DATA SCIENCE ALLIANCE

White Paper

Guiding Principles Of Responsible Data Science

How to Build a World Where Data Helps Without Harm

Authors

Adir Mancebo Jr., Ph.D.

Data Analyst, Data Science Alliance

Bradley Voytek, Ph.D.

Professor, Cognitive Science,
Data Science, and Neuroscience, UC San Diego
Founding Board Member, Data Science Alliance

Ilkay Altintas, Ph.D.

Chief Data Science Officer,
San Diego Supercomputer Center
Founding Board Member, Data Science Alliance

Di Le

Lead AI/ML Design Strategist, ServiceNow
Working Group Member, Data Science Alliance

Contributors

Michael Limbo, Czarina Argana, Tyler Tran, and Jaemin Ko

Table of Contents

1.0 Introduction

In the 18th century, the First Industrial Revolution used water and steam power to mechanize production, igniting a major societal change. In the subsequent centuries, the Second and Third Industrial Revolutions continued to transform human life as they introduced electric power, electronics, and information technology to expand and automate production [30]. However, those impacts pale in comparison to the potential consequences of the advancements in digital technologies, process automation, workflow management, and Artificial Intelligence (AI) which characterize the Fourth Industrial Revolution. Klaus Schwab, the founder and executive chairman of the World Economic Forum, argues that "we stand on the brink of a technological revolution that will fundamentally alter the way we live, work, and relate to one another." (cont.)

"In its scale, scope, and complexity, the transformation will be unlike anything humankind has experienced before." [30]

The Fourth Industrial Revolution is one powered by astronomical amounts of data, enabling the rapid progress of data science and the evolution of AI systems. In parallel, the Fourth Paradigm marks the transformation of the scientific discovery process with the influence of data, computing, and AI [13]. A striking aspect of this period is that there is no historical precedent for the speed of current advances and the pervasiveness of these evolving technologies affecting entire systems of production, management, innovation, and governance. Furthermore, it is reshaping every industry globally [30].

The world faces a vast potential for great benefit and improvement to human life, but concurrently the risks for lasting negative consequences due to the irresponsible use of data are tremendous. There is a rapidly growing need to expand the development and reach of responsible data science (RDS) values, and to establish an ecosystem in which everyone involved with data science acts responsibly. The Data Science Alliance (DSA) was born from this call to action. Our vision is "a world where data helps without harm" in which the large-scale impact of data science is realized ethically. We are pursuing this goal by: (i) building a common set of principles and framework for RDS practices; (ii) engaging with the data community and its students from academia to professional settings; and (iii) working in partnership with mission-driven organizations to execute RDS projects for societal good. Four core principles are the foundation of our work: Fairness, Transparency, Privacy, and Veracity.

In this white paper, we will:

1. Evidence the importance of RDS
2. Educate the public about its four guiding principles
3. Demonstrate why these principles are important to advance our vision and mission.

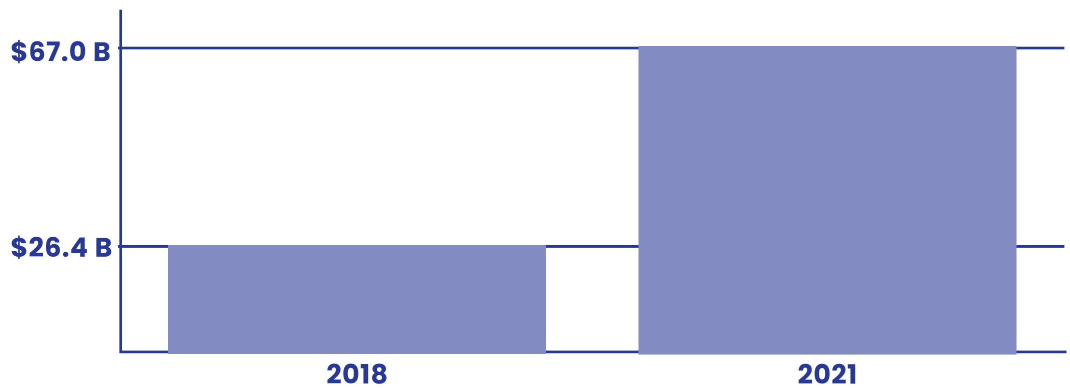
2.0 The Era of Data

In the past two decades, the exponential expansion in the amount of data produced and stored by companies, government institutions, and the Internet of Things, along with the increasing ability to dynamically process such data, has enabled a revolution in how organizations operate. What is commonly known as “big data” has fueled a new economy. Just as oil power changed the world in the 20th century, big data allows the development and implementation of new technologies that are transforming the way we live [1]. At the forefront of these advancements is AI, a subdiscipline of data science that aims to mimic human cognition, learning, and behaviors through data. Modern AI algorithms rely on innumerable amounts of data to operate and significantly improve their capabilities as more data is fed into their systems [12].

