Evolving Role of Data Scientist in the Age of Personalization

Author, Lalith Kumar Vemali

A Data Science Foundation White Paper

March 2020

----

www.datascience.foundation

Authors: Arun Bharadwaj Dintyala, Lalith Kumar Vemali

This point of view is an exploration of the possibilities engendered by rethinking the role of data scientists in the wake of industrial revolution. It might be claimed that current trends in industrial revolution reflect a paradigm shift towards data centric processing with data science playing an increasingly critical role. This point of view also explicitly highlights the potential role of Data scientists as an emerging phenomenon, and then to show some of the benefits that this role can bring as we move towards industrial disruption.
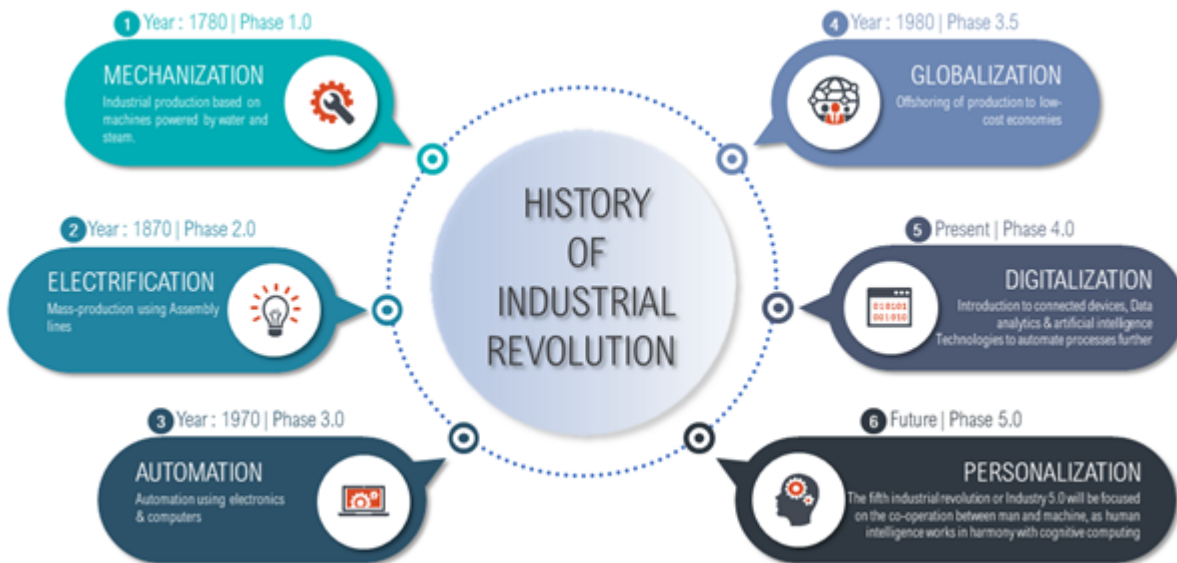
Industrial Revolution

Over a period, humans have evolved themselves by perfecting the industry with the help of technical evolution and by reinventing new resources which are more robust in nature thereby making the life easier.

Industry has benefitted from all these technically qualitative advancements which were so ingrained in a certain time and have had such an overwhelming impact that we dubbed them "revolutions". These revolutions are always important as they have always dominated and changed the world in a big way. Industrial revolution resulted in far reaching changes in the history of humanity and the influence of these changes continues to sweep us even today in our daily life's.

In short, the Industrial Revolution is the "**game changer**" of modern world history. More than anything else, it's what makes the modern world, well, "**modern.**"

History of Industrial Revolution

History of Industrial revolution can broadly be categorized into below given phases

Digitalization

Today Industry 4.0 is all about digitalization. The world is running on data and intelligent agents. This entire digitalization in the new world is driven by 4 important design principles of system. They are ---

- **Interoperability —** Machines, devices, sensors and humans need to seamlessly connect through an ecosystem of connected devices.
- **Information Transparency —** Sensors collect data about the physical world and use smart analytics to index this unstructured data to help with decision making.
- **Technical Assistance —** Systems support people in the decision-making process and assist with tasks that are difficult or unsafe.
- **Decentralized Decisions** Intelligent systems operate autonomously and can make simple decisions and resolve conflicts faster

All the above 4 design principles are related to **Data Analytics & Artificial Intelligence (DAAI)** which requires data driven insights.

