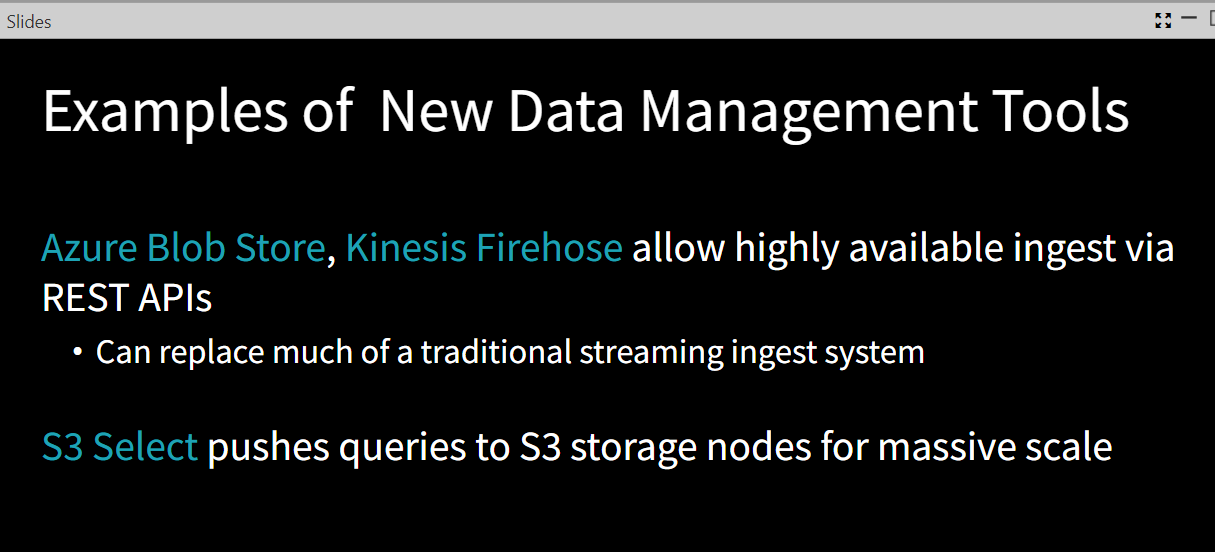


“transfer learning” technology is one of the best way to use deep learning on small datasets

