Balsam: A Collaborative Platform to Support ML and ML-Ops initiatives

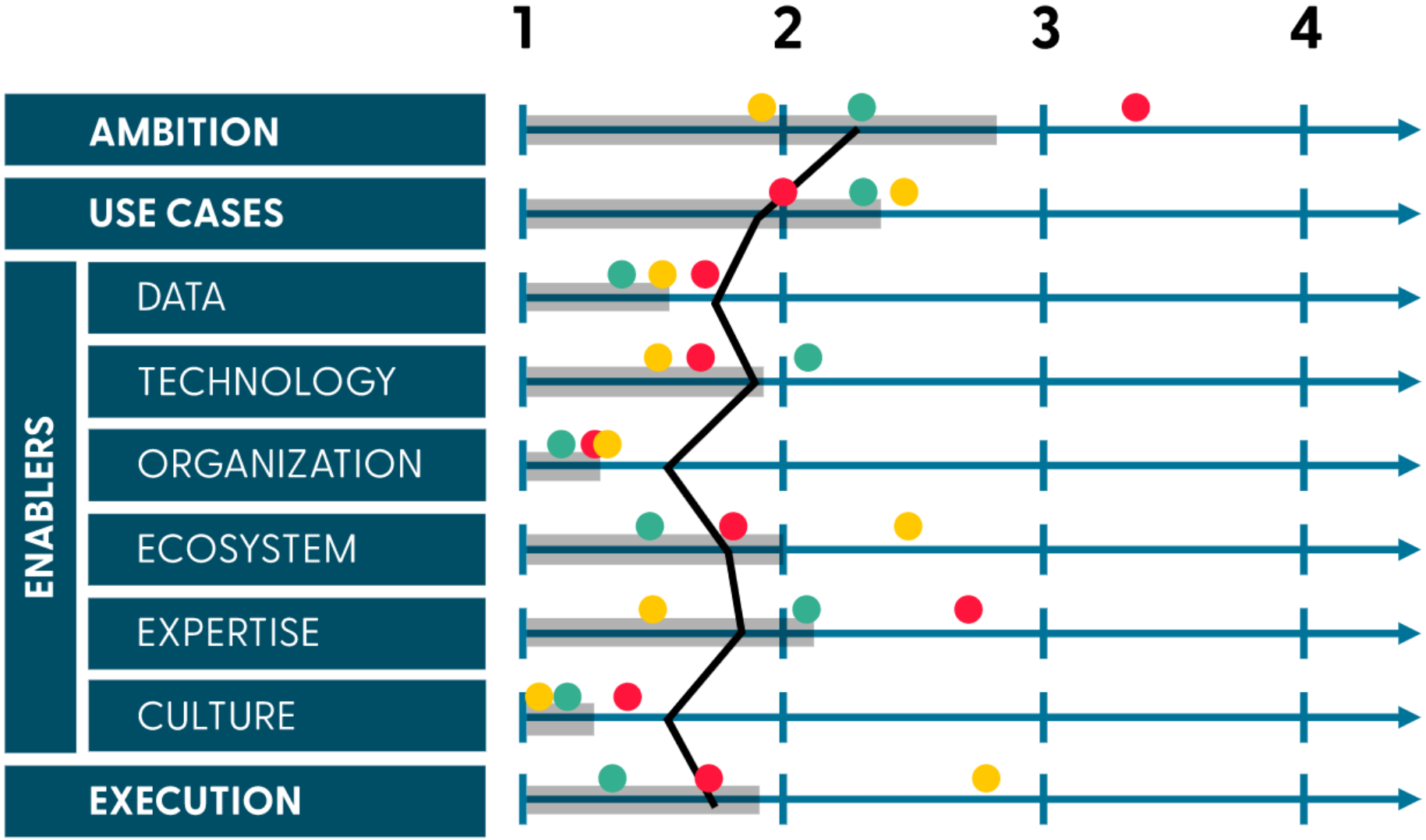