Elements such as the **Digital Twins, Cognitive Computing, Industrial Internet of Things (IIoT)** and **Cyber Physical Systems** have aligned with DAAI as unseparated elements, providing the necessary ingredients for a paradigm shift.

Therefore, **data science** is becoming a guiding source for the industry to transform the information into knowledge using algorithms created in the pre-digital era and data Scientist plays a key role in adapting to these transformations.

Evolution of Data and Data Scientist along with Industrial Disruption

Year	Phase	Industrial Disruption	Knowledge of Data Science	Role of Data Scientist
1780	1.0	Mechanization	NA	Limited knowledge of Data
1870	2.0	Electrification	Manual computations	Manual evaluation of data
1970	3.0	Automation (Computer Era)	Microsoft Excel, Microsoft Access	Retrospective Analysis
1980	3.5	Globalization	OLAP, DataMart's, Data Warehouses	Analytical Decision Making
Present	4.0	Digitalization	Big Data, Data Science	Predictive analytics, AI
Future	5.0	Personalization	Self-Serving Analytics	Citizen Data Scientists

The data scientists are playing a fundamental role in the NEW to understand the enterprise activity. Nevertheless, a new challenge appears where the data engineers might face a challenge in the industry: they must deal with unknown processes, procedures, operations & science. A new set of multidisciplinary skills must be added to the data scientist expertise to provide the required added value in the specialized analysis for each challenge.

Recent survey conducted by Data Scientists and Analytics society has given a clear indication that data scientists are going to play a key role in all the industries. Below table gives an industry wise summary of data scientists utilization and this is only going to grow further in the organization as we move towards personalization (Industry 5.0).
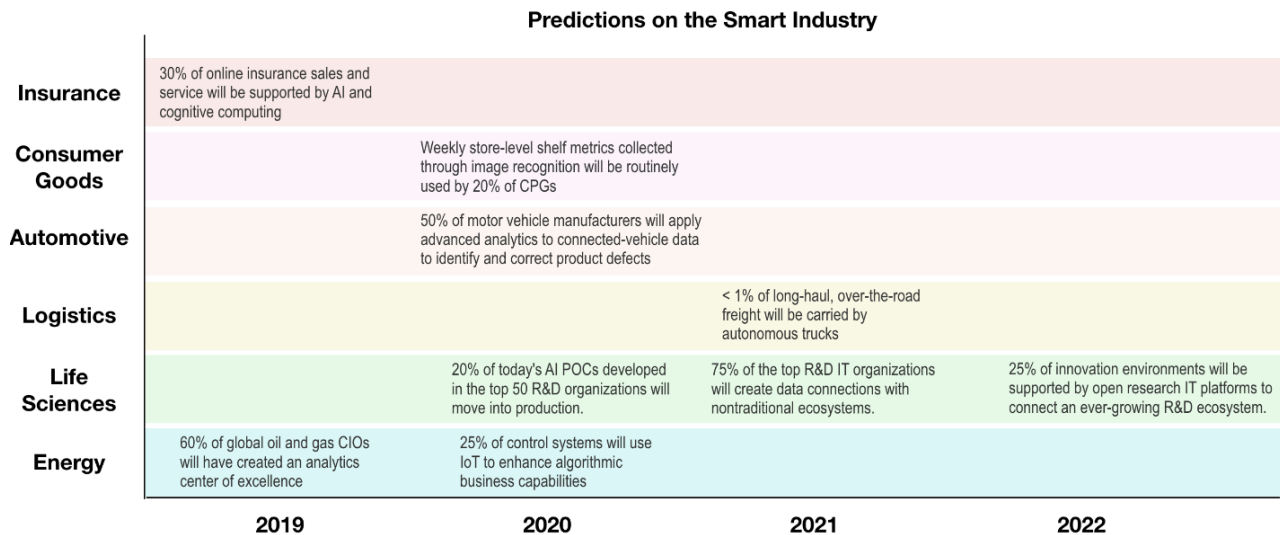
DSA Framework Category	Professional Services	Finance & Insurance	Manufacturing	Information	Health Care & Social Assistance	Retail Trade
Data-Driven Decision Makers	23%	17%	16%	10%	6%	6%
Functional Analysts	23%	34%	9%	5%	8%	4%
Data Systems Developers	41%	14%	14%	10%	5%	3%
Data Analysts	34%	25%	9%	6%	7%	3%
Data Scientists & Advanced Analysts	31%	23%	12%	10%	6%	4%
Analytics Managers	21%	41%	9%	9%	6%	3%