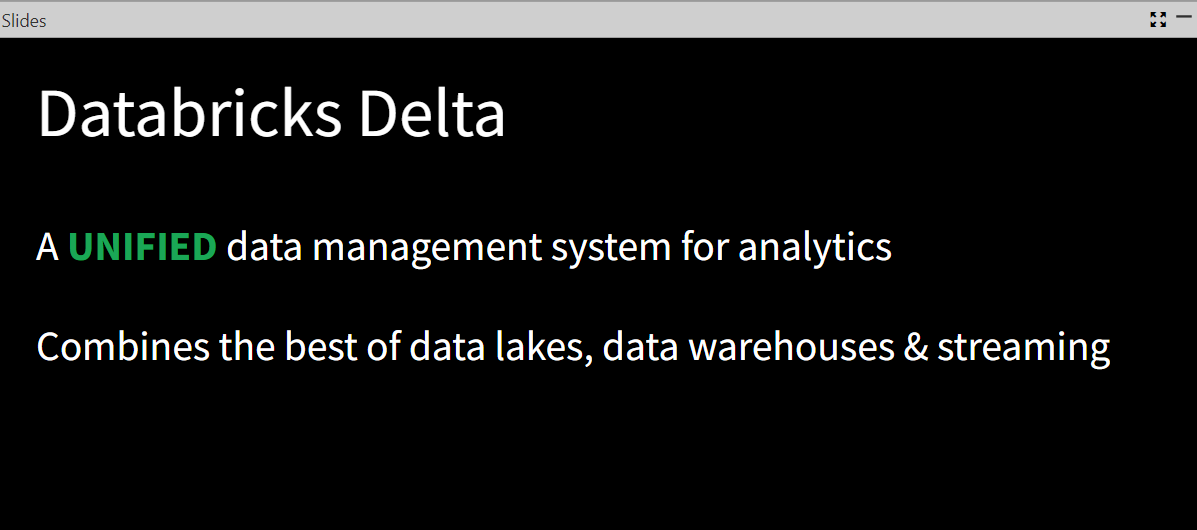


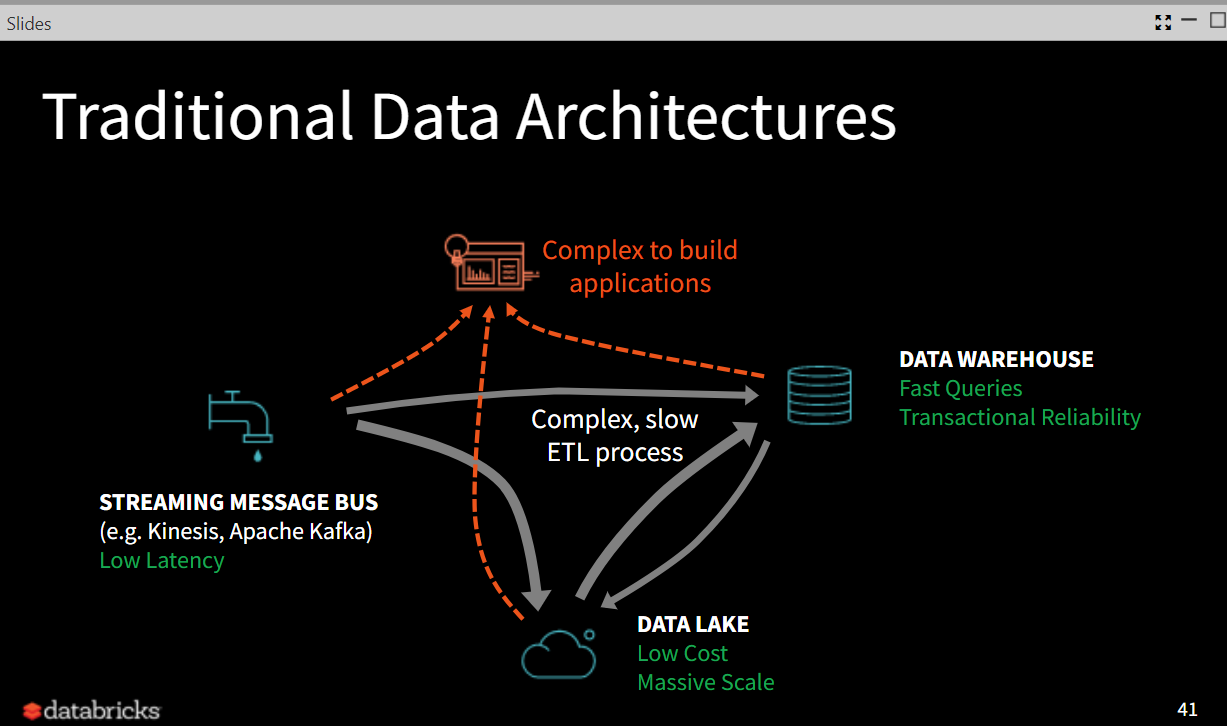


New, simpler data architectures will be enabled by the cloud

