AI Assessment — Where Are We in 2022?

Since the 2012 Harvard Business Review issue presenting Data Scientist as the sexiest job title of the 21st century[1], much has changed in the field. Data Scientists are not a rarity any longer, and are often embedded in core business operations. In addition, more tools are available for Data Science teams to analyse bigger data, and train even larger models, while various platforms offer services for "effortless" deployment.

Data Science in the real world

The field has shifted from primarily generating actionable insights, to integrating Machine Learning (ML) models into large enterprise systems. These models may replace existing services built around rules or create new ones. Entire businesses are occasionally built solely on such models.

There is still ambiguity surrounding the responsibilities and skills of Data Scientists, but publicity and investments around Data Science have attracted more talent over recent years. This fueled the third AI renaissance, with impressive breakthroughs in research and engineering alike. ML models can now outperform humans (e.g., Question Answering in Natural Language Processing), which was considered unlikely not too long ago.

Businesses have been benefiting from this progress for a while now. Leaders have become more comfortable making the necessary investments in data science as more success stories are shared. There are now many ML models in production making predictions and decisions across many industries and enterprise functions. One of the risks this proliferation has revealed is bias in AI which the traditional model performance metrics were never designed to automatically expose. Research groups and innovation labs may be pushing the envelope of what is possible, but in the wild, biased data and poorly designed models preserve systemic prejudices.

As algorithms will learn from patterns in the data, they can reiterate or amplify biases, even when these are unknown to humans. The issue of bias is not new. It is a research subject in social sciences in and of itself. In this context, bias is often the focal point and understanding its cause and effect produces important insights affecting policies worldwide.

In the enterprise however, the goal is somewhat different. Instead of studying bias and causal inference, algorithms and models are used primarily to deliver predictions and decisions at scale.

---

1   https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century

FTI CONSULTING™

For example, in healthcare, algorithms are leveraged to support medical doctors and policy makers assessing risk and make decisions for millions of patients every year. However, as it was discovered a couple of years ago[2] such algorithms can falsely conclude black patients are healthier than they truly are. As a result, it recommended inadequate (and cheaper) treatments than required. This has various implications, but the most important consequence is the introduction (or rather reinforcement) of unfairness in healthcare and its effect on human lives.

Another example is when both Amazon and LinkedIn had to retract their recruiting algorithms as they exhibited a bias towards male candidates.

AI algorithms with inherent biases may pose various risks in the enterprise but they usually fall into three key categories:

1. **Operational:** The algorithm makes the biased decision which may be socially unfair, or simply costly for a business; In some cases, the data may encourage the algorithm to develop predatory "behaviour"

2. **Regulatory & Compliance:** In certain areas, e.g., insurance, non-compliance may result in fines and penalties

3. **Reputational:** Brands can be permanently damaged by a biased and unfair algorithm

Interestingly, regardless of the type of risk or industry, the resulting issues can be preemptively mitigated mostly by following a well-established, iterative process.

It's called Data Science

Despite the ambiguity of what Data Science is, its name enforces certain well-defined principles on the entire practice. Being *scientific* implies expectations which should not be negotiable, as e. In the era of "story-telling Data

Science", lack of *scientific* rigor may have gone unnoticed or have limited impact. However, when algorithms independently make millions of decisions every day, algorithmic predictions affect human lives or may translate directly into sizable financial loses or gains, deciding the fate of a business.

Science is not easy, and according to Richard Feynman, a celebrated physicist, *"The first principle is that you must not fool yourself — and you are the easiest person to fool"*.[3] In reality, Data Scientists are often under time pressure to deliver and deploy algorithms. This creates conditions which are not always ideal for the scientific process. It is important to remember that training a model which results in astonishingly high-performance metrics on the first iteration is more likely a sign of something gone wrong rather than reason to celebrate - testing hypotheses and interrogating the data to ensure nothing will teach the model to be biased is almost never straightforward.

The good news is that despite these real-world constraints Data Scientists face, it is not impossible to follow a scientific approach. As the realisation of algorithm deployment challenges became commonplace, so has a proliferation of tools which facilitate a more scientific workflow.

Data Scientists no longer have to build their own scaffolding and write all the boilerplate code to keep track of underlying data and its changes. Models can be developed while effortlessly keeping track of all parameters and performance metrics in purpose-built registries. From these registries, models can often be deployed with a few simple steps. Algorithm performance can be monitored and if the aforementioned structures are in place, identifying an issue is simplified by following the trail. Development and production environments can be easily controlled ensuring the results are reproducible anywhere.

"The first principle is that you must not fool yourself - and you are the easiest person to fool."

– Richard Feynman, Physicist

2   Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464):447–453, 2019.

3   Richard Feynman, "Surely You're Joking, Mr. Feynman!": Adventures of a Curious Character.

The tools however need to be used effectively. Data Scientists must be disciplined and dedicate ample time to understand and clean the data. The models should be trained carefully while choosing appropriate algorithms and optimised against more than a single metric or target group. It is crucial to not blindly maximise the performance metric of choice and naïvely ignore the idiosyncracies of the problem.

Assessing a ML algorithm development workflow under this lens and asking the question "is this a scientific process" is a first step towards identifying and evaluate bias-related risks. Ideally, this should be asked before a model is deployed and embedded into a fabric of automated processes.

However, the responsibility of what AI does, still lies with us. Although regulations will most certainly play a role in the coming years, organisations already can and should work to build better and ethically responsible AI. The rigor should be the same whether it's regarding decisions on healthcare or making consumer products recommendations. The lessons learned from striving toward a scientific approach in any use-case are transferable. AI is learning from us, via the data we generate. But is it learning from our mistakes or simply learning our mistakes is the question that should be always on top of mind.

The views expressed herein are those of the author(s) and not necessarily the views of FTI Consulting, its management, its subsidiaries, its affiliates, or its other professionals.

DIMITRIS KORRES
Managing Director
Data & Analytics
dimitris.korres@fticonsulting.com

MELORIA MESCHI
Senior Managing Director
Economic Consulting
meloria.meschi@fticonsulting.com

CLAUDIO CALVINO
Senior Managing Director
Data & Analytics
claudio.calvino@fticonsulting.com

FTI ™
CONSULTING