In the same manner that natural resources can be treated as a form of currency, human data has become a new currency in today’s economy. However, most people don’t know how to control their own data or what they’re giving up in return for daily tech conveniences. The terms and conditions of personal data management are usually written in very complicated and confusing language, sometimes intentionally.

The pervasiveness of AI algorithms keeps growing rapidly. According to Gartner Research, the number of organizations reporting that they either already used AI technology, or would be doing so in the near future, jumped from 10% to 37% between 2015 to 2019 [9]. The worldwide pandemic that started in 2020 profoundly affected many business operations, increasing the demand for automated solutions [15]. McKinsey's Global AI Survey reported that 50% of companies were already adopting AI in at least one function by the end of 2020. This share increased to 56% in 2021 and nearly two-thirds of respondents said their companies' investments in AI would continue to rise over the following three years [5]. Moreover, CB Insights reported that global AI funding increased almost 154% between 2018 and 2021, from $26.4 billion to $67.0 billion [4].

Global AI Funding

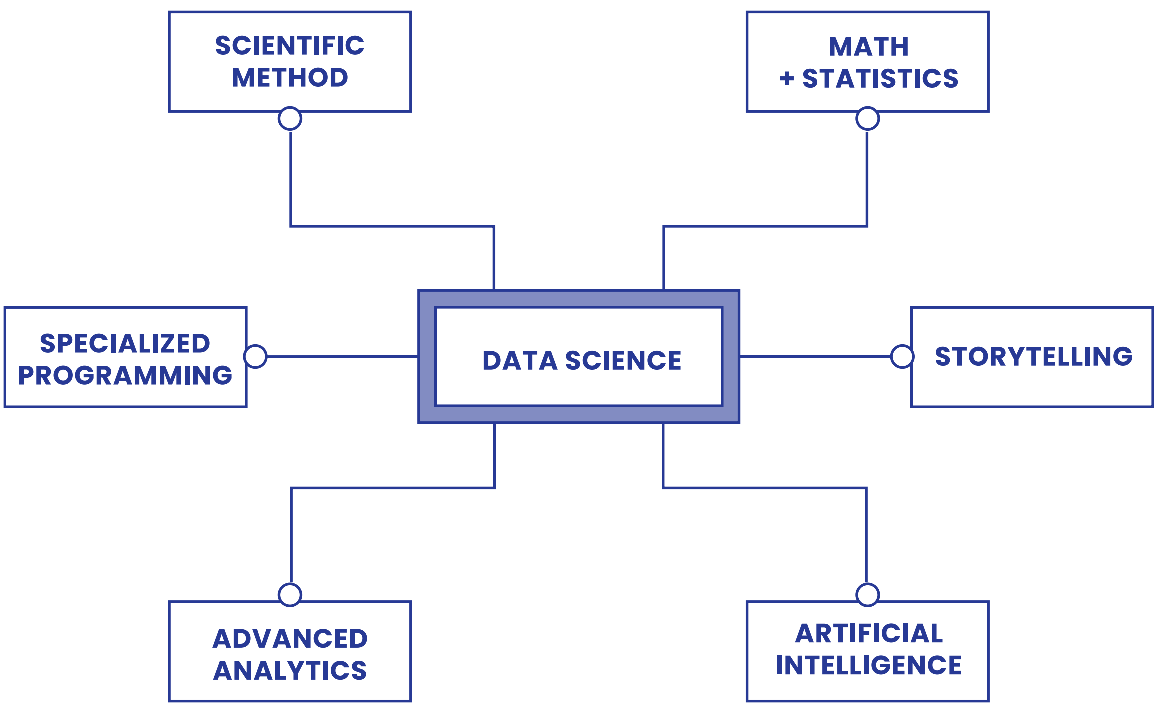


Source: State of AI - Q2 2022, CB Insights

Recent job market trends reflect the increasing investment in AI technology. Linkedin's 2020 Emerging Jobs Report listed "AI Specialist" as the #1 emerging job, with an annual growth rate of 74% [19]. This steadfast growth is generating demand for a specialized workforce to facilitate the research, development, governance, and management of AI systems. In 2021, the "machine learning engineer" position was the single fastest-growing job worldwide [14]. In the coming years, AI applications will only become more ubiquitous and their impact on our lives increasingly more powerful. Hence, it is imperative that these new technologies are designed, developed, and used responsibly—in a way that puts humans at the center of it all.