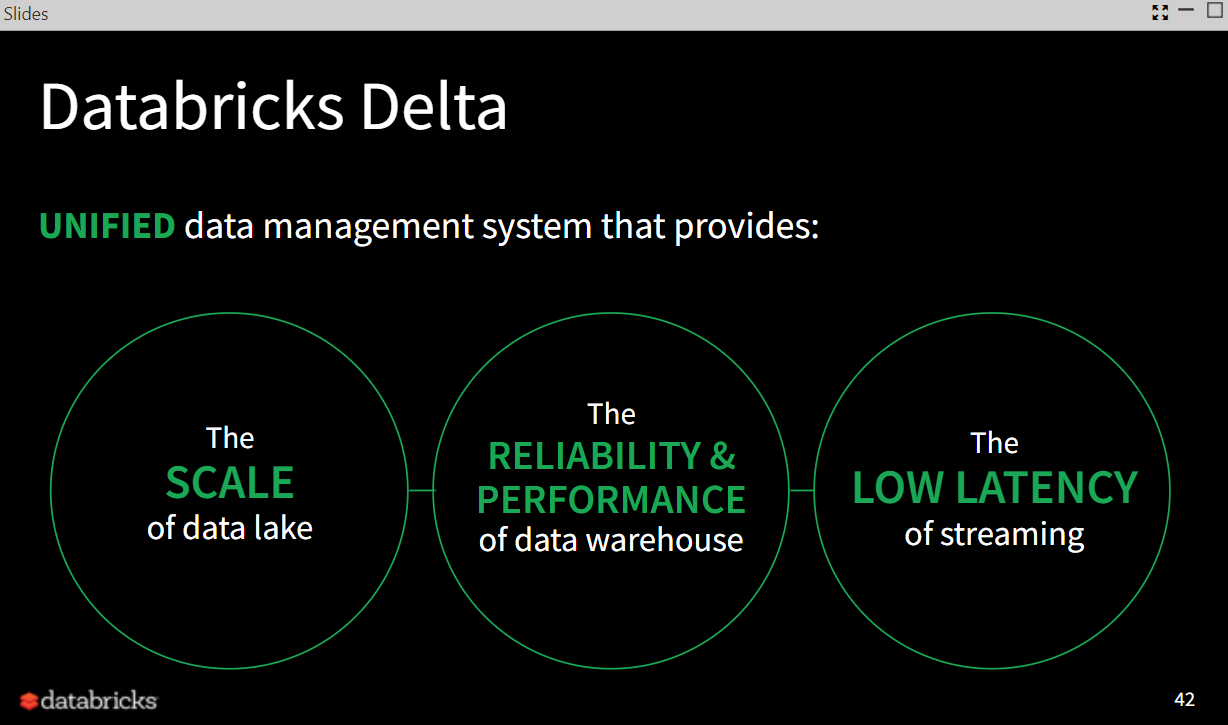


There are some examples about new data management tools

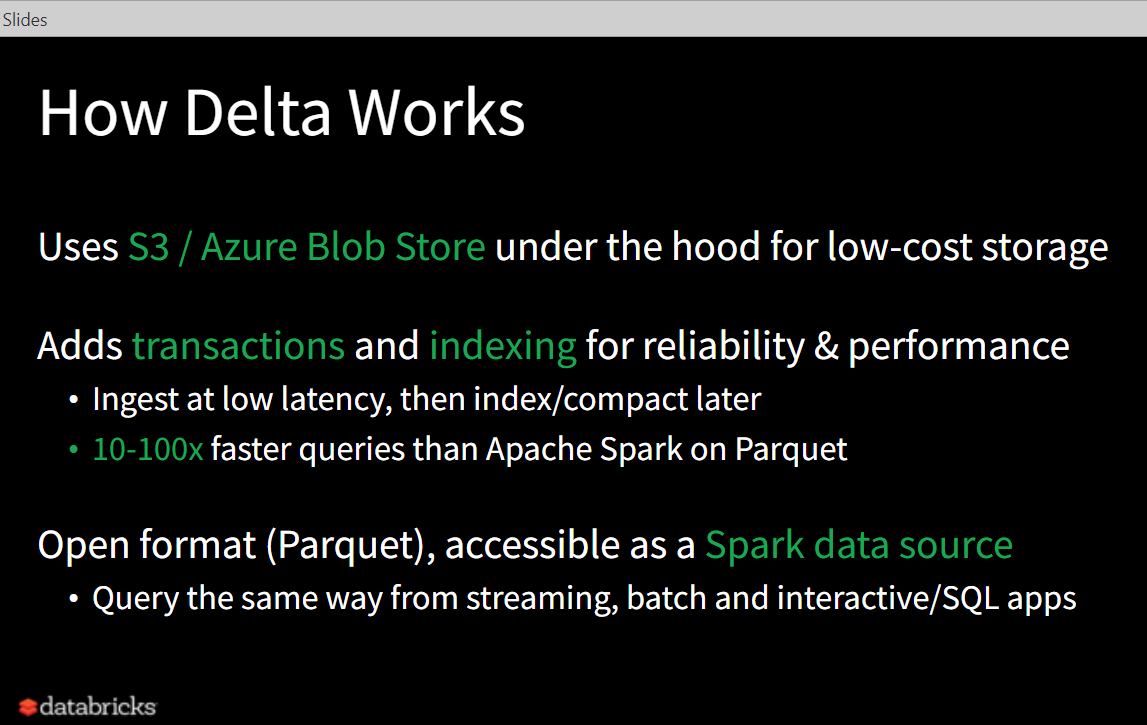




