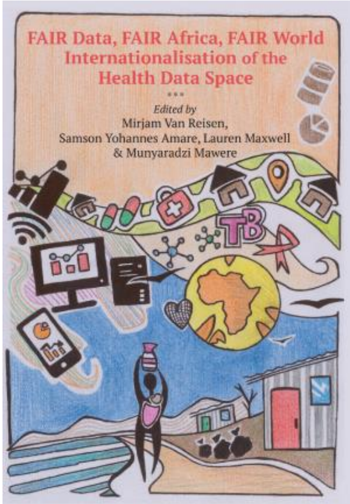
A Higher Education Curriculum for Cultural Competence, Representation and Social Responsibility in AI and FAIR Data Practices

Sakinat Folorunso, Francisca Oladipo, Mirjam van Reisen & Ibrahim Abdullahi

Chapter in:
Fair Data Fair Africa Fair World:
Internationalisation of the Health Data Space



Cite as: Folorunso, S. O., Oladipo, F. O., & Van Reisen, M. (2025). A higher education curriculum for cultural competence, representation, and social responsibility in AI and FAIR data practices. https://doi.org/10.5281/zenodo.15383056. In Van Reisen, M., Amare, S. Y., Maxwell, L. & Mawere, M. (Eds.), FAIR data, FAIR Africa, FAIR world: Internationalisation of the Health Data Space. (pp. 685–720). Bamenda: Langaa. URL: https://www.researchgate.net/publication/391750151_FAIR_Data_FAIR_Africa_FAIR_World_The_Internationalisation_of_the_Health_Data_Space

The About the Authors note can be found here: https://raee.eu/wp-content/uploads/2025/05/About-the-Authors-and-Editors.pdf
The list of figures and tables can be found here: https://raee.eu/wp-content/uploads/2025/05/List-of-Figures-and-Tables.pdf

Contents

Chapter 21

A Higher Education Curriculum for Cultural Competence, Representation and Social Responsibility in AI and FAIR Data Practices

Sakinat Folorunso, Francisca Oladipo, Mirjam van Reisen & Ibrahim Abdullahi

Abstract

The rise of data-driven decision-making points to the urgent need for ethical, transparent, and inclusive data practices. While implementing Findable, Accessible, Interoperable, and Reusable (FAIR) principles remains inconsistent, challenges in data science concerning cultural competence, representation and social responsibility remain. We propose a curriculum designed to advance FAIR practices in academia. It blends theoretical foundations with applied skills, enabling participants to assess datasets for FAIR compliance, refine metadata, and publish FAIR-compliant datasets on open science platforms. Leveraging cutting-edge tools, learners learn to detect and mitigate biases in datasets and artificial intelligence systems, ensuring ethical and representational outcomes. Beyond technical training, the programme highlights the importance of advocacy and leadership, empowering learners to develop data-driven campaigns and mentorship initiatives that address gender and societal disparities in STEM fields. The curriculum is structured into six interconnected modules: (1) Introduction to FAIR principles, (2) Data Governance, Ethics and Inclusion, (3) Data Management and Governance, (4) Inclusive Data Science and AI, (5) Open Science and FAIR Data Tools, and (6) Advocacy and Leadership in Data Practices. Each module combines lectures, case studies, and hands-on laboratories, culminating in a capstone project with a real-world dataset. Targeting higher education institutions, this curriculum bridges skill gaps in FAIR and ethics practices, aligning with global standards.

Keywords: FAIR principles, cultural competence, representation, social responsibility, Curriculum Development, Higher Education, AI

Introduction

The rapid expansion of data science has transformed decision-making across sectors such as healthcare, education, governance, and business. With its profound impact, data science is increasingly relied upon to solve complex problems, enhance efficiency, and improve outcomes. However, the growing reliance on data-driven technologies also brings to light a range of challenges, particularly in ensuring that these systems are both technically robust and socially equitable. One of the foundational frameworks for addressing these challenges is the adoption of the FAIR principles—Findable, Accessible, Interoperable, and Reusable—which aim to enhance data transparency, reproducibility, and accessibility (Wilkinson *et al.*, 2016). Despite their promise, the implementation of these principles remains inconsistent across domains (Van Reisen *et al.*, 2021; Van Reisen *et al.,* 2022). Gaps in technical knowledge, resource limitations, and a lack of integration with frameworks that address systemic inequities contribute to the challenges in achieving widespread adoption (Mons *et al.,* 2017; Van Reisen *et al.,* 2022).

Incompleteness in data to trace novel coronavirus (COVID-19) in Africa, showed the urgency of this problem (Van Reisen et al., 2022). Women who are pregnant or breastfeeding are often excluded from inclusion of clinical trials (Rubin, 2018). Increasing attention is given to what is invisible in data, including for education as shown in a learning corner of the EU (European Union, n.d.). These examples underscore the urgent need for a more inclusive and ethical approach to data science that addresses both technical precision and social responsibility (Green et al., 2023).

Alongside technical challenges, issues of cultural competence, representation and social responsibility within data science remain equally pressing. Historically, biases embedded in datasets and artificial intelligence (AI) systems have perpetuated systemic inequalities, disproportionately impacting marginalised communities. For instance, predictive policing algorithms, such as those used in the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) system, have been shown to unfairly target certain racial groups, exacerbating existing social inequities (Kirchner *et al.*, 2016).

COMPAS is a risk assessment algorithm used in the U.S. criminal justice system to predict the likelihood of a defendant reoffending. It analyses various factors, such as age, criminal history, and socioeconomic background, to generate risk scores that influence decisions on bail, sentencing, and parole. However, COMPAS has been widely criticised for potential racial bias, as studies have shown that it disproportionately assigns higher risk scores to Black defendants compared to white defendants with similar backgrounds (Kirchner *et al.*, 2016). The system is an example of predictive policing and criminal risk assessment tools that raise ethical and fairness concerns regarding algorithmic decision-making in the justice system. Biases in hiring algorithms and facial recognition software also highlight the risks of uncritical data practices that reinforce societal discrimination (Buolamwini & Gebru, 2018).

To tackle these intertwined challenges, it is imperative to develop educational frameworks that integrate the technical rigor of FAIR principles with the ethical imperatives of practices concerning cultural competence, representation and social responsibility. This curriculum is designed to prepare data scientists, researchers, and Science, Technology, Engineering, and Mathematics (STEM) professionals to navigate the complexities of modern data systems. By embedding FAIR principles within a framework of cultural competence, representation and social responsibility, this curriculum provides learners with not only the technical skills required to apply FAIR standards but also the tools to identify, assess, and mitigate biases. In doing so, it ensures that data science becomes a force for innovation and social justice rather than a mechanism that reinforces existing inequalities.

The curriculum is structured around six interconnected modules that balance theoretical foundations with applied skills, covering key topics, including FAIR data science, ethical governance, inclusive Artificial Intelligence (AI) design, open science practices, data management, and advocacy for equity in STEM. Learners will learn how to assess datasets for FAIR compliance, enhance metadata using international standards such as Dublin Core, and publish FAIR-compliant datasets on open science platforms like Zenodo. Furthermore, hands-on labs and real-world case studies will enable learners to evaluate biases in datasets, implement fairness interventions using tools like AI Fairness 360, and design advocacy campaigns that address inequities in STEM fields.

Aligned with the established EDISON Data Science Framework (EDSF), this curriculum ensures that its learning outcomes meet international standards and industry demands. Through this integrated approach, the curriculum not only bridges the skill gap in FAIR and inclusive data practices but also fosters ethical leadership. By empowering graduates to lead transformative change, this programme contributes to creating a data-driven world that prioritises transparency, accountability, and social justice.

This curriculum underscores the critical role of education in shaping a new generation of data scientists who value both technical expertise and ethical responsibility. By blending FAIR principles with practices of cultural competence, representation and social responsibility, the curriculum offers a holistic model for ensuring that data science not only drives technological innovation but also serves the greater good of society.