2.1 A Call for Responsible Data Science

With proper AI investments and the adoption of correct AI systems and applications, organizations can support and elevate a workforce to higher value tasks. AI can aid in reducing redundancies, provide business insights, and create enhanced engagements and experiences for companies and their customers. As such, AI systems affect a large spectrum of people with potential benefits and possible negative outcomes without proper attention to what might happen.



In most of its current business applications, we can regard AI as one element of the broad data science category. According to IBM's definition, "data science combines the scientific method, math and statistics, specialized programming, advanced analytics, AI, and even storytelling to uncover and explain the business insights buried in data." [16] Therefore, although it is adequate to attribute to AI systems much of the progress related to the data revolution, the impact of data science applications in society goes beyond automated models. Moreover, the development of those technologies involves not only AI specialists and data scientists but many other contributors from researchers (UX and applied sciences) and engineers to developers, product owners, and UX designers.

Projects based on data science methods have brought visible positive impacts to the world. They provide insights, diagnostics, predictions, and automated decisions that have improved the lives of millions (see [37] for a list of examples across different sectors). These improvements range from the commonly experienced, such as personalized movie or shopping recommendations, to the life-changing, such as the improved identification and prediction of diseases that lead to more individualized medical treatments. However, the irresponsible use of data can also produce serious negative externalities. It can reinforce systemic discrimination and bias, reaffirm economic and social inequities, violate privacy, and produce non-transparent consequential decisions that further harm marginalized or susceptible groups.

Companies and governments have overlooked these damaging impacts of irresponsible data science for many years. In fact, the United States still lacks governance and guidance around the responsible use of data in comparison to other parts of the world like Singapore and the European Union (see, e.g., [27] and [8]).

In 2019, researchers found that a health care risk-prediction algorithm used on more than 200 million people in U.S. hospitals strongly favored white patients over black patients [26]. The issue was only mitigated by the company providing the service because an external investigation was conducted and uncovered the problem.

With the pervasiveness and scale of data usage for AI systems, through mishaps, whistleblowers, and internal leaks, transparency into nefarious and irresponsible data practices has been receiving more media attention. The harmful effects of the irresponsible use of data have been disseminated to the public (e.g., [2]), and pressure is mounting forcing corporate entities to address the matter seriously. Moreover, many of those participating in data science projects, especially when the projects involve AI, recognize that the risks associated with their work demand great responsibility. Despite this general awareness, when confronted with the possibility that their data-driven, AI-based tools can be used for harm, researchers have expressed surprise. For example, researchers did not expect the results when an AI-powered drug-discovery platform was “inverted” by its creators to instead develop maximally harmful toxic agents [36].

In recent years, companies like Google, Microsoft, LinkedIn, and OpenAI have established initiatives to define and guide the production of responsible AI pipelines. Though the movement towards the diffusion of RDS is needed and celebrated, it is still in its infancy. The number of corporations and government institutions actively practicing and promoting RDS is small. Academic efforts like the Association for Computing Machinery (ACM) conference on Fairness, Accountability, and Transparency (FAccT) are relatively nascent. Since these initiatives are independent of one another, there is no unified definition of what “responsible” means for data science. Besides, most actions only focus on AI systems, which may be the most prominent part of data science due to the scope of its use, the complexity of its models, and the amount of data it uses, but it is not all-encompassing.

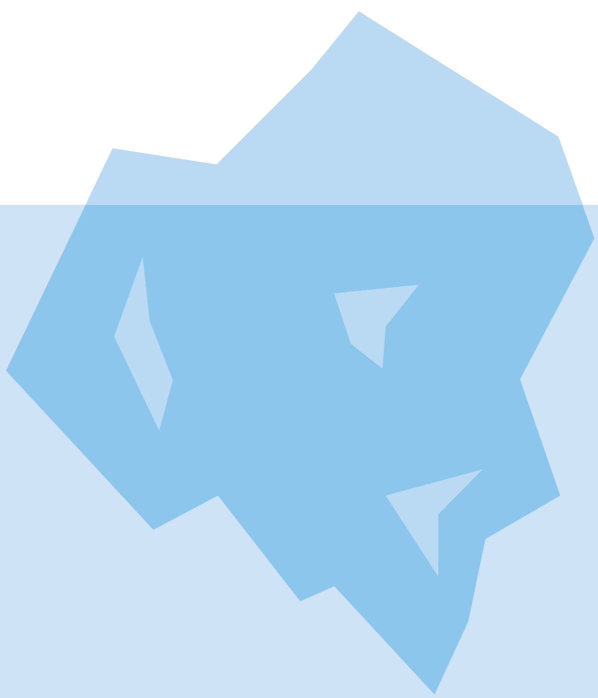
3.0 Guiding Principles of Responsible Data Science