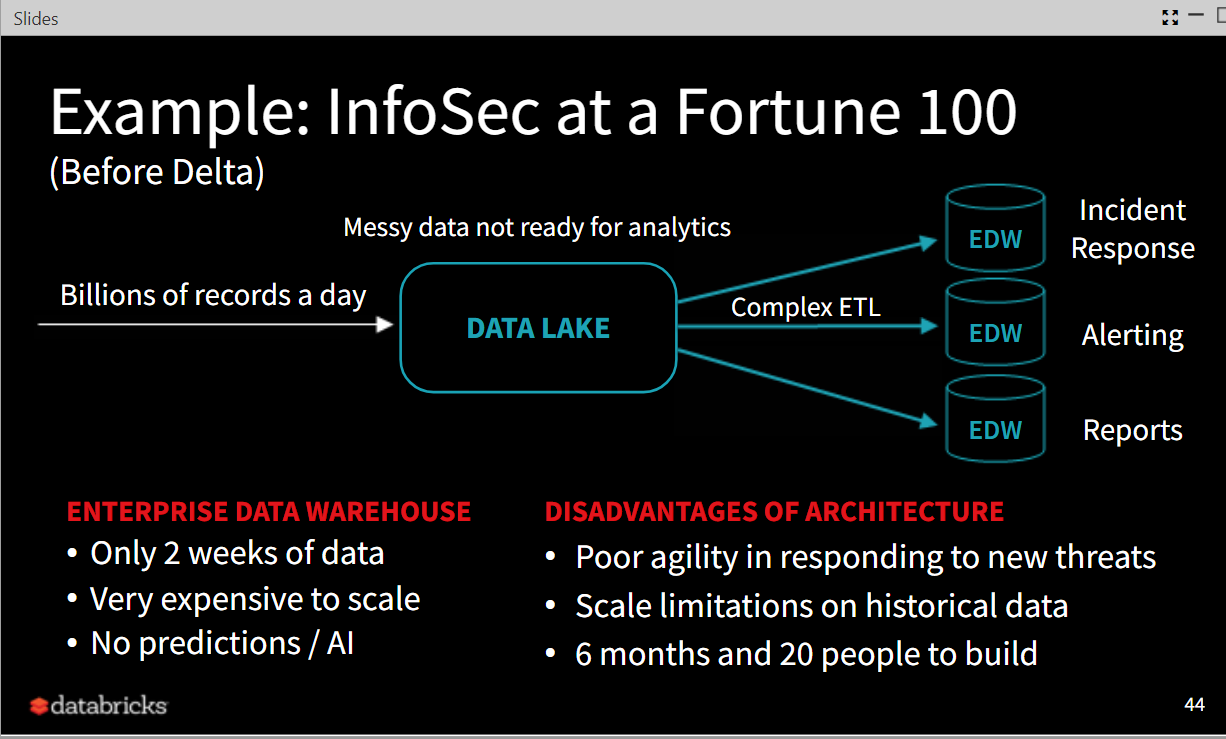
The drawback of process of ETL process is complex and slow, also the data architectures is very complex to build applications