Integrating FAIR (Findable, Accessible, Interoperable, and Reusable) data principles into higher education curricula is essential for preparing learners to navigate the data-intensive landscape of modern research and industry. By embedding these principles, universities enhance research integrity and accessibility, as well-structured data promotes transparency and reproducibility in scientific findings.

Aligning educational programs with FAIR principles bridges the gap between academia and industry. Many sectors, including healthcare and finance, require professionals to be adept in data ethics and compliance. Implementing FAIR practices ensures that graduates possess the competencies necessary to meet these industry demands.

FAIR data practices support open education by making datasets more discoverable and reusable, fostering inclusivity and democratising access to knowledge. This approach aligns with global efforts in open science, contributing to transparent and equitable research ecosystems.

Incorporating FAIR principles into higher education curricula also promotes interdisciplinary learning. Learners from diverse fields can collaborate using standardised, interoperable datasets, enhancing data literacy and innovation.

A computer science and data science programme focusing on FAIR data skills was created in 2021 and tested by the Digital Innovation Skills Hub (DISH) in 2022 (Oladipo et al., 2022). This curriculum was the first curriculum available targeting learners globally on FAIR data science. The curriculum was targeting pre-university learners, particularly from a vulnerable background. The courses are particularly aimed at youth and women who have been particularly impacted by unemployment and who struggle to access (further) education. This includes second-school graduates, but also dropouts. DISH courses can be adapted for learners with internet access or without internet access, in order to facilitate learners who do not have constant connectivity.

The programme was tested in Sudan, South Sudan, Ethiopia, Somalia, Uganda and Kenya. The programme is set up as an e-learning programme. The DISH programme provides modules of 3 months courses, and provided four courses in computer science, data science and FAIR data science (Africa Health Data Space, n.d.; "Digital Innovation", 2022; DISH Kenya, 2023; Disk Mekelle, 2022; DISH Mekelle, 2023). The curriculum is accompanied by courses on mental health and wellbeing, health, legal and human rights frameworks, as well as business and administration skills. It has a practical orientation and is focused on skills that are attractive for the labour market.

Some of the graduates continued to go to university and moved through graduate and undergraduate informatics programmes, strengthening their capacities to integrate their notions of FAIR data stewardship in their computer science training.

The DISH programme is embedded in the Africa University Network on FAIR Open Science (AUN-FOS). The network has been set up to support FAIR and open science in Africa. Other networks, such as the UNA Europa network, have also followed suit, by offering specialised training programmes in FAIR data science for PhD learners, targeting the postgraduate level of researchers.

Bring Your Own Data (BYOD) workshops play a crucial role in fostering expertise in FAIR principles by providing hands-on, evolving approaches tailored to data, software, and data management. The article highlights how these workshops have developed over time to address different aspects of FAIRification, catering to various stakeholder needs. It underscores the importance of practical engagement in improving FAIR data practices and the necessity of adapting workshop methodologies to keep pace with advancements in data management and technology (Bernabé *et al.* 2024; "Una Europa", n.d.; "Unlocking data potential", 2025). Value-driven Ownership of Data and Accessibility Network (VODAN) Africa has adapted the DISH programme into a manual that has been used as summer schools and training of data stewards (Africa Health Data Space, n.d.). Field labs in the Computer Science curriculum at the University of Leiden, have functioned as an initiation in FAIR data science and in some instances leading to further research (Kievit et al., 2024; Utami, 2025; Lin, 2025; Jati & Lasroha, 2025; Smits et al., 2025).

A structured programme targeting undergraduate learners in universities is currently lacking. This cohort is particularly significant, as they represent the future workforce in data science and AI-related fields. Given the rapid advancement and integration of AI into various sectors, it is essential to equip these learners with the necessary skills and knowledge to navigate emerging challenges effectively. Preparing undergraduates for the evolving demands of AI-driven industries will enhance their employability and ensure a well-trained workforce capable of addressing the societal and ethical implications of AI deployment.

In addition to the gap identified above, research on the integration of FAIR-principles in innovation, point consistently to the importance of capacity building (Aktau, 2025; Zhang, 2022).

Theoretical framework

The theoretical framework serves as the academic and philosophical foundation for the curriculum, ensuring that it is rooted in well-established theories and best practices. By drawing from diverse theoretical perspectives, the framework provides a structured approach to understanding FAIR data science, ethics, and inclusive education. It integrates principles from data ethics, critical race studies, and education theory, shaping how learners engage with data governance, bias mitigation, and responsible AI practices.

A key pillar of this framework is the FAIR Data Principles (Wilkinson et al., 2016), which establish technical guidelines for ensuring that data is Findable, Accessible, Interoperable, and Reusable. These principles are essential for fostering transparency and reproducibility in scientific research. Within the curriculum, learners engage with these principles by evaluating datasets for FAIR compliance, using criteria such as persistent identifiers, metadata richness, and open licensing. By applying these concepts in practical scenarios, learners develop proficiency in data stewardship, interoperability, and ethical data sharing.

The 'Terminology of European Education and Training Policy' is a publication by the European Centre for the Development of Vocational Training (CEDEFOP). It defines a curriculum as a collection of "activities related to the design, organisation and planning of an education or training action, including definition of learning outcomes, content, methods (including assessment) and material, as well as arrangements for training teachers and trainers" (CEDEFOP, 2014). It serves as a comprehensive glossary of terms used in European education and training policies. The publication has undergone multiple editions, with the latest being the second edition published in 2014. The curriculum is a structured plan that defines what, why, how, and how well students learn, serving as a means to quality education (UNESCO IBE, 2011). The United Nations Educational, Scientific and Cultural Organization (UNESCO) emphasises that definitions vary, from a planned course of study to a structured sequence of learning outcomes (Australian Curriculum, Assessment and Reporting Authority, 2013; Donnelley, 2009). It incorporates educational content, sequencing, teaching methods, and evaluation (Braslavsky, 2003) which have tremendous influence on the thinking of the transformative capacity of curricula for a more inclusive society (Cox, 2005). Curriculum design is a dynamic, often debated process, with high political stakes. Probably ahead of her time, Cox summarises the relevance of Braslavsky on curricula development as follows: "her faith in education to re-create politics and improve collective life" (Cox, 2005).

At the intersection of data and social justice, Critical Race Theory (CRT) (Delgado & Stefancic, 2023; Gillborn *et al.,* 2023) provides a framework for analysing how systemic inequities manifest in datasets and AI systems. This perspective is crucial for addressing biases embedded in data-driven decision-making, particularly in areas such as criminal justice, healthcare, and hiring algorithms. Through case studies of real-world datasets, such as the COMPAS recidivism prediction model or racial disparities in healthcare data, learners critically assess how biases perpetuate discrimination. By applying bias detection tools and algorithmic auditing techniques, they learn how to design AI systems that actively mitigate rather than reinforce inequities.

Alongside the technical and critical dimensions of data science, the curriculum is also informed by the Ethics of Care (Gilligan, 1993), which advocates for human-centred decision-making in technology and governance. In contrast to purely utilitarian or deontological ethical models, this perspective emphasises the relational and contextual aspects of ethical responsibility, particularly in AI and data collection frameworks. Learners engage with this principle by designing AI models that prioritise societal wellbeing, such as frameworks for equitable healthcare resource allocation or community-driven data initiatives. This approach ensures that ethical considerations extend beyond abstract policy discussions and are integrated into real-world data science applications.

In designing an engaging and experiential learning environment, the curriculum also incorporates Constructivist Learning Theory (Jones, 1995). This theory emphasises active knowledge construction, where learners learn through hands-on projects, collaborative problem-solving, and real-world applications. Rather than relying solely on passive learning, learners are immersed in labs, case studies, and capstone projects that require them to apply theoretical knowledge to practical challenges. For example, in Module 6, learners design advocacy campaigns addressing gender disparities in STEM. This project applies Critical Race Theory to analyse inequities, the Ethics of Care to prioritise human-centred solutions, and Constructivist Learning Theory to encourage experiential engagement.