The responsibility in data science is not limited to how we build a model or system. Data science is a multi-step activity that involves acquiring, cleaning, modeling, and preparing data before any analysis happens. Communication of insights is also an important step that involves visualization, summarization, and presentation, often through a digital environment. Every step of the data science process has to be developed with safeguards, human oversight, and consistent evaluation. From acquiring data to delivering the final product using the results, which are data themselves, every action requires careful consideration [29]. Negative consequences can arise from any phase of the process. More subtly and perniciously, the most well-intentioned model can still result in detrimental outcomes if it is supplied with biased or faulty data through a non-repeatable and unexplainable process.

Data Science Workflow

20% Model Development

80% Data Preparation



Accordingly, in the past few years, scholars and practitioners of data science have been working on identifying, understanding, and elaborating on solutions for the harmful uses of data across different segments of society [18]. This has led to the recent emergence of a number of academic and corporate initiatives centered on the dissemination of RDS practices.

The principles that characterize RDS across these groups may vary depending on how they define the responsible use of data. They include among others: fairness, privacy, veracity, transparency, accuracy, reproducibility, repeatability, reusability, interoperability, accessibility, reliability, safety, security, inclusiveness, accountability, equity, diversity, and equality [6, 7, 10, 22]. After a series of discussions involving the founding board members, academic experts, and seasoned data science professionals, the DSA has chosen to focus on the first four of the mentioned principles. We understand them to be the core foundation of RDS. In the following sections, we will endeavor to define and discuss each principle.

Fairness

Understanding that biases exist and taking into account that technology has the potential to harm.

Transparency

Making methodology and practices broadly available for public scrutiny.

Privacy

Ensuring that we answer questions without exposing sensitive information.

Veracity

Assuring that our methods have integrity in their scope and accuracy.

3.1 Fairness

A responsible practitioner of data science must understand that biases exist. Sometimes, biases are preferred and intentionally included in systems to perform specific tasks. For instance, there are products created to bias toward detecting defects in manufacturing. However, biases can be extremely harmful as well. We live in an unfair world in which the human language and actions are biased along with the data collected about them. AI models learn from data and if we are not paying attention, algorithms can discriminate just like humans do—but on a larger scale. RDS requires that the systems we build should be fair and equitable. They should not persist or worsen the present inequities in our society.

A challenge is defining fairness when its meaning varies across individuals from diverse cultural, social, and economic backgrounds. According to [24], there are at least 21 definitions of fairness. Hence, although it is crucial that data systems must not be biased against one single person or group of people, it will not be easy to identify a set of appropriate fairness criteria to follow. As such, each project must take into consideration the context surrounding it. Here, we suggest a few ways to improve fairness.

Fairness demands diversity

Fairness reaches far beyond the technical components of data science. Because of its relation to the cultural, social, historical, political, legal, and ethical contexts in which an entity deploys the product, only a diverse team actively involved in every step of the system development can achieve fairness [10]. This is crucial since processes can incorporate bias at any stage. Hence, everyone involved should be actively thinking about fairness.

Fairness evaluation needs to be a part of every step

Fairness starts at the very beginning of any project. Even in the product research and design stage, the team must implement actions to mitigate biases. For instance, when evaluating a design or idea, or researching a product idea, it is important that the team interviews people that are a representative sample of the intended product demographic. They should include people from various socioeconomic and cultural backgrounds, diverse abilities, and technical literacy as well.

Biased data needs to be eliminated

Data professionals must carefully evaluate the data they are using. Technically sound and accurate models can deliver unfair outcomes when trained on data that are biased against groups based on their gender, race, religion, or other characteristics. For example, [31] found that machine learning models trained on two large, publicly available image datasets, ImageNet [28] and Open Images [17], produced Americentric and Eurocentric bias in their results. As such, it is paramount for RDS to ensure that the data used to build a system is not biased and that those involved in the process correct the biases within a model throughout its development.