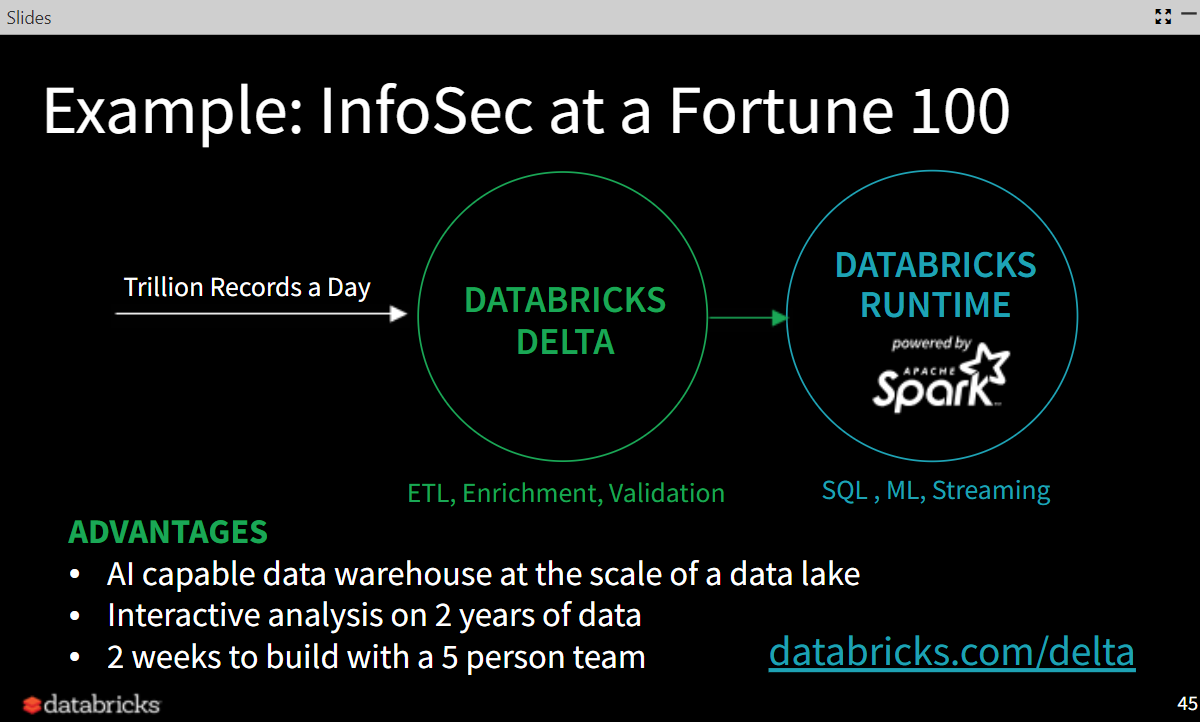


Introduction to Databricks Delta



The principle for delta works

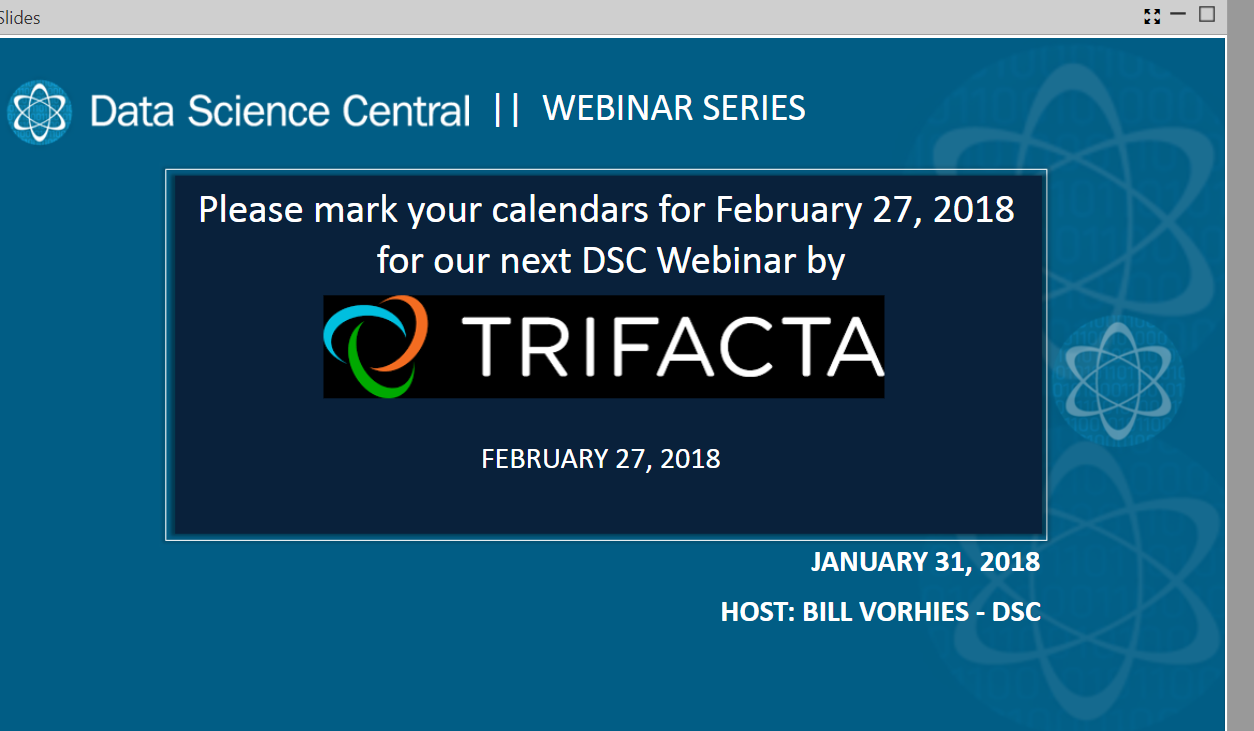




Example of using Databricks Delta and its advantages



More successful AI use cases in domain; Data will be the key to AI; rich, high-quality data is important to those organizations



## My own predictions

Make your own predictions (10 statements) around big data and artificial intelligence based on the arguments of the webcast and other material from Google, Google Scholar and Scopus.

1. There will be more products that are developed for collecting data which has diverse and high-quality quality(The quality of data might be detected or tested in those products) Because in the webcast speaker said that data is the driver and key to AI product, also the diversity and high-quality of data is so important, so I think there will be more demanding on the “good” data, which leads to new products has related functions.
2. There will be more organization sell data that has diversity and high-quality, because the data demanding is also increasing, and more organizations pay attentions to AI, so the data is the products they need.
3. There will be some services that help data scientist to prepare data, because in the webcast, the speaker said that data scientists take usually 80% time to do data preparation, which enable collaboration, visualization and one-click deployment of pipelines, so there will be some services that can help to reduce the pressure of them, it could be part of the preparation of data.
4. Data security services will be improved. (Because the importance of data has been mentioned in webcast, for those organizations, data will be the driver of AI product, so protect data from been stolen or disclosure are important and related to business interests. The demanding for protecting data is increasing, so there will be more improved services to protect data)
5. 20% of enterprise will deploy AI to make decisions and provide real-time instructions[1](AI will suggest the option for customer and recommend terms to give suppliers, also it is able the guide the employee on what to do in real time, which means saving a lot for those related organizations)