By the end of the last module, the capstone project, students will be able to apply FAIR principles in practice by assessing and enhancing the FAIR compliance of a dataset, ensuring it is findable, accessible, interoperable, and reusable in accordance with best practices. They will also develop the ability to identify and mitigate bias by conducting a systematic audit of datasets to detect gender, racial, or other biases and implementing appropriate fairness interventions to improve inclusivity in data-driven decision-making. Furthermore, students will demonstrate ethical leadership by critically engaging with ethical dilemmas in data management and advocating for responsible and transparent data practices that promote equity and inclusion. They will improve data documentation and metadata by evaluating and enhancing metadata quality, ensuring clarity, completeness, and alignment with FAIR standards to support better data reuse, increase representativity and enhance reproducibility. Lastly, students will effectively communicate their findings through a comprehensive capstone report and presentation, articulating the significance of FAIR and inclusive data practices, supported by data visualisations, a bias audit, and a Digital Object Identifier (DOI) linked dataset. These learning outcomes ensure that students integrate FAIR science principles with ethical leadership to address real-world data challenges while promoting inclusivity in research and policy.

By weaving together these theoretical perspectives, the curriculum provides a holistic and interdisciplinary approach to FAIR data science, management, and inclusion. It ensures that learners not only develop technical expertise in data science and AI governance but also cultivate a critical awareness of ethical, social, and systemic implications. This integration empowers future data professionals to contribute to a more equitable, responsible, and impactful data-driven society.

Methodology: Framework integration

Figure 1 shows how the competency, conceptual, and theoretical frameworks interact to create a comprehensive, interdisciplinary approach to FAIR data science education. Each framework plays a distinct role while also overlapping to form a holistic learning experience that integrates technical expertise, ethical considerations, and societal impact.

At its foundation, the competency framework focuses on technical and applied skills, ensuring that learners develop proficiency in data acquisition, FAIR principles, AI governance, and analytical methods. This framework equips learners with practical expertise in data science, preparing them to navigate industry demands with confidence. However, technical skills alone are insufficient to address the broader ethical and societal implications of data-driven decision-making.

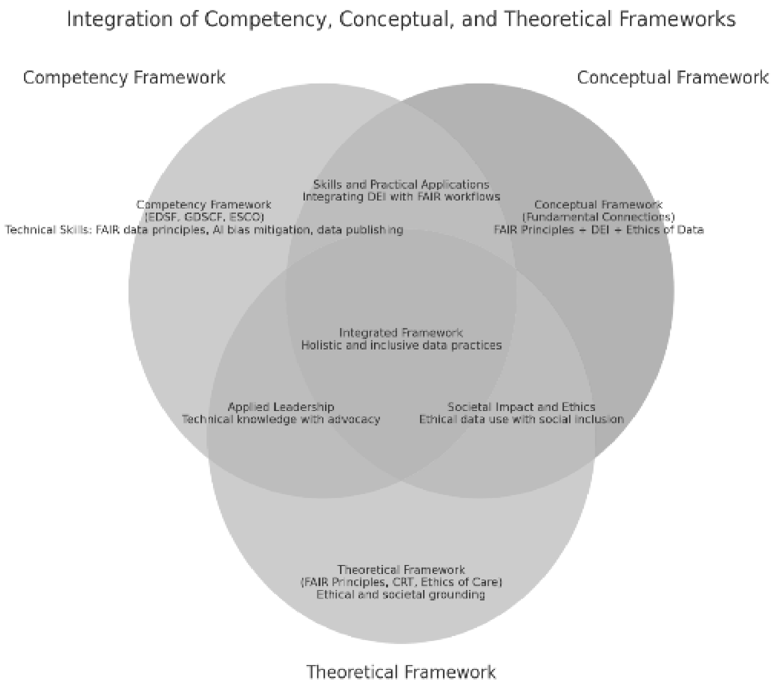


Figure 1. Framework integration

Bridging this gap, the conceptual framework connects technical practices with social dimensions, embedding cultural competence, representation and social responsibility, ethical AI, and responsible data governance into the learning process. It encourages learners to think critically about the impact of their work, recognising how biases in datasets, accessibility issues, and governance policies shape real-world outcomes.

The theoretical framework, meanwhile, provides ethical and academic grounding, drawing from disciplines such as critical data studies, FAIR and open science, and ethical AI development. This foundation ensures that learners are not only technically skilled but also equipped to analyse data practices through an ethical lens, understanding the historical, social, and legal contexts that influence data science.

The overlapping sections of the Venn diagram highlight key areas where these frameworks converge, reinforcing the interconnected nature of competency-based learning, conceptual thinking, and ethical inquiry. The intersection of competency and conceptual frameworks emphasises skills and practical applications, where learners apply FAIR principles to real-world cultural competence, representation and social responsibility challenges. For instance, they might conduct bias audits on AI models, improve metadata accessibility, or develop inclusive data policies, ensuring that technical solutions are aligned with societal needs.

Where conceptual and theoretical frameworks overlap, the focus shifts to societal impact and ethics, underscoring the broader consequences of data practices. Learners explore case studies on algorithmic bias, data privacy, and misinformation, engaging in ethical debates and policy analysis that highlight the responsibilities of data scientists in shaping equitable digital environments. The intersection of competency and theoretical frameworks fosters applied leadership, where learners leverage their technical expertise to advocate for responsible AI governance, data ethics, and digital equity. This dimension prepares them to lead interdisciplinary teams, influence policy decisions, and drive ethical innovation in data science.

At the centre of the diagram, the Integrated Framework emerges, representing the holistic nature of the curriculum. This central space embodies technical excellence, ethical responsibility, and societal awareness, reinforcing the idea that data science is not just about algorithms and efficiency but also about fairness, accountability, and impact. This integrated approach ensures learners graduate with the skills to address technical challenges, the ethical grounding to make responsible decisions, and the leadership capabilities to promote equity and inclusion in data science.

Curriculum structure

The course comprises 5 core modules and a capstone project, blending theoretical foundations, hands-on labs, and practical applications.

Module 1: Introduction to FAIR Principles
Week 1–2

Topics

What are FAIR Principles?

- Definitions: Findable, Accessible, Interoperable, Reusable
- Why FAIR matters in modern data science
- FAIR in higher education and research workflows

Key Metrics for FAIR Compliance:

- Metadata completeness
- Licensing and accessibility
- Interoperability standards

Challenges to FAIR Implementation:

- Lack of awareness, resource limitations, and tools

Module objectives

To introduce FAIR principles and their applications in higher education and research.

Understand the FAIR principles framework

- Explain the meaning and importance of Findability, Accessibility, Interoperability, and Reusability
- Recognise how FAIR principles support open science and research reproducibility

Recognise the value of FAIR data practices

- Identify benefits to individual researchers, institutions, and the broader scientific community

- Explain how FAIR practices enhance research visibility, citation, and collaboration

Apply basic FAIR concepts to research workflows

- Identify key intervention points in the research lifecycle for implementing FAIR practices
- Recognise common barriers to FAIR implementation and potential solutions

Module 2: Data Governance, Ethics, and Inclusion
Week 3–5

<u>Topics</u>

Ethics in Data Practices:

- Privacy, security, equitable access, and consent in data collection.
- Addressing ethical dilemmas in AI systems.
- Ethical considerations in data sharing and reuse
- Addressing bias and inequities in datasets
- Data privacy, intellectual property, and governance frameworks

Data Governance Frameworks:

- GDPR, UNESCO's Data Ethics Framework, and local policies.
- Ethical considerations in data sharing and reuse.

Bias in Datasets:

- Types of bias: Sampling, algorithmic, and societal.
- Gender biases in datasets: Identifying and addressing them
- Real-world consequences of biased systems (e.g., COMPAS dataset).

Module objectives

Explore ethical practices in data governance, focusing on inclusion, equity, and transparency.

