# Technical Specification Double-side Page

1. **TECHNICAL SCOPE:** Summarise the mock-up devised during the EXPLORE phase: how have you addressed the challenge/Theme Challenges and tackled with its requirements and data. Include a diagram.

The mock-up portrays the deep learning model configured to identify objects classified. However, as we gear up for the REACH Challenge, our focus now pivots from object detection to object segmentation. This strategic change in focus is expected to enhance the system's efficiency and accuracy in construction site surveillance and management.

2. **ALGORITHMS, TOOLS AND CONCLUSIONS:** Detail the algorithms and tools identified to accomplish the challenge/Theme Challenges. Show clear understanding of the used REACH dataset/s and addressed challenge/Theme Challenges.

The main algorithms and tools to be used for project AI-SCEND are the following:

**A suite of analysis tools** from the REACH Toolsuite.

**Deep learning algorithms** used for tasks related to image analysis, including object detection and segmentation. These algorithms are great candidates for the base model of our approach for a secure and robust system for detailed object segmentation in construction sites.

**A web-based tool** designed to annotate video and images for computer vision algorithms. It provides a user-friendly interface and a powerful backend with advanced features like an interpolation of shapes between keyframes, automatic annotation using deep learning models, and collaboration abilities.

3. **SCALABILITY AND FLEXIBILITY OF THE SOLUTION:** Discuss whether the solution can truly cope with humongous and increasing datasets and how flexible it is to adapt to other related domains and integrate into Data Value Chains (DVC).

VERTLINER aims to capture part of the BIM ecosystem providing accurate digital measurements of building assets, thus informing BIM systems with direct data from physical assets. As the BIM workflow tends to become a leading industry standard for construction site operations, its integration with VERTLINER's asset assessment system will make the proposed solution scalable to the vast majority of the projects under development.

As new data is collected, AI models can be continually trained and fine-tuned to improve performance, effectively adapting to the growth of the dataset.

On the subject of flexibility, AI-SCEND demonstrates adaptability to related domains. Given that our core processes involve image data collection, labelling, and the training and deployment of an image segmentation model, these processes can be adapted to various domains where image data is relevant.

When it comes to integration into Data Value Chains (DVC), our solution fits naturally at several stages. During the data acquisition step, our image data collection process can be integrated. In the data pre-processing step, our labeling methodology helps in data validation and cleaning. Lastly, in the data analysis step, our model can be utilized to extract valuable insights from the data. Therefore, our solution can be a valuable component in a DVC, enhancing the value derived from data at various stages.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951981

4. **DATA GOVERNANCE AND LEGAL COMPLIANCE:** Describe the security level of the proposed solution, i.e. how authentication, authorization policies, encryption or other approaches are used to keep data secure. Explain how will be compliant with the current data legislations concerning security and privacy (e.g. GDPR).

Project AI-SCEND is designed with a high level of security and adherence to legal compliance, employing several tools and strategies to maintain data privacy and security.

We prioritize code-level security from the REACH Toolsuite and plugin for C++ programs. These tools allow us to ensure confidentiality, enforce strict coding standards, and protect our algorithms against potential security flaws, effectively identifying and mitigating potential data leakages.

To ensure secure data exchange between our system and the data providers, we use established data-sharing platforms like Google Workspace Drive. Access to these platforms is controlled through authentication and authorization policies, ensuring that only authorized users can access the data. This approach maintains the privacy and confidentiality of the data being exchanged.

In terms of legal compliance, our solution is designed to be fully compliant with GDPR. The images collected are strictly limited to these objects and do not include people or other sensitive information. This approach ensures that we maintain privacy and adhere to data protection regulations.

5. **QUALITY ASSURANCE AND RISK MANAGEMENT:** Describe the quality process planned for the final product. Technologically, which are the potential risks in all the phases of the project (design of the solution, development, testing, deployment...) and indicate mitigation plans to still fulfil the challenge/Theme Challenges and data provider requirements.

The quality of AI-SCEND is intrinsically tied to effective risk management throughout all project phases. Successful handling of potential issues directly influences the quality of the end product. Let's delve into some of the critical challenges that might emerge.

Design issues can arise within the solution. By employing agile methodologies, we can iteratively address and correct any problems that surface during the project's progression. During the development of the model, several high-impact risks can occur:

- Limited quality datasets: To tackle this, we plan to utilize data augmentation, expanding and diversifying our AI model training dataset.
- Class imbalance: This issue will be handled using techniques like oversampling, undersampling, or class-weighting during model training.
- Image quality sensitivity: Enhancement techniques will be applied for better image quality, and efficient architectures will be used to balance performance with high-resolution images.

Two key considerations ensure model quality:

- Overfitting: This will be mitigated using cross-validation and regularization techniques.
- Lack of Generalization: We'll ensure a diverse, representative training dataset and possibly employ data augmentation.

Finally, during deployment, we anticipate potential challenges:

- Computational Resources: We plan to apply model optimization techniques, such as quantization and pruning. Utilizing NJ modules and their deep learning capabilities will further address computational resource constraints.
- Model Drift: Continuous model monitoring and periodic retraining will be implemented to handle this issue.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951981