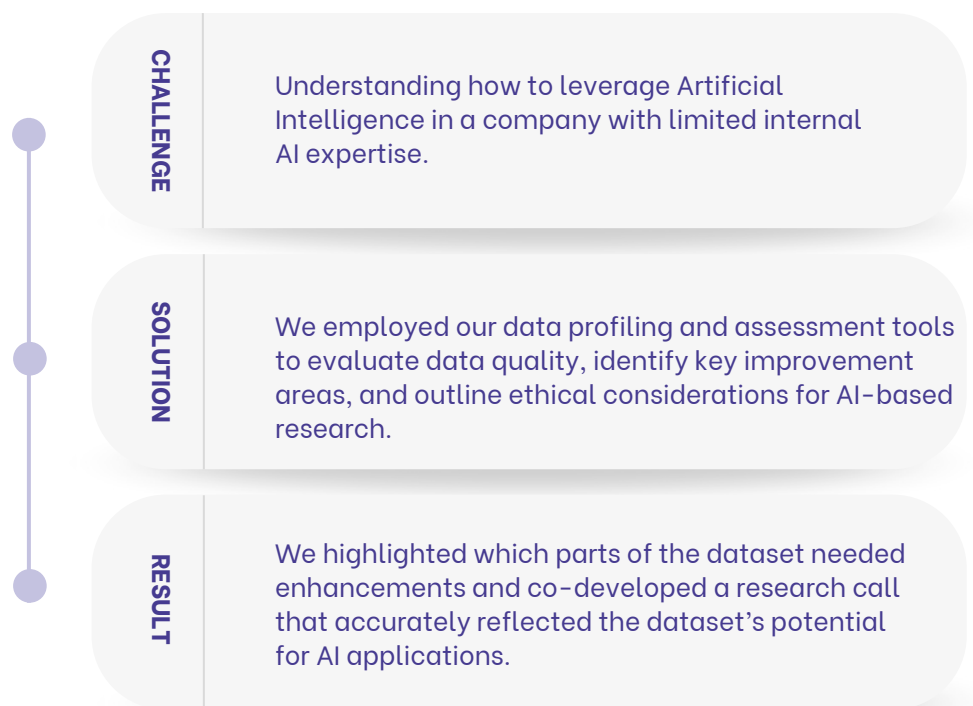# clearbox AI

humans && machines

Data Assessment & AI Feasibility

Assessing the AI readiness of a health dataset

clearbox.ai

Introduction

Assessing the AI readiness of a hospital dataset is vital to ensure that advanced machine learning solutions can be successfully implemented and deliver meaningful insights. In this use case, we collaborated with a foundation who wished to organize and fund a research call. Their goal was to determine how effectively the hospital datasets could be used for AI research, set realistic project boundaries, and ultimately co-author a call that would leverage this data for future studies.

CHALLENGE

Understanding how to leverage Artificial Intelligence in a company with limited internal AI expertise.

SOLUTION

We employed our data profiling and assessment tools to evaluate data quality, identify key improvement areas, and outline ethical considerations for AI-based research.

RESULT

We highlighted which parts of the dataset needed enhancements and co-developed a research call that accurately reflected the dataset's potential for AI applications.

Challenge

Medical data typically involves complexities like privacy constraints, fragmented records, and inconsistent formats. The foundation needed to confirm that the hospital dataset could support AI models—from ensuring data completeness and representativeness, to verifying that compliance and ethical considerations were addressed. They sought our expertise to profile the data thoroughly and assess any risks or limitations that could affect the viability of AI projects.

Solution

Leveraging our data profiling and assessment methodology, we began by examining the dataset's overall health—spotting missing values, outliers, and potential biases. We also assessed metadata, privacy risks, and data linkages to gauge whether it was suitable for machine learning tasks. Having established a clear baseline, we collaborated with the foundation to co-draft the research call, highlighting the dataset's readiness and proposing enhancements. This alignment allowed the foundation to invite proposals that were both innovative and feasible, ensuring realistic project outcomes.

Result

Our analysis offered a structured roadmap for improving the dataset and making it more AI-ready. The foundation successfully published a targeted research call grounded in our findings. By clarifying data capabilities and limitations, the call attracted proposals that could maximize the dataset's value while adhering to ethical and privacy standards. As a result, the foundation is now poised to fund research projects with a solid basis for innovative and responsible AI solutions.