Introduce basic data governance concepts

- Define what data governance means in practical terms
- Connect governance to FAIR principles

Explore fundamental ethical considerations

- Identify common ethical issues in research data management
- Understand the importance of informed consent

Recognise inclusion challenges in data practices

- Understand how bias can enter data collection and sharing
- Learn simple steps toward more inclusive data practices

Module 3: Data Management and Governance

Week: 6–8

Topics

Data lifecycle management

- Data collection
- Data storage
- Data processing

Module objectives

Implementing Data Management Plans (DMPs)

- Introduction to Data Management Plans (DMPs) and FAIR Data Policies.
- FAIR & Ethical Data Storage Solutions (Cloud Repositories, Institutional Repositories).
- Open Science & Open Data Policy Implementation in Institutions.

- Evaluating Data Governance Success Metrics (Compliance Scorecards).

FAIR data stewardship, Interoperability & long-term preservation.

- Understanding Data Lifecycle Management (Collection, Processing, Storage, Distribution, Archiving, Reuse).
- Data Documentation Initiative (DDI) & Metadata Standards for research data.
- Interoperability & Linked Data: How datasets interact across platforms (Query Language and Resource Description Framework (SPARQL), Resource Description Framework (RDF).
- FAIRification of Research Data: Ensuring long-term usability and compliance.

Data Governance, Security, and Compliance frameworks:

- ISO 27001 (Information Security), Data Documentation Initiative (DDI).
- Data Governance Models (Centralised, Decentralised, Hybrid).
- Data Access & Privacy Regulations (International Organization for Standardization (ISO) 27001, GDPR, Health Insurance Portability and Accountability Act (HIPAA) Compliance).
- Data Security & Encryption Techniques (Advanced Encryption Standard (AES), Transport Layer Security (TLS), Secure Hashing).
- Data Ownership & Licensing (Creative Commons, Open Data Licensing).

<u>Module objectives</u>

Teach data stewardship, interoperability, security, and compliance

- Ensure Data Integrity and Security – Maintain accurate, consistent, and reliable data while implementing privacy and security measures to protect against breaches and unauthorised access.

- Enable FAIR Data Principles – Ensure data is Findable, Accessible, Interoperable, and Reusable, enhancing transparency, reproducibility, and responsible data sharing.

- Ensure Compliance with Regulations – Align data management practices with legal and ethical standards such as GDPR, ISO 27001, HIPAA, and Open Data Directives to uphold data rights and governance.

- Optimise Data Governance and Interoperability – Establish clear frameworks for data ownership, stewardship, and accountability, while ensuring seamless data exchange through standardised metadata (e.g., Dublin Core, RDF, DataCite).

- Foster Open Science and Ethical AI Development – Promote open data sharing through repositories (Zenodo, Figshare) while ensuring bias mitigation and fairness in AI-driven decision-making

Module 4: Inclusive Data Science and AI

Week: 9–10

<u>Topics</u>

Understanding Bias in AI:

- How bias propagates through the AI lifecycle

- Examples: Hiring algorithms, and facial recognition systems

Designing Inclusive AI:

- Principles for inclusive datasets and models

- Tools for bias mitigation (e.g., IBM AI Fairness 360, Google What-If Tool)

Addressing Gaps in Data Representation:

- The importance of diverse datasets

Module objectives

Learn how to design and implement an inclusive dataset and build AI models and data practices.

- Promote Fairness and Equity in AI Development – Ensure AI models and data practices are designed to reduce biases and promote equitable outcomes, particularly for marginalised communities.
- Identify and Mitigate Bias in Data and Algorithms – Equip students with tools and techniques to audit datasets, assess algorithmic bias, and implement fairness metrics to create more responsible AI systems.
- Develop Ethical and Transparent AI Practices – Integrate ethical frameworks such as AI Fairness 360, UNESCO AI Ethics, and Organisation for Economic Co-operation and Development. (OECD) AI Principles, fostering transparency, accountability, and trust in AI decision-making.
- Enhance Diversity and Representation in Data Collection – Advocate for inclusive dataset design that accurately represents gender, racial, disability, and socioeconomic diversity to ensure fair AI applications.
- Apply Inclusive AI Solutions in Real-World Scenarios – Engage students in hands-on projects and case studies that address bias mitigation, inclusive policy development, and AI-driven social impact initiatives.

Module 5: Open Science and FAIR Data Tools

Week: 11–12

Topics

Open Science Framework:

- Principles of open science and its relation to FAIR.
- Benefits of open-access datasets.

FAIR Data Tools:

- Tools and platforms for sharing FAIR data (e.g., Zenodo, Figshare).

- Metadata standards: Dublin Core, DataCite.

Licensing and Accessibility:

- Selecting appropriate open licenses (e.g., CC-BY-4.0).
- Challenges faced by women and underrepresented groups in open data.

<u>Module objectives</u>

Provide hands-on experience with tools for open science and FAIR-compliant data publishing.

- Apply FAIR Data Principles in Research and Data Management – Ensure datasets are Findable, Accessible, Interoperable, and Reusable (FAIR) to enhance transparency and reproducibility.
- Utilise FAIR Data and Open Science Tools – Train students to use FAIRsharing.org, Zenodo, Figshare, and metadata standards like Dublin Core and DataCite for effective data management and sharing.
- Ensure Ethical and Legal Data Sharing – Educate students on open data licensing frameworks such as Creative Commons (CC-BY) and Open Government Data (OGD) for responsible data use.
- Promote Open and Reproducible Research Practices – Teach students to document research workflows using Jupyter Notebooks, R Markdown, and Open Science Framework (OSF) to ensure transparency.
- Encourage Collaboration and Open Science Advocacy – Empower students to engage with global Open Science initiatives, interdisciplinary collaborations, and policy discussions to drive Open Science adoption.

Module 6: Advocacy and Leadership for Diversity and Equity in Data Practices

Week: 13–14

<u>Topics</u>

Policy and Advocacy for Equity for Women and Girls in STEM:

- The role of leadership in creating inclusive environments.
- Global initiatives for gender equity in STEM (e.g., UNESCO programs).
- International policies promoting inclusion (e.g., Sustainable Development Goals (SDGs), UNESCO).
- Strategies for mentorship and capacity building.

Data-Driven Advocacy:

- Using data visualisation to highlight disparities.
- Designing impactful campaigns.

Mentorship and Community Building:

- Supporting underrepresented groups in data science.
- Foster leadership and advocacy skills for promoting diversity and inclusion in STEM and data science.

<u>Module objectives</u>

Advocacy and Leadership for Diversity and Equity in Data Practices

- Promote Inclusive Data Governance and Ethical AI – Advocate for fair, unbiased, and transparent data practices that ensure equitable representation in AI and decision-making systems.
- Develop Leadership Skills for Diversity in Data Science – Empower students to take leadership roles in mentorship, policymaking, and advocacy initiatives that foster diversity in STEM and data-driven fields.
- Address Bias and Disparities in Data Collection and Use – Equip students with strategies to identify, mitigate, and challenge biases in datasets, algorithms, and research methodologies.
- Engage in Policy and Community-Based Advocacy – Train students to influence policy discussions, industry standards, and institutional reforms that advance diversity and inclusion in data practices.

- Foster Cross-Disciplinary and Inclusive Collaboration – Encourage partnerships across academia, industry, and marginalised communities to develop equitable data solutions and inclusive innovation.

Module 7: Capstone Project – FAIR Data and Inclusion in Action

Weeks. 15–17

Tasks

- Select a dataset in a chosen domain (e.g., healthcare, education, environment).
- Audit the dataset for gender or racial biases.
- Evaluate its FAIR compliance and improve metadata.
- Conduct a bias audit and implement fairness interventions.
- Publish the dataset on a FAIR-compliant repository.
- Present findings in a written report and oral presentation
- Check FAIR compliance checklist.
- Prepare bias audit report with visualisations.
- Provide a DOI of the published dataset.
- Capstone presentation.