Fairness should be evaluated in every context

A great challenge to fairness in data science is the confidence that a system will be fair in every context it is employed [10]. It is possible that a model thoughtfully developed to deliver fair outcomes in one specific context, may not be fair when used in another, i.e., for a determined country or for a particular age group. Moreover, we must consider that these unintentional biases may only appear after a system is deployed. This demonstrates the importance of continuously monitoring outcomes and acquiring feedback from users.

3.2 Transparency

Government agencies and corporations rely heavily on data science methods to make decisions that can affect the lives of everyone. However, the processes underlying such actions are often not available for public scrutiny. For instance, [3] submitted 42 requests for information about algorithms that were used by various government agencies to make decisions regarding citizens' lives. Almost all the entities involved either declined or were unable to provide any information that would elucidate how these tools were generating their results. Here, we suggest a few ways to improve transparency.

Methods should be transparent

Transparency is not limited to the governmental sector; nor is it only a matter of hesitation to share the work. Data science techniques are evolving rapidly and are becoming more complex. Machine learning experts have difficulty trying to explain how some AI models, such as deep neural networks, are giving the results and predictions that they do. Still, every effort must be made to provide a clear concept of the algorithmic structure.

Motivations and limitations should be clear

Transparency goes beyond the understanding of how data scientists build systems and models. It applies to the other dimensions of data science projects. While it is important that people understand the behavior of the applications affecting them, or every component of the process, they should also be able to comprehend the data feeding the models, the reasons why those data systems were developed, and the limitations of their use.

Data should be transparent

Without the transparency of data, we cannot achieve meaningful transparency of algorithmic processes [32]. A machine learning model is heavily influenced by the data it is trained on. Therefore, one cannot consider a system fully transparent if the algorithm is open to the public but the data behind the results are concealed. When there is a clear tension between data privacy and transparency, data science practitioners can implement alternatives to ensure openness such as releasing privacy-preserving synthetic datasets.

Everyone must be transparent

RDS requires transparency in every step of the process, from the decision to build to the recommendation to use the outcomes. Everyone involved in data science applications must be diligent in ensuring that their methods and practices are broadly available for public scrutiny. Transparency is essential to being able to question, understand, and trust data science products.

3.3 Privacy

The Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights, among many other international and regional treaties, declare privacy as a fundamental human right [34, 35]. Hence, there can be no RDS without the assurance that professionals working with data will not make sensitive information public. Here, we suggest a few ways to improve privacy.

Data privacy should be regulated by law

The issue of data privacy is so relevant that many countries around the world have enacted laws to regulate the use of personal data. The most prominent is the General Data Protection Regulation (GDPR) of the European Union [8], which confers data rights to data owners and directs how companies can use that data. Several states in the U.S. have created their own regulation based on the GDPR, but the U.S. federal government has yet to enact a data privacy law with that scope [21].

Privacy matters from the beginning

Privacy concerns must be addressed in all phases of the data science and data product workflow, from interviewing individuals in market and competitive research, to UX research and validating the designs and ideas of products. It is imperative that organizations put measures in place to prevent sensitive data from being accidentally disclosed to the public and also shared with team members that have no need to access that information.

Unintentional breach of privacy is a risk

When developing a system, data science practitioners may inadvertently reveal sensitive data in the outcomes of the model they are using. It is extremely important to consider the information at hand and create safeguards to protect privacy. This is especially true for projects involving machine learning models that have the capability to remember details of the data they have processed.

Privacy requires rigorous security measures

Responsibility regarding data privacy does not lie solely with the data science team. It also requires rigorous data security. Organizations must regularly update measures to secure data storage and access. As such, companies and public entities must ensure that data is well protected against the most common attacks on the internet. Whenever possible, data should be fragmented and stored in multiple locations. This ensures that should one specific site be compromised, the complete picture of what the data represents is not revealed [20]. Moreover, every employee should have to participate in regular cybersecurity training.

3.4 Veracity

The transparency of the algorithmic processes we use for decision-making is essential for RDS practice; but it is not a guarantee that the outcomes of the systems are correct. The knowledge to understand a transparent model is not yet standardized or widespread. Therefore, to be responsible, data science professionals must ensure that their methods have integrity in their scope and accuracy in their results.

RDS is not postulating or ad hoc. Accurate systems begin with a solid definition of the problem. They also demand serious data investigation combined with rigorous statistical analysis. While it is not expected that everyone involved in the development of a data application will know every detail and every method employed, it is crucial that the professionals working in each step understand that phase well and are held accountable for the outcomes. Here, we suggest a few ways to improve veracity.

