# ePotentia, Michael Sluydts, PhD., CEO [https://www.epotentia.com](https://eur02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.epotentia.com%2F&data=05%7C02%7Cgreta.boonen%40vito.be%7C39f28f02f1614e0654d808dd5b32cebe%7C9e2777ed82374ab992782c144d6f6da3%7C1%7C0%7C638766995969881422%7CUnknown%7CTWFpbGZsb3d8eyJFbXB0eU1hcGkiOnRydWUsIlYiOiIwLjAuMDAwMCIsIlAiOiJXaW4zMiIsIkFOIjoiTWFpbCIsIldUIjoyfQ%3D%3D%7C0%7C%7C%7C&sdata=xoQeFZjs0A02kUm2efB61SCxtrmBg3hk5NiwtbzGABs%3D&reserved=0)

# Pitch Title: From Real Scarcity to Synthetic Abundance: how generative AI can complete materials datasets

Transforming Scarcity into Synthetic Abundance: Leveraging Generative AI to Enhance Materials Datasets

Abstract: Modern AI techniques can greatly accelerate materials R&D, but are often data-hungry. Creating new datasets can be time-consuming and expensive, while collecting historical data is often not much better due to incomplete labels and inconsistent quality. Generative AI can help solve this problem by creating synthetic data to overcome data scarcity issues. But this approach brings new challenges: How good does synthetic data need to be? Can it become too good, raising ethical considerations about scientific integrity? We'll explore these questions using a steel microscopy use case from the AID4GREENEST project, a Horizon Europe collaboration focused on accelerating green steel production through AI-accelerated characterization, including Flemish partners ePotentia, UGent, and OCAS. We are initially investigating how AI combined with expert knowledge, can rapidly label historical data, which can then be used to generate synthetic microstructure images that are virtually indistinguishable from real samples. This project highlights both the tremendous potential of AI to accelerate the development of green steels through enhanced characterization and the important ethical considerations that arise as synthetic data approaches perfect fidelity - considerations that can may have far-reaching impact for the wider materials science community as these techniques become more widespread.