Expected learning outcomes, deliverables, laboratory guides, and assessments

Module 1 will equip students with the knowledge and skills to apply FAIR principles in data management, enhancing research transparency and accessibility. Students will begin by defining and explaining the components of FAIR, distinguishing them from open data concepts. They will explore common metadata standards relevant to their fields, including discipline-specific schemas and persistent identifiers (DOIs, Open Researcher and Contributor ID [ORCIDs]), ensuring proper data attribution and accessibility. Through practical activities, students will conduct a FAIR compliance audit using FAIRsharing.org, analyse case studies such as the Human Genome Project, and engage in discussions comparing FAIR and non-FAIR datasets to identify gaps. Hands-on labs will allow them to assess, improve, and document FAIR compliance for sample datasets. Assessment includes submitting a FAIR compliance checklist and an evaluation report with proposed improvements for a non-compliant dataset. By the end of the module, students will develop a basic data management plan, ensuring sustainable and ethical data-sharing practices.

This module 2 equips students with the skills to evaluate datasets for bias and fairness, critically examining the ethical implications of data-driven decision-making. By exploring real-world cases of algorithmic bias, students will understand how flawed data governance can reinforce discrimination and learn strategies to mitigate these risks. Through hands-on activities, students will conduct a bias audit of the COMPAS dataset, a well-known case of racial bias in predictive policing. Using Python (Pandas, Scikit-learn) and AI Fairness 360, they will generate fairness metrics, identify disparities, and propose corrective measures. A case study analysis will deepen their understanding of data governance failures, while a workshop on gender and racial biases in AI (Noble, 2018) will encourage critical discussions. For assessment, students will submit a bias audit report detailing their findings, fairness metrics, and recommendations for improving dataset inclusivity, ensuring they can apply ethical data science practices in real-world scenarios.

Module 3 will equip learners with the knowledge and skills to develop and implement data management plans (DMPs) aligned with FAIR principles, ensuring that datasets are findable, accessible, interoperable, and reusable. Learners will learn to evaluate data interoperability using metadata standards such as Dublin Core, DataCite, and RDF, enhancing dataset structure and accessibility. Additionally, they will explore data governance policies, ensuring compliance with GDPR, ISO 27001, and HIPAA, while integrating privacy, security, and ethical considerations into data practices. Through hands-on exercises, learners will implement data protection mechanisms, including AES encryption and role-based access control (RBAC), reinforcing secure data storage and management. They will also publish FAIR-compliant datasets using open science repositories like Zenodo and Figshare, applying licensing frameworks such as CC-BY and OGD policies. Assessment includes developing a DMP, conducting metadata structuring exercises, drafting a data governance policy, and publishing a FAIR-compliant dataset, ensuring practical application of learned concepts.

Module 4 equips learners with the skills to train AI models with inclusivity in mind, ensuring that machine learning systems are designed to mitigate bias and promote fairness. Through a combination of hands-on practice, case studies, and critical analysis, learners will learn to identify, measure, and address biases in AI models using tools such as AI Fairness 360. In the lab activity, learners will build a machine learning model on a biased hiring dataset, assessing its fairness before and after applying bias mitigation techniques. They will also engage in a design exercise, proposing inclusive AI solutions to societal challenges, such as equitable healthcare access. Case studies will further explore the impact of AI on women and marginalised communities, reinforcing the importance of ethical AI design. For assessment, learners will submit a comparative report evaluating the performance of their model before and after bias mitigation, ensuring a practical, evidence-based approach to responsible AI development.

Module 5 will equip students with the skills to create, validate, and publish FAIR-compliant datasets, ensuring alignment with Findability, Accessibility, Interoperability, and Reusability (FAIR) principles. By understanding the technical and ethical dimensions of data sharing, students will learn how to structure metadata, apply licensing frameworks, and assess dataset compliance using recognised standards. Through a lab activity, students will publish a dataset on an open repository such as Zenodo, Fishare or Mendeley, ensuring proper metadata documentation and licensing. A workshop will guide students in evaluating datasets from different repositories and identifying compliance gaps and best practices for FAIR data management. This comparative analysis will help them recognise how different platforms handle accessibility, interoperability, and reuse. For assessment, students will submit the DOI of their published dataset along with a compliance validation report, demonstrating their ability to apply FAIR principles and evaluate dataset quality within an open science framework.

Module 6 empowers students to develop advocacy and mentorship strategies while using data-driven insights to promote diversity and inclusion in STEM. By combining analytical skills with social impact initiatives, students will learn how to identify challenges, propose actionable solutions, and drive meaningful change in underrepresented communities. Through a group activity, students will explore local gender equity challenges in STEM, designing solutions with clear, measurable action steps. They will then create a data-driven advocacy campaign, using visualised data to support their message. In the lab activity, students will leverage Tableau or Python (Matplotlib) to analyse and visualise STEM representation disparities, reinforcing the importance of evidence-based advocacy. For assessment, students will present their campaign, incorporating visualisations and key action points, demonstrating their ability to combine data science with social impact strategies to advance gender equity in STEM.

To ensure a comprehensive and impactful learning experience, this section integrates cutting-edge tools, inclusive pedagogical strategies, and meaningful partnerships that support FAIR data science, bias mitigation, and gender equity in STEM. Students will gain hands-on experience with data analysis tools such as Python (Pandas, NumPy, Matplotlib), R, Waikato Environment for Knowledge Analysis (WEKA), Kofler-Nielsen-Mittelbach-Eckert (KNIME), and Orange Data Mining (ORANGE) while leveraging FAIR compliance platforms like FAIRsharing.org and FAIR Metrics tools to evaluate dataset accessibility and interoperability. They will also work with the IBM AI Fairness 360 Toolkit and Google What-If Tool to identify and mitigate bias in AI models and create compelling visualisations using Tableau, Canva, or Python libraries.

The pedagogical approach centres on interactive learning through workshops, hackathons, and mentorship programs, connecting women students with industry mentors. Case studies and community outreach further reinforce practical learning. Evaluation metrics will measure programme success, tracking women's participation, dataset quality, advocacy efforts, and feedback. Key partnerships include data science organisations, women's advocacy groups, UNESCO, and universities, ensuring sustainability and policy impact.

Discussion on frameworks and practical examples

This section explains the competency, conceptual, and theoretical frameworks with additional practical examples to clarify how these frameworks guide the FAIR data science and inclusion curriculum.

Competency framework

The Competency Framework defines the key knowledge, skills, and abilities learners will acquire to master FAIR principles and practices concerning cultural competence, representation and social responsibility. This framework ensures practical application in real-world contexts, fostering a balance of technical expertise and ethical leadership.

The EDISON Data Science Framework (EDSF) (Demchenko *et al.,* 2016) provides a comprehensive structure to define and standardise the competencies required for data science professionals. At its core, EDSF integrates three key components: the Data Science Competence Framework (CF-DS), the Data Science Body of Knowledge (DS-BoK), and the Data Science Model Curriculum (MC-DS). Together, these elements bridge the gap between academic training and industry demands, ensuring that data science education remains relevant and aligned with professional expectations.

Conceptual foundation of the EDISON Data Science Framework

The EDSF establishes a structured approach to identifying and categorising the essential competencies for data science professionals. It provides the foundation for developing standardised curricula, competency-based certification programmes, and training pathways, ensuring that data scientists acquire the skills necessary to thrive in an evolving, data-driven world. By offering a systematic model, the framework supports universities and professional institutions in designing education programs that reflect the diverse applications of data science across industries.

Theoretical underpinnings: The Data Science Competence Framework (CF-DS)

A cornerstone of the EDSF is the Data Science Competence Framework (CF-DS), which defines the specific competencies and skill sets required for data science roles. The framework is structured into five main competency groups, each addressing a critical aspect of data science:

- Data Analytics (DSDA) – Focuses on the ability to analyse and interpret complex data, ensuring that professionals can derive meaningful insights from structured and unstructured datasets.

- Data Science Engineering (DSENG) – Covers the design and implementation of data systems, including computational infrastructure, software pipelines, and scalable architectures that support large-scale data analysis.