6. More money will pour into AI enterprise projects than ever before[2](Because more and more businesses will get involved in AI after seeing many successes achieved by innovators and market leaders in AI such as self-driving cars and ships, life-saving medical advances in AI)
7. More jobs will be created by AI related fields (According to the speaker in webcast, even there are more and more people that are paying attention to data science and computer science, there is still a huge demanding in those fields; also, AI will created jobs for people who can do AI products related things such as building, selling and so on)[3]
8. The process of data will be improved (Because in the webcast, the speaker shows that in traditional architectures, data is been complex and slow processed, it is a drawback in data process, so there is demanding in improve it, so I think more people will pay attention to the improvement of processing data )
9. AI will more fully integrate into daily Life[4] (Because with time going by, more data will be generated and the learning for things like robot will be deeper, which means robots will get better at complex tasks; also more and more products will contains AI in our life cause the benefits we can get from AI is a lot)
10. There will be some standardization in the AI related framework and architecture (there are so many framework that related to AI nowadays, as the speaker mentioned, many differences for those framework, so there might be the integration for those framework or make those products obey some rules which help the use of customers)

Reference:  
[1] <https://www.forbes.com/sites/gilpress/2017/11/09/10-predictions-for-ai-big-data-and-analytics-in-2018/#4ff34922403c>

[2] <https://www.forbes.com/sites/bernardmarr/2017/12/18/5-key-artificial-intelligence-predictions-for-2018-how-machine-learning-will-change-everything/2/#a9c6d3fb05ea>

[3] <https://towardsdatascience.com/2018-predictions-for-ai-484cd15e61ec>

[4] <https://blogs.nvidia.com/blog/2017/12/03/ai-headed-2018/>