Interview Report on Data Science Practice

Introduction

For this report, I interviewed a professional with four years of experience working as a data scientist at a mid-sized technology company in the healthcare analytics sector. The individual (referred to as "Lena") has been directly involved in machine learning model development, data governance, and cross-functional collaboration with clinicians and business teams.

Our conversation centered around ethical challenges Lena has faced, both technical and workplace-related. Two particularly memorable issues stood out: (1) pressures to deploy models with insufficient validation, and (2) bias and fairness concerns in demographic data usage.

Ethical Issue 1: Data Privacy and Pressure to Share Identifiable Records

Summary of Lena's Experience

Lena recalled a situation where a partner organization requested detailed patient-level data to run their own analyses. The request included identifiers like dates of birth and partial addresses. While the partner argued that such information was "necessary for research accuracy," sharing it would have violated HIPAA guidelines and company privacy policies.

Follow-up Responses:

- Handled well? Lena felt this issue was handled moderately well. Her team pushed back, explaining the risks and legal implications, and ultimately released only de-identified datasets. However, leadership initially leaned toward granting the request to preserve the business relationship.
- Difficulty in ethical action? Yes, the partner was influential and represented significant future revenue. Saying "no" risked damaging trust and contracts, which made the ethical stance harder to uphold.

Analysis and Reflection

What struck me here was how business partnerships can test privacy safeguards. It was surprising that leadership's first instinct was accommodation rather than strict compliance.

This could have been resolved more ethically by:

- Establishing a clear contractual boundary around data use before collaboration began.
- Building automated de-identification pipelines so that "privacy-first" becomes the default.
- Empowering technical staff like Lena to veto unsafe data-sharing requests.

Through a deontological framework, protecting patient privacy is a duty, not a negotiable condition. Even if business relationships are at stake, the obligation to uphold confidentiality should take precedence.

Ethical Issue 2: Transparency in Model Performance

Summary of Lena's Experience

Another challenge Lena described involved reporting model performance to a client hospital. The model predicting readmission risk had moderate accuracy overall but performed poorly for certain subgroups, such as older adults. Some managers wanted to present only the overall accuracy score (which looked stronger) without disclosing subgroup weaknesses.

Lena strongly objected, arguing that clinicians relying on the model needed a full picture of its limitations. After heated discussions, the team compromised by including subgroup performance in a technical appendix—something most clients might never read.

Follow-up Responses:

- Handled well? Lena felt this was not handled well. Transparency was compromised, as the most important limitations were effectively hidden from frontline decision-makers.
- Difficulty in ethical action? Yes, the organization feared that highlighting weaknesses would undermine confidence in the product and hurt sales. Speaking up risked Lena being seen as "too negative."

Analysis and Reflection

This issue demonstrates how commercial incentives can conflict with ethical transparency. I found it concerning that leadership prioritized optics over accuracy, especially in a healthcare context where poor model performance can directly affect patient care.

What might have helped:

- Standardizing performance reporting that requires subgroup breakdowns.
- Creating joint review panels with clinicians to ensure results are presented responsibly.
- Linking executive incentives to long-term trust, not short-term client acquisition.

From a utilitarian perspective, hiding subgroup weaknesses may maximize short-term benefit (sales, reputation) but produces long-term harm by eroding trust and risking patient safety. True utility lies in candid disclosure and iterative improvement.

Ethical Issue 3 Workplace Pressure to Work Beyond Scope

Lena also mentioned feeling pressured to perform tasks outside her professional scope, such as making clinical recommendations rather than limiting her input to data-driven insights. While not as acute as the first two issues, this blurred professional boundaries and created potential liability.

What I Found Most Interesting

From Issue 1, I found it reassuring that Lena and her colleagues ultimately resisted the push to share identifiable data. It reinforced that strong professional norms can check business pressure.

From Issue 2, I was most concerned by how easy it was to “bury” critical limitations in an appendix. This shows how transparency can be compromised in subtle ways, not just outright deception.

Reflecting on Lena’s experiences, several broader lessons about ethical practice in data science became evident. One key takeaway is the importance of ethical literacy not just among technical staff, but across leadership and business teams**.** In both the privacy and transparency issues Lena described, the tension often arose not from malicious intent, but from a lack of awareness or prioritization of ethical consequences at the leadership level.

Lena emphasized the value of embedding ethics into everyday workflows, not treating them as separate checkpoints. For example, she suggested regular "ethics standups" alongside agile sprints, where team members raise potential concerns early and normalize these discussions as part of the project lifecycle.

She also reflected on how isolating it can feel to raise ethical objections in a fast-paced, business-driven environment. This points to the need for stronger organizational support systems, such as anonymous reporting mechanisms or designated ethics advocates,so professionals don’t have to bear the burden alone.

Lastly, Lena noted that ethical clarity often comes down to courage and communication. The moments where she felt most conflicted were not due to a lack of knowledge, but fear of professional repercussions or being seen as an obstacle. Her advice to early-career data scientists is to build confidence in ethical reasoning just as much as technical skill, and to seek allies who value integrity over short-term gains.

Conclusion

The interview underscored the dual ethical pressures in data science: (1) safeguarding privacy despite external demands, and (2) ensuring transparency despite internal commercial pressures.

Applying deontological ethics to Issue 1 highlights the duty to protect privacy. Applying utilitarianism to Issue 2 shows that the greatest good comes from honesty, even if it slows short-term business growth.

Overall, Lena's experiences suggest that stronger organizational policies, reporting standards, and protections for professionals raising concerns would make it easier to act ethically. Embedding ethics at both the technical and workplace levels remains a critical challenge for the field.