Veracity requires good documentation

The smallest mistakes can pose great impacts and dangers in big systems. Errors can potentially snowball as the product is used many times across large groups of people. That is why meticulous documentation is so important to ensure veracity; it provides a guide to finding errors and omissions. It is also important to always clarify the uncertainty of results for statistical analyses [1]. For AI systems, the use of Model Cards is recommended. It is a framework developed "to clarify the intended use cases of machine learning models and minimize their usage in contexts for which they are not well suited" [23]. It's a due diligence item that practitioners can implement to increase veracity and transparency early on.

There are challenges to ensuring veracity

The ad hoc nature of model development and error analysis is definitely a challenge to ensure veracity. In data analysis and modeling, it is often necessary to make assumptions and simplifications in order to make progress. These decisions are usually based on the specific data set and problem at hand and are not always easy to generalize. This can lead to errors and inaccuracies in the results. Therefore, it is important to be flexible and to continually revise and improve the models and methods as new information and insights are gained. Moreover, the attempt to analyze when, how, and why models fail is a crucial part of the development cycle, although error analysis is often complicated, uncharted, and unsystematized.

Veracity needs to be prioritized in data collection and manipulation

Veracity also refers to the trustworthiness and the manipulation of the input data [1]. For instance, a model built to analyze population sentiment using Twitter posts may deliver inaccurate results due to the presence of bot-written posts in the data. Furthermore, if a professional is not rigorous when dealing with missing data in a dataset, the results of the model may be far from the truth.

Inaccurate systems are dangerous

Veracity cannot be neglected. In some instances, an inaccurate system can be a direct threat to human life, e.g., self-driving cars and disease diagnostic algorithms. If we want to avoid harm, we must strive for robust systems that are auditable at every step, and that are constantly reevaluated to verify that no unintended consequences have occurred since their implementation.

4.0 DSA in Action

To foster an RDS ecosystem and build a world where data helps without harm, guiding principles are necessary but not sufficient. We need concrete measures to fulfill our mission. Therefore, the DSA is directing its course of action to three segments:

1	A pledge and common Framework for RDS practice.
2	Engagement with the data science community to promote RDS.
3	Partnership with mission-driven organizations to execute RDS projects for the social good.

RDS Pledge and Framework

There are a number of initiatives proposing guiding principles for RDS, but there are few specific guidelines for how to put those principles into practice. The DSA aims to fill that gap by creating an RDS framework that is simple, comprehensive, and publicly available for any individual professional or company to use.

This living document will be regularly updated with feedback from users and adapted to any innovations that may emerge. We also developed a pledge for RDS based on our four guiding principles. This document is displayed at the end of the White Paper. Everyone that shares our values and is committed to using data with responsibility is invited to sign the pledge on our website.

RDS for Societal Good

Many nonprofit organizations lack the resources to develop data science projects that could improve their operations and provide more help to the communities they serve. Hence, the DSA is committed to partnering with such institutions to implement RDS in practice to exhibit data science as a great force for the societal good.