Source: [Forbes Article](#)

From the above table, we can clearly see that data scientists have already started their work on most of the industries and once they take over, the roles like data analysts, data driven decision makers and

---

Data Science Foundation

Data Science Foundation, Atlantic Business Centre, Atlantic Street, Altrincham, WA14 5NQ
Tel: 0161 926 3641   Email: admin@datascience.foundation Web: www.datascience.foundation
Registered in England and Wales 4th June 2015, Registered Number 9624670

functional analysts will slowly be taken over by data scientists. If we further deep pe, it is imperative that every industry needs their own specific skills to understand the data and reap the benefits out of it. So, industry should also have data scientists specific to line of business. Below chart shows the projections on how various industries are slowly moving towards **DAAI** soon
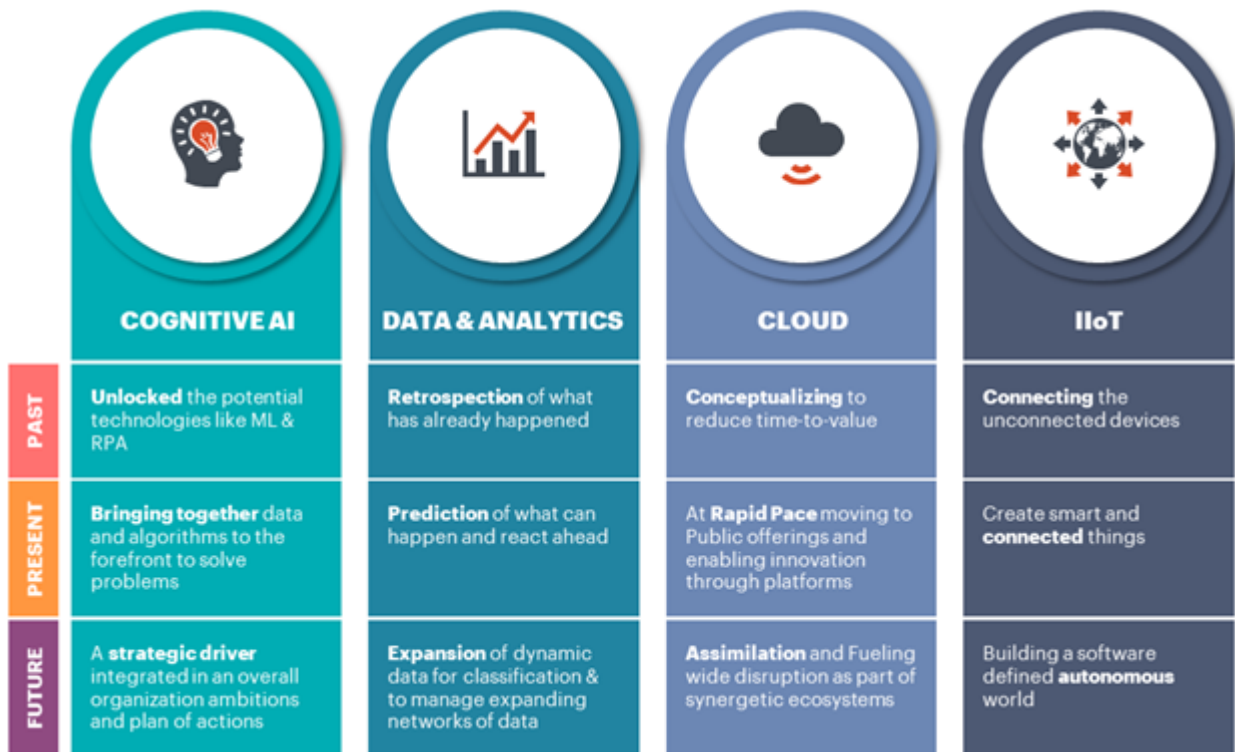
Predictions on the Smart Industry