- Data Management (DSDM) – Encompasses data governance, curation, quality control, and compliance, ensuring that data is well-organised, standardised, and adheres to ethical and legal standards.

- Research Methods (DSRM) – Focuses on the application of scientific methodologies, experimental design, and hypothesis testing in data-driven research.

- Business Process Management (DSBPM) – Relates to optimising business processes through data insights, enabling organisations to leverage data for strategic decision-making and operational efficiency.

By aligning these competence groups with the European e-Competence Framework (e-CF3.0), the CF-DS ensures integration with existing professional standards, making it applicable across both academia and industry.

Competency frameworks: The DS-BoK and MC-DS

The Data Science Body of Knowledge (DS-BoK) expands on the CF-DS by detailing the knowledge areas associated with each competency group. It provides a structured reference for essential topics, methodologies, and best practices that professionals must master to excel in the field. Complementing this, the Data Science Model Curriculum (MC-DS) operationalises these competencies into structured learning modules, ensuring that educational programs equip graduates with industry-relevant skills.

Through the DS-BoK and MC-DS, the EDSF ensures that universities and training institutions have a well-defined roadmap for curriculum development, competency evaluation, and alignment with real-world applications of data science.

Integrating the EDISON Framework into a comprehensive data science curriculum

To effectively integrate these frameworks into a cohesive data science curriculum, it is essential to align the six core learning modules with the EDSF competency groups. This mapping ensures that graduates acquire both theoretical knowledge and applied skills needed to address contemporary data science challenges.

1. Data acquisition & integration

This module focuses on data ingestion, API integration, and ensuring interoperability across systems. It aligns with competencies in Data Science Engineering (DSENG) and Data Management (DSDM), equipping learners with the ability to collect, transform, and integrate data from diverse sources while ensuring data quality and accessibility.

2. Data storage & FAIR repositories

Learners in this module learn to implement FAIR principles (Findability, Accessibility, Interoperability, and Reusability) in data storage solutions. It aligns with Data Management (DSDM) and Data Science Engineering (DSENG), teaching learners how to structure and manage repositories, optimise metadata, and ensure compliance with open science initiatives.

3. Metadata curation & FAIRification

This module ensures that learners develop expertise in metadata structuring, data annotation, and enhancing data findability and accessibility. The competencies from Data Management (DSDM) and Data Analytics (DSDA) are incorporated to help learners create standardised and machine-readable metadata frameworks that facilitate seamless data exchange and reuse.

4. Data quality & preprocessing

Ensuring data quality is fundamental to any data-driven process. This module covers cleaning, validation, and preprocessing techniques while aligning with Data Analytics (DSDA) and Data Management (DSDM). Learners are trained to detect and correct inconsistencies, reduce noise, and prepare datasets for analysis using industry best practices.

5. Data security & privacy compliance

As data privacy regulations continue to evolve, this module emphasises data protection, security protocols, and regulatory compliance (such as GDPR and ISO 27001, the EU AI Act and Digital Services Act and Digital Markets Act and legal provisions from China Personal Information Protection Law (PIPL) and the African Union Convention on Cyber Security and Personal Data Protection). It integrates competencies from Data Science Engineering (DSENG), Data Management (DSDM), and Business Process Management (DSBPM), ensuring that learners understand encryption, access control, and ethical considerations in data governance.

6. Data governance, AI FAIRness & ethics

Addressing the ethical dimensions of data science, this module teaches learners how to establish governance frameworks, promote fairness in AI applications, and ensure transparency in decision-making systems. It aligns with Business Process Management (DSBPM), Research Methods (DSRM), and Data Analytics (DSDA), emphasising interdisciplinary collaboration between data science, policy, and ethics.

Ensuring comprehensive competency development in data science education

By aligning each module with the relevant competency groups from the EDSF, educational institutions can design holistic curricula that equip learners with technical expertise and ethical awareness. This structured approach ensures that graduates are proficient in data engineering, analytics, and governance and capable of addressing the broader social, legal, and ethical implications of data-driven decision-making.

As data science continues to shape critical aspects of society, from artificial intelligence to policy-making, universities must actively foster responsible and inclusive data practices. The EDISON framework, through its integration of competence-based learning, standardised knowledge areas, and interdisciplinary collaboration, provides a roadmap for creating future-ready data professionals who are equipped to navigate the complexities of a rapidly evolving field.

Conceptual framework

The conceptual framework provides a structured foundation for understanding how FAIR principles, cultural competence, representation and social responsibility practices, and ethical data governance interact to create a responsible and inclusive data ecosystem. By integrating technical rigor with social responsibility, this framework ensures that data science professionals are equipped with both the analytical skills and the ethical mindset needed to address real-world challenges. Through this lens, the curriculum fosters a holistic approach to data education, bridging technical proficiency with accountability and impact.

At the heart of this framework, the FAIR principles serve as the technical backbone, ensuring that data practices are transparent, accessible, reproducible, and interoperable (Wilkinson et al., 2016). By embedding these principles into coursework, learners engage with the foundational aspects of data stewardship, from metadata completeness and licensing to data accessibility and reusability. For instance, in Module 1, learners conduct FAIR compliance audits, analysing datasets for persistent identifiers, standardised metadata, and machine-readable formats. Through this process, they develop the ability to evaluate and improve data transparency and interoperability, ensuring that datasets meet the highest standards for usability.

Beyond technical considerations, the framework incorporates cultural competence, representation and social responsibility as a social layer, addressing systemic biases in data representation and AI-driven decision-making. Data science has historically been shaped by unequal power structures, often leading to underrepresentation of marginalised groups in datasets (Noble, 2018). To confront these challenges, the curriculum integrates critical examinations of data biases, equipping learners with the tools to identify, critique, and propose solutions for equity gaps. For example, in Module 4, learners analyse demographic representation in healthcare datasets, evaluating whether certain populations, such as ethnic minorities or lower-income groups, are underrepresented. By applying cultural competence, representation and social responsibility principles, they recommend alternative data collection strategies to ensure more inclusive and representative datasets.

Alongside cultural competence, representation and social responsibility, ethics and data governance serve as the guiding framework, ensuring that accountability, transparency, and societal impact remain at the forefront of all data practices. As AI-driven systems increasingly influence criminal justice, finance, and hiring decisions, ethical considerations must be deeply embedded in data science education. The curriculum emphasises global ethical standards and regulatory frameworks, such as the GDPR, ISO 27001, and UNESCO AI Ethics Guidelines. In Module 2, learners engage with real-world ethical dilemmas, such as the racial biases found in facial recognition software, critically analysing case studies where misclassification has disproportionately affected individuals of colour. This exercise not only sharpens their technical evaluation skills, but also reinforces the importance of responsible AI deployment.

To bridge the technical and social dimensions, education serves as the critical link, transforming theoretical understanding into practical, hands-on learning experiences. The curriculum is designed around experiential learning, where learners apply FAIR principles, conduct bias audits, and implement ethical governance strategies in real-world scenarios. This approach reaches its peak in the capstone project, where learners synthesise their knowledge by improving a dataset's FAIR compliance, auditing it for bias, and publishing their findings. This culminating exercise underscores the interconnected nature of FAIR data management, cultural competence, representation and social responsibility practices, and ethical responsibility, demonstrating how these components work together to create a more equitable and responsible data science ecosystem.

These elements are not isolated; rather, they function as an integrated system, where FAIR principles provide the technical foundation, cultural competence, representation and social responsibility principles ensure inclusivity, and ethical governance guarantees accountability. Their dynamic interaction enables learners to tackle complex data science challenges, preparing them for roles where they must balance technical expertise with social impact. This interdisciplinary model ensures that graduates are not only proficient in data analytics and engineering but also capable of making informed, ethical, and equity-driven decisions.