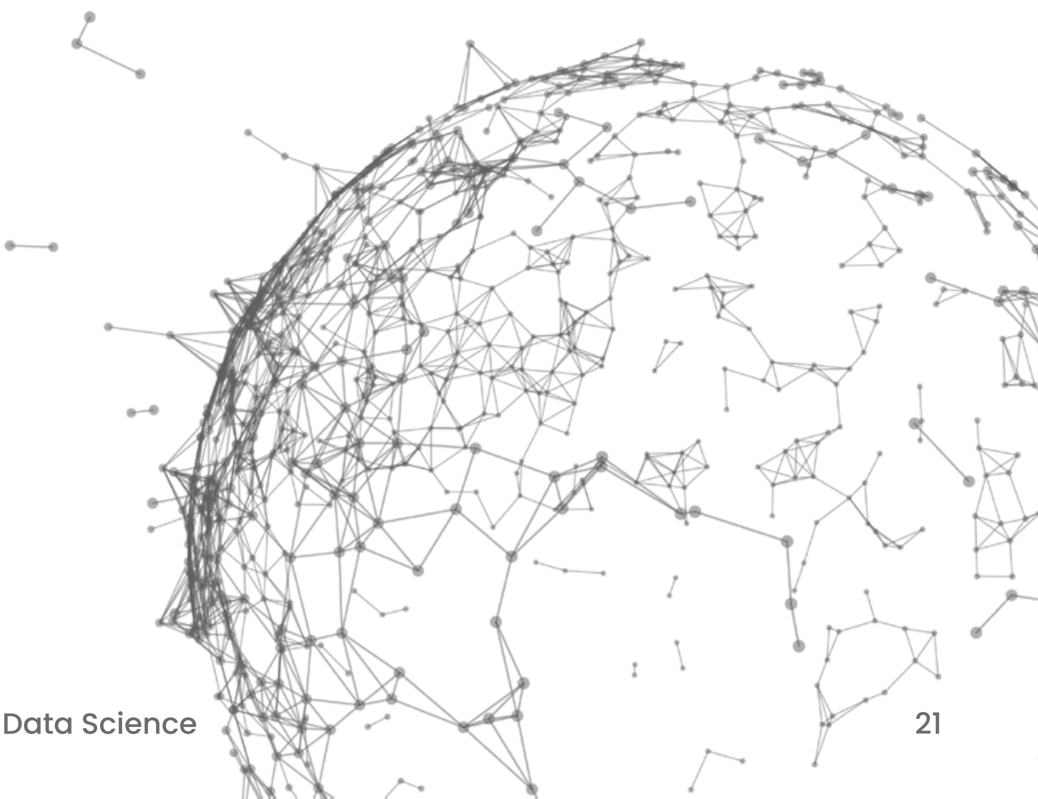
Community Engagement

Isolated projects are not enough to change the way data science is practiced across the world. DSA fosters a culture of RDS through active engagement with the data science community to build an RDS ecosystem from within. DSA brings together experienced practitioners from different companies, accomplished members of the academic community, and most importantly, data science students who will carry the torch for the future of RDS.

5.0 Conclusion

We all have an obligation to uphold data science that is responsible. A world where data helps without harm is an achievable goal, and the DSA is working diligently to accomplish it. Guided by the four principles of Fairness, Transparency, Privacy, and Veracity, we are building an RDS framework, partnering with nonprofit organizations, and engaging with the data community to cultivate an ecosystem where RDS can be a reality.

DATA SCIENCE ALLIANCE



DATA SCIENCE ALLIANCE

Pledge for Responsible Data Science

The Data Science Alliance is committed to building a community of responsible Data Science practitioners, and creating a framework to support responsible practices. The following values are what guide us.

- **I aim to provide a positive impact to our peers, community, and society as a whole, while minimizing harm, discrimination, or inequities that would result as a consequence of my work.**
- **I respect privacy and will adhere to localized laws and governance of personal information.**
- **I am rigorous, scientic, and strive to ensure my work is accurate and reflective of the truth to the best of my ability.**
- **I am transparent and will make my intentions clear with how I use personal information. I am proactive in communicating when there are risks to the above.**

In joining the Data Science Alliance, I expect our members to have the same principles and values, and provide role-model leadership that demonstrates responsible data science practice. Together, our members preserve these principles as we build a diverse, collaborative, and inclusive community of responsible data users.

References

[1] Wil van der Aalst. 2016. Responsible Data Science: Using Event Data in a “People Friendly” Manner. In International Conference on Enterprise Information Systems, 3–28.

[2] Julia Angwin, Lauren Kirchner, Jeff Larson, and Surya Mattu. 2016. Machine Bias. ProPublica. Retrieved June 10, 2022 from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[3] Robert Brauneis and Ellen P. Goodman. 2018. Algorithmic transparency for the smart city. Yale JL & Tech. 20, (2018), 103.

[4] CB Insights. 2022. State of AI.

[5] Michael Chui, Hall Bryce, Alex Singla, and Alex Sukharevsky. 2021. McKinsey Global Survey: The state of AI in 2021.

[6] Data, Responsibly. 2022. Responsible Data Science Principles. Retrieved from https://dataresponsibly.github.io/

[7] DSCE Research Program. Responsible Data Science. Retrieved June 10, 2022 from https://www.tue.nl/en/research/research-areas/data-science/responsible-data-science/from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

[8] European Parliament. 2016. General Data Protection Regulation (GDPR). Retrieved July 25, 2022 from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

[9] Gartner Research. 2019. 2019 CIO Survey: CIOs Have Awoken to the Importance of AI.

[10] Google AI. Responsible AI practices. Retrieved June 10, 2022 from https://ai.google/responsibilities/responsible-ai-practices/

[11] Thilo Hagendorff. 2020. The ethics of AI ethics: An evaluation of guidelines. Minds and Machines 30, 1 (2020), 99–120.

[12] Alon Halevy, Peter Norvig, and Fernando Pereira. 2009. The unreasonable effectiveness of data. IEEE Intelligent Systems 24, 2 (2009), 8–12.