	2019	2020	2021	2022
Insurance	30% of online insurance sales and service will be supported by AI and cognitive computing			
Consumer Goods		Weekly store-level shelf metrics collected through image recognition will be routinely used by 20% of CPGs		
Automotive		50% of motor vehicle manufacturers will apply advanced analytics to connected-vehicle data to identify and correct product defects		
Logistics			< 1% of long-haul, over-the-road freight will be carried by autonomous trucks	
Life Sciences		20% of today's AI POCs developed in the top 50 R&D organizations will move into production.	75% of the top R&D IT organizations will create data connections with nontraditional ecosystems.	25% of innovation environments will be supported by open research IT platforms to connect an ever-growing R&D ecosystem.
Energy	60% of global oil and gas CIOs will have created an analytics center of excellence	25% of control systems will use IoT to enhance algorithmic business capabilities		

Graphic 1. Predictions representation based on the report published by Gartner. Reference: Douglas L., Ankush J. 100 Data and Analytics Predictions Through 2021. 2017. Gartner, published on 20 June 2017. Report ID: G00332376

So, to align and grow with the industry, a data scientist should make himself more comfortable with the line of business he is working with. Here are few examples across each industry –

- **Banking** and **finance** sectors are extensively using data science for Customer segmentation, fraud detection, risk management and offering personal services - data scientist is helping them in achieving their goals. In future, data scientist 5.0 might further need to work towards algorithmic trading, automated risk management techniques, deep personalization and customization services
- **In retail** sector, a data scientist is already helping the customers by providing recommendation engines & location services. Data scientist 5.0 could further move towards market basket analytics, warranty analytics, price optimization for the customers. For the industry, they should think towards implementing automated inventory management, customer life time value prediction
- **In Travel & Transportation** sector, a data scientist is helping the travelers by doing predictive analytics of the dynamic pricing and fair costing, but data scientist 5.0 could further grow and upgrade towards travel bots, route optimization mechanisms and customer segmentation analysis resulting in minimizing travel costs, time & distance.
- **In Pharma & health care Industry,** a data scientist can blend data between discovery and clinical development to smash silos that previously caused clinicians and researchers to miss important elements. The advance predictive analytics & data driven drug trials enables data scientist 5.0 to curb the Skyrocketing Costs of Pharmaceutical R&D

Data Science Foundation

Data Science Foundation, Atlantic Business Centre, Atlantic Street, Altrincham, WA14 5NQ
Tel: 0161 926 3641   Email: admin@datascience.foundation Web: www.datascience.foundation
Registered in England and Wales 4th June 2015, Registered Number 9624670

Evolution of Technology Disruption

	COGNITIVE AI	DATA & ANALYTICS	CLOUD	IIoT
PAST	**Unlocked** the potential technologies like ML & RPA	**Retrospection** of what has already happened	**Conceptualizing** to reduce time-to-value	**Connecting** the unconnected devices
PRESENT	**Bringing together** data and algorithms to the forefront to solve problems	**Prediction** of what can happen and react ahead	At **Rapid Pace** moving to Public offerings and enabling innovation through platforms	Create smart and **connected** things
FUTURE	A **strategic driver** integrated in an overall organization ambitions and plan of actions	**Expansion** of dynamic data for classification & to manage expanding networks of data	**Assimilation** and Fueling wide disruption as part of synergetic ecosystems	Building a software defined **autonomous** world

If we look at the various disruptions which are happening at a technology level, we can clearly say that data is playing a pivotal role in all the technology disruptions. Data scientist should be more engaged with all the disruptions which are happening at a technology level along with the industry level.

The following table provides a deep insight as to how we expect the role of data scientist in the future

Past (Pre-Digitalization)	Present (Post-Digitalization)	Future (Personalization)
Backward Looking	Forward Looking	Exploratory Analysis
Slice & Dice Data	Interact with Data	Personalization with Data
Analyze the past	Predict the future	Expansion of dynamic data
Create reports	Create actionable insights	Data Optimization
Warehouse and siloed data	Distributed & Realtime Data	Quantum Data

What's ahead for Data Scientist's?

With GDPR and regulatory rules, it is even more important that the laws of countries must be adapted and thus, appear a new generation of data scientist and technological experts with a clear orientation for regulatory advising or legislators consultancy, initiating the path to a consensual data management.