A prime example of this integrated approach is found in Module 5, where learners validate datasets for FAIR compliance while simultaneously assessing their broader societal impact. Through this process, they recognise that technical improvements, such as metadata enhancement and standardised licensing, must align with ethical considerations, including privacy, representation, and accessibility. By engaging in this holistic analysis, learners gain a deep understanding of how data science intersects with human rights, social justice, and policy-making, ensuring that they enter the workforce as responsible, critically engaged data professionals.

Through this conceptual framework, the curriculum fosters a new generation of data scientists who are not only skilled in data management, machine learning, and analytics but also deeply aware of the ethical and social dimensions of their work. By weaving together FAIR principles, cultural competence, representation and social responsibility, and ethical governance, this approach ensures that data science serves as a force for transparency, accountability, and inclusivity—paving the way for a more equitable digital future.

Practical examples across frameworks for capstone projects

Example 1: FAIR compliance audit

Activity: learners evaluate a publicly available dataset for compliance with FAIR principles.

- Competency framework: learners learn to identify missing metadata and enhance compliance.

- Conceptual framework: Highlights the relationship between technical improvements (FAIR principles) and their impact on accessibility and inclusivity.

- Theoretical Framework: Grounded in FAIR principles, with an experiential learning approach.

Example 2: Bias audit and mitigation

Activity: Analyse bias in the COMPAS dataset and propose mitigation strategies.

- Competency framework: learners gain technical skills in bias detection and mitigation.

- Conceptual framework: Links technical analysis (bias audit) with ethical implications and societal impact.

- Theoretical framework: Draws on CRT to understand systemic biases and ethics of care to prioritise equitable solutions.

Example 3: Advocacy campaign

Activity: Develop a campaign highlighting disparities in STEM fields using real-world data.

- Competency framework: learners develop leadership and advocacy skills.

- Conceptual framework: Connects the technical (data visualisation) and social (STEM inequities) aspects of the curriculum.

- Theoretical framework: Grounded in ethics of care, prioritising societal wellbeing, and constructivist pedagogy for collaborative learning.

Conclusion

The development of the curriculum in this study represents a transformative approach to addressing the intertwined challenges of ensuring data systems are technically robust and socially equitable. As data science increasingly shapes decision-making in critical sectors such as healthcare, education, and governance, the need for transparency, reproducibility, and accessibility in data practices has become paramount. At the same time, addressing systemic biases and fostering cultural competence, representation and social responsibility within data science is essential to ensure that data-driven technologies do not perpetuate or exacerbate societal inequities.

By embedding FAIR principles into the curriculum, this framework equips learners with the technical tools and methodologies needed to enhance data quality and usability. The incorporation of global competency frameworks such as the EDISON Data Science Framework (EDSF), and the ESCO Framework ensures that learners acquire globally recognised competencies in FAIR data management, bias mitigation, and ethical governance. This alignment with international standards not only enhances the technical rigour of the curriculum but also ensures its relevance across diverse contexts and disciplines.

Beyond technical expertise, the curriculum's integration of cultural competence, representation and social responsibility practices addresses critical societal challenges. By enabling learners to identify and mitigate biases in datasets and artificial intelligence systems, the programme ensures that future data scientists are equipped to develop technologies that are inclusive and just. Real-world case studies, hands-on labs, and advocacy projects provide opportunities for learners to connect theoretical principles with practical applications, fostering a deep understanding of how data science can drive both innovation and social justice.

The curriculum's five interconnected modules, spanning FAIR principles, ethical governance, inclusive AI design, open science practices, and advocacy for equity, offer a comprehensive and interdisciplinary learning experience. This modular structure ensures a balance between theoretical foundations and applied skills, empowering learners to navigate complex technical challenges while promoting transparency, accountability, and equity. The capstone projects, case studies, and practical exercises embedded within the curriculum further enhance its relevance, preparing learners to address real-world data challenges in diverse domains.

In addition to its immediate impact on learner's learning, this curriculum serves as a model for embedding ethical and inclusive practices into data science education. By fostering leadership and advocacy skills, it equips graduates to not only excel in technical domains but also to lead transformative change in their organisations and communities. The curriculum underscores the role of education in shaping a new generation of data scientists who prioritise both innovation and social responsibility.

FAIR data science plays a crucial role in promoting fairness through FAIRification, ensuring that data is Findable, Accessible, Interoperable, and Reusable while upholding ethical principles of cultural competence, representativity, inclusion, and social responsibility. By embedding these values in data governance and AI development, FAIRification enables more equitable and transparent decision-making processes that address biases and support diverse communities. Leadership in this field requires a commitment to ethical stewardship, the ability to navigate complex socio-technical challenges, and the skills to foster interdisciplinary collaboration. Strengthening AI for global good demands leaders who not only advocate for FAIR principles but also actively implement inclusive frameworks that enhance trust, accountability, and fairness in data-driven innovation. Through responsible leadership, FAIR data science can serve as a catalyst for a more just and representative digital future.

Authors' Contributions

Sakinat Folorunso conceptualised the research and wrote the first draft. **Francisca Oladipo** helped elaborate the concept and commented on the various drafts. **Mirjam van Reisen** wrote sections of the chapter and edited the final draft. **Ibrahim Abdullahi** provided insights and suggestions for the curriculum.

Ethical Considerations

Tilburg University, Research Ethics and Data Management Committee of Tilburg School of Humanities and Digital Sciences REDC#2020/013, June 1, 2020-May 31, 2024, on Social Dynamics of Digital Innovation in remote non-western communities. Uganda National Council for Science and Technology, Reference IS18ES, July 23, 2019-July 23, 2023.

References

Africa Health Data Space. (n.d.). *Africa Health Data Space training portal.* https://aun.mu.edu.et/ahds/training/

Aktau, A. (2025). The Potential of Adoption of FAIR Guidelines in Digital Healthcare in Kazakhstan. In M. Van Reisen, M., Amare, S. Y., Maxwell, L. & Mawere, M. (Eds.), *FAIR data, FAIR Africa, FAIR world: Internationalisation of the Health Data Space.* Bamenda: Langaa

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2013). *General capabilities in the Australian Curriculum*. Sidney: ACARA.

Barrantes, S. A., Rodriguez, M., & Pérez, R. (2009). Information management and communication in emergencies and disasters: manual for disaster response teams. Pan American Health Organization, Washington, DC.

Behr, F. J., & Huber, M. (2012). Enabling Interoperability: Data models for humanitarian emergency situations. AGSE 2012–FOSS4G-SEA, 205.

Bernabé, C. H., Thielemans, L., Kaliyaperumal, R., Carta, C., Zhang, S., van Gelder, C. W. G., Benis, N., Bonino da Silva Santos, L. O., Cornet, R., Vieira, B. D. S., Lalout, N., Henriques, I., Ballesteros, A. C., Burger, K., Kersloot, M. G., Ehrhart, F., van Enckevort, E., Evelo, C. T., Gray, A. J. G., ... Jacobsen, A. (2024). Building expertise on FAIR through evolving Bring Your Own Data (BYOD) workshops: Describing the data, software, and management-focused approaches and their evolution. *Data Intelligence, 6*(2), 429–456. https://doi.org/10.1162/dint_a_00236

Braslavsky, C. 2003. *The curriculum*. UNESCO International Bureau of Education.

Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77-91). PMLR.

CEDEFOP. 2014. *Terminology of European education and training policy*. (2nd ed). Luxemburg: Publications Office of the European Union.

Cox, C. (2005). Cecilia Braslavsky and the curriculum. *Prospects, 35*(4), 415–427. https://doi.org/10.1007/s11125-006-7258-9

Delgado, R., & Stefancic, J. (2023). *Critical race theory: An introduction* (Vol. 87). NyU press.