[13] Anthony JG Hey, Stewart Tansley, and Kristin Michele Tolle. 2009. The fourth paradigm: data-intensive scientific discovery. Microsoft Research Redmond, WA.

[14] Laura Hilgers. 2022. The Fastest-Growing Jobs Around the World. LinkedIn Talent Blog. Retrieved July 6, 2022 from https://www.linkedin.com/business/talent/blog/talent-strategy/fastest-growing-jobs-global

[15] IBM and Morning Consult. 2021. Global AI adoption index 2021.

[16] IBM Cloud Education. Data Science. IBM. Retrieved July 6, 2022 from https://www.ibm.com/cloud/learn/data-science-introduction#:~:text=Data%20science%20combines%20the%20scientific,business%20insights%20buried%20in%20data

[17] Ivan Krasin, Tom Duerig, Neil Alldrin, Vittorio Ferrari, Sami Abu-El-Haija, Alina Kuznetsova, Hassan Rom, Jasper Uijlings, Stefan Popov, and Andreas Veit. 2017. Openimages: A public dataset for large-scale multi-label and multi-class image classification. Dataset available from https://github.com/openimages (2017).

[18] Armanda Lewis and Julia Stoyanovich. 2021. Teaching responsible data science: Charting new pedagogical territory. International Journal of Artificial Intelligence in Education (2021), 1–25.

[19] LinkedIn. 2020 Emerging Jobs Report. Retrieved from https://business.linkedin.com/content/dam/me/business/en-us/talent-solutions/emerging-jobs-report/Emerging_Jobs_Report_U.S._FINAL.pdf

[20] Luna. 2022. The Role of Cybersecurity in the Management of Data Privacy . Retrieved June 10, 2022 from https://www.lunadna.com/data-security-privacy/

[21] Luna. 2022. The Evolution of Data Privacy: A Q&A with Bojil Velinov, Head of DevOps & Automation at Luna. Retrieved June 10, 2022 from https://www.lunadna.com/evolution-data-privacy/

[22] Microsoft. Microsoft responsible AI principles. Retrieved June 10, 2022 from https://www.microsoft.com/en-us/ai/our-approach?activetab=pivot1:primaryr5

[23] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model cards for model reporting. Proceedings of the conference on fairness, accountability, and transparency (2019), 220–229.

[24] Arvind Narayanan. 2018. Translation tutorial: 21 fairness definitions and their politics. In Proc. Conf. Fairness Accountability Transp., 3.

[25] Peter Naur. 1974. Concise survey of computer methods. Petrocelli Books.

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

[27] Personal Data Protection Commission. 2020. Model Artificial Intelligence Governance Framework. Singapore. Retrieved July 25, 2022 from https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.ashx

[28] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, and Michael Bernstein. 2015. Imagenet large scale visual recognition challenge. International Journal of Computer Vision 115, 3 (2015), 211–252.

[29] Monica Scannapieco and Antonino Virgillito. 2021. How to be responsible in all the steps of a data science pipeline: The case of the Italian public sector. Patterns 2, 12 (2021), 100393.

[30] Klaus Schwab. 2017. The fourth industrial revolution. Currency.

[31] Shreya Shankar, Yoni Halpern, Eric Breck, James Atwood, Jimbo Wilson, and D Sculley. 2017. No classification without representation: Assessing geodiversity issues in open data sets for the developing world.

[32] Julia Stoyanovich. 2019. TransFAT: Translating fairness, accountability, and transparency into data science practice. In CEUR Workshop Proceedings.

[33] Julia Stoyanovich, Bill Howe, Serge Abiteboul, Gerome Miklau, Arnaud Sahuguet, and Gerhard Weikum. 2017. Fides: Towards a Platform for Responsible Data Science. In Proceedings of the 29th International Conference on Scientific and Statistical Database Management, 1–6.

[34] UN General Assembly. 1948. Universal Declaration of Human Rights. UN General Assembly 302, 2 (1948), 14–25.

[35] UN General Assembly. 1966. United Nations International Covenant on Civil and Political Rights. Retrieved June 10, 2022 from https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights

[36] Fabio Urbina, Filippa Lentzos, Cedric Invernizzi, and Sean Ekins. 2022. Dual use of artificial-intelligence-powered drug discovery. Nature Machine Intelligence 4, 3 (2022), 189–191.

[37] Darrell West and John Allen. 2018. How artificial intelligence is transforming the world. Retrieved July 25, 2022 from https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/