Today, data scientists 4.0 already are aware of the algorithms to apply on top of the data ecosystem and this fact requires to be multidisciplinary and specialized at the same time. They are already working towards growing demands of data science services applied across industries.

In future, when industry moves to personalization (Industry 5.0), it is imperative to understand that the skills they have now may not meet the then organization goals and thus they should have additional skills on top of the skills they have today. They should have extensive industry specific knowledge to understand the Line of Business rather than having only knowledge on algorithms implementing applied intelligence. They should understand that Industry 5.0 is more about Human-Machine collaboration and should be ready to accept the fact and ensure that human emotions work with cognitive computing in the new world.

Data Scientist 5.0

Key enablers to be a Data Scientist in the personalization era –

Data science will get democratized

It will become more democratized with greater automation, platforms with embedded solutions and easy to use tools. This will enable the non-specialists do significant model building, analysis and prediction using DIY kits. These citizen data-scientists will affect a substantial decentralization of the analytics function within the organization. The data scientists will devote their time and energy to create more advanced techniques that lead to better and faster solutions.

Data science will evolve to accommodate specializations

Data Science is all about statistics and math's. As technologies grew broader and deeper, data science will become more and more specialized, we expect that the superman data scientist will give life to several forms — specializations across techniques and/ or domains like Risk Data Scientists and Cyber Security Data Scientists just to name a few. We will also have the Data Visualizer, the Machine Learning expert, the AI star, and data scientists who are skilled at providing solutions to specific business problems.

Industry / Domain orientation of data scientists

Data scientist will move far closer to business than today. Once a good number of key algorithms get templatized and menu driven, the focus of the data scientist will shift to the business side of things. Data Scientists will enter an era of sharp, function or vertical focused applications — for example, Govt of AP is providing DWCRA (Development of Women and Children in Rural Areas) assistance to all the women to empower them and grow towards small scale industry entrepreneurship and amounts are disbursed over a period, based on the progress they have made. Can we develop an application to help Govt to identify the ability of these people to pay back after a period of duration? This will require deep knowledge of their financial, socio-economic conditions and understanding of all the external / internal environmental factors or variables that may impact a person's ability to pay.

The next data scientist will have to be a Data Pundit.

Deeper domain knowledge and availability of a much wider source base for data will drive personalization. In the DWCRA example, data on financial and socio-economic conditions over a period of time will be critical to assess the financial health of the beneficiary according to external/internal environmental factors, given the amount of benefits she receives from the Govt.

Is the Data Scientist ready for 5.0?

The future looks bright and exciting for data scientists of tomorrow. But the future starts today since most of the technologies mentioned above are available and data scientist should now look out for new disruptions to act as an interface between technology and Industry.

In this continuous evolving DAAI powered world, data scientists are definitely on the better side of the deal. Data scientists do a varied set of jobs to help businesses grow. Few activities like data governance framework, data cleansing, data visualization and diagnostic analysis with set frameworks can be automated. There are certain activities like identifying the right business problem, Tweaking ML techniques to fit the problem, creatively leveraging solutions from cross domains and creating business strategy from analytics which surely requires data scientist's intervention. With good certainty it can be said that even though the data science stream will stay up and running for long term, the roles and responsibilities of these jobs are up for big changes. Data Scientists who have challenges upgrading to these new roles and responsibilities will face strong setbacks in progressing in career.

Considering the above, Data Scientists across the globe are not yet ready for Industry 5.0. In order to reach there, Data Scientists need to upgrade himself on domain / industry specific skills and also evolve in specialized roles like - Risk Data Scientists and Cyber Security Data Scientists etc.

The data scientist of 2022 will have to be a **data pundit** – not from a technology perspective but from the perspective of understating, identifying and accessing the right sources of data and bringing it all together to enrich the analysis.

About the Data Science Foundation

The Data Science Foundation is a professional body representing the interests of the Data Science Industry. Its membership consists of suppliers who offer a range of big data analytical and technical services and companies and individuals with an interest in the commercial advantages that can be gained from big data. The organisation aims to raise the profile of this developing industry, to educate people about the benefits of knowledge based decision making and to encourage firms to start using big data techniques.

Contact Data Science Foundation

Email:admin@datascience.foundation
Telephone: 0161 926 3641
Atlantic Business Centre
Atlantic Street
Altrincham
WA14 5NQ
web: www.datascience.foundation