Demchenko, Y., Belloum, A., Los, W., Wiktorski, T., Manieri, A., Brocks, H., Becker, J., Heutelbeck, D., Hemmje, M., & Brewer, S. (2016). EDISON data science framework: A foundation for building data science profession for research and industry. In *2016 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)* (pp. 620–626). IEEE. https://doi.org/10.1109/CloudCom.2016.0092

Digital Innovation and Skills Hub (DISH). (2022, October 15). *DISH portal.* https://aun.mu.edu.et/Dish-portal/

DISH Kenya. (2023, March 21). *Tangaza University College graduates students from DISH programme in Kenya*. Digital Innovation and Skills Hub. https://aun.mu.edu.et/Dish-portal/2023/02/26/tangaza-university-college-graduates-students-from-dish-programme-in-kenya/

DISH Mekelle. (2022, September 3). *Mekelle University to train 2800 youths under the DISH project in Tigray*. Digital Innovation and Skills Hub. https://aun.mu.edu.et/Dish-portal/2022/09/03/mekelle-university-to-train-2800-youths-under-the-dish-project-in-tigray/

DISH Mekelle. (2023, April 23). *DISH programme created a convenient environment for students to learn in difficult situation – Kokob Gebru*. Digital Innovation and Skills Hub. https://aun.mu.edu.et/Dish-portal/2023/03/21/dish-programme-created-a-convenience-for-students-in-difficult-situations-kokob-gebru/

Gilligan, C. (1993). *In a different voice: Psychological theory and women's development*. Harvard University Press.

Gillborn, D., Warmington, P., & Demack, S. (2023). QuantCrit: Education, policy, 'Big Data' and principles for a critical race theory of statistics. In *QuantCrit* (pp. 10–31). Routledge.

Gray, B. J., & Colling, M. (2021). Understanding Interoperability in Humanitarian Aid Organizations. In ISCRAM (pp. 439-447).

Kievit, R., Kawu, A. A., Van Reisen, M., O'Sullivan, D., & Hederman, L. (2024). Exploring the design of low-end technology to increase patient connectivity to electronic health records. In *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies – Volume 2: HEALTHINF* (pp. 194–200). https://doi.org/10.5220/0012460900003657

Lin, Z. (2025) Implementation of De Novo FAIRification in Relational Legacy Systems: the Case of The Electronic Medical Record system for maternal health in Afya.ke. In Van Reisen, M., Amare, S. Y., Maxwell, L. & Mawere, M. (Eds.), *FAIR data, FAIR Africa, FAIR world: Internationalisation of the Health Data Space.* Bamenda: Langaa

Kirchner, J., Angwin, S., Mattu, J., & Larson, L. (2016). Machine bias: There's software used across the country to predict future criminals. And it's biased against Blacks. *ProPublica*. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Donnelley, R. R. (2009). *Curriculum for excellence: Building the curriculum 4: Skills for learning, skills for life and skills for work*. Scottish Government

Green, S., Hillersdal, L., Holt, J., Hoeyer, K., and Wadmann, S. (2023) 'The Practical Ethics of Repurposing Health Data: How to Acknowledge Invisible Data Work and the Need for Prioritization', *Medicine, Health Care and Philosophy, 26*(1), 119–132. https://doi.org/10.1007/s11019-022-10128-6

Jati, P. H., Lasroha, M. S. T., & Van Reisen, M. (2025). Data privacy and interoperability in Indonesia's SATUSEHAT platform. In Van Reisen, M., Amare, S. Y., Maxwell, L. & Mawere, M. (Eds.), *FAIR data, FAIR*

Africa, FAIR world: Internationalisation of the Health Data Space. Bamenda: Langaa
Mons, B., Neylon, C., Velterop, J., Dumontier, M., da Silva Santos, L. O. B., & Wilkinson, M. D. (2017). Cloudy, increasingly FAIR; revisiting the FAIR data guiding principles for the European Open Science Cloud. *Information Services & Use, 37*(1), 49–56. https://doi.org/10.3233/ISU-170824
Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.
Oladipo, F., Folorunso, S., Ogundepo, E., Osigwe, O., & Akindele, A. (2022). Curriculum development for FAIR data stewardship. *Data Intelligence, 4*(4), 991–1012. https://doi.org/10.1162/dint_a_00140
Rancho BioSciences. (2025, March 3). Unlocking data potential: Understanding FAIR vs open data in life sciences. *Rancho BioSciences*. https://ranchobiosciences.com/2025/01/unlocking-data-potential-understanding-fair-vs-open-data-in-life-sciences/
Rubin, R. (2018). Addressing barriers to inclusion of pregnant women in clinical trials. *JAMA, 320*(8), 742–744. https://doi.org/10.1001/jama.2018.8401
Smits, K., Upase, M., Pandey, N., Bala, S., & Van Reisen, M. (2025). FAIR-data implementation for analysis of research data in human trafficking and migration. In Van Reisen, M., Amare, S. Y., Maxwell, L. & Mawere, M. (Eds.), *FAIR data, FAIR Africa, FAIR world: Internationalisation of the Health Data Space*. Bamenda: Langaa
UNA Europa Fair Data Hackathon for PhD candidates. (n.d.). *LUMC*. https://www.lumc.nl/over-het-lumc/agenda/2025/april/una-europa-fair-data-hackathon-for-phd-candidates/
UNESCO IBE. (2011). *What makes a good quality school curriculum?* [Background paper developed by Stabback, P., Male, B., & Georgescu, D.]. UNESCO IBE. (Unpublished)
Utami, I.K. (2025) Satu Data Indonesia and FAIR Data: Advancing Coherent Data Management in Government Administration. In M. Van Reisen, M., Amare, S. Y., Maxwell, L. & Mawere, M. (Eds.), *FAIR data, FAIR Africa, FAIR world: Internationalisation of the Health Data Space.* Bamenda: Langaa.
Van der Graaf, R., van der Zande, I. S., den Ruijter, H. M., Oudijk, M. A., van Delden, J. J., Oude Rengerink, K., & Groenwold, R. H. (2018). Fair inclusion of pregnant women in clinical trials: An integrated scientific and ethical approach. *Trials, 19*, Article 78. https://doi.org/10.1186/s13063-018-2472-z
Van Reisen, M., Oladipo, F., Stokmans, M., Mpezamihigo, M., Folorunso, S., Schultes, E., Basajja, M., Aktau, A., Yohannes Amare, S., Tadele Taye, G., Hadi Purnama Jati, P., Chindoza, K., Wirtz, M., Ghardallou, M., van Stam, G., Ayele, W., Nalugala, R., Abdullahi, I., Osigwe, O., Graybeal, J., Abrha Medhanyie, A., Abubakar Kawu, A., Liu, F., Wolstencroft, K., Flikkenschild, E., Lin, Y., Stocker, J., Musen, M.A.

(2021). Design of a FAIR digital data health infrastructure in Africa for COVID-19 reporting and research. In: *Advanced Genetics, 2*(2). http://doi.org/10.1002/ggn2.10050

Van Reisen, M., Oladipo, F., Mpezamihigo, M., Plug, R., Basajja, M., Aktau, A., Purnama Jati, P.H., Nalugala, R., Folorunso, S., Amare, Y.S., Abdulahi, I., Afolabi, O.O., Mwesigwa, E., Taye, G.T., Kawu, A., Ghardallou, M., Liang, Y., Osigwe, O., Medhanyie, A.A., Mawere, M. (2022). Incomplete COVID-19 data: The curation of medical health data by the Virus Outbreak Data Network-Africa. *Data Intelligence 4*(4). https://doi.org/10.1162/dint_e_00166

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data, 3*, 160018. https://doi.org/10.1038/sdata.2016.18

Zhang, S., Dos Santos Vieira, B., Bernabé, C. H., Abaza, H., Benis, N., Cámara, A., Cornet, R., Le Cornec, C. M. A., ‘t Hoen, P. A. C., Schaefer, F., van der Velde, K. J., Swertz, M. A., Wilkinson, M. D., Jacobsen, A., & Roos, M. (2022). Towards FAIRification of sensitive and fragmented rare disease patient data: Challenges and solutions in European Reference Network registries. *Orphanet Journal of Rare Diseases, 17*, 436. https://doi.org/10.1186/s13023-022-02558-5