Jakob Engdahl
+ChatGPT & midjourney

SCB

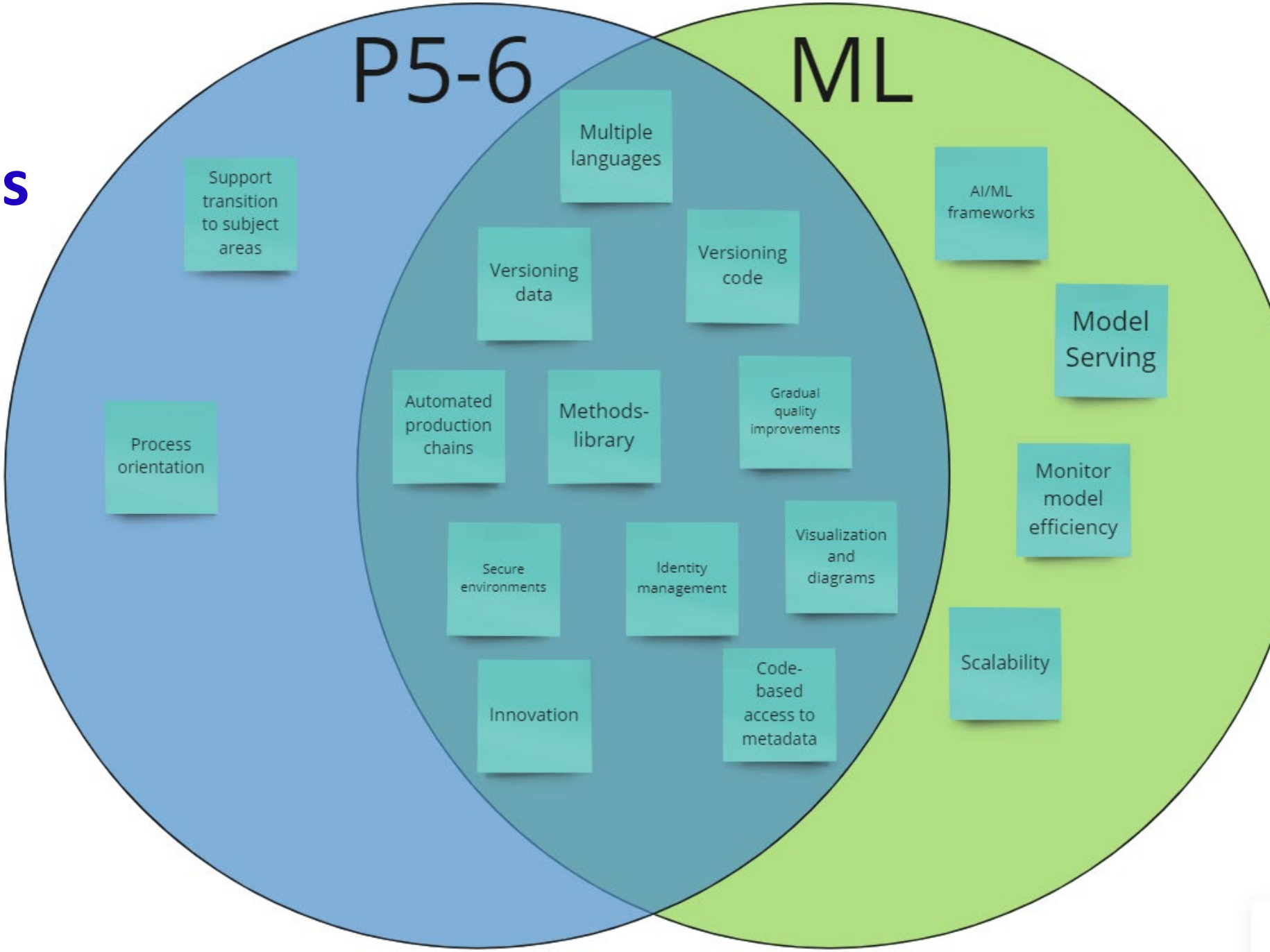
The BALSAM Project: An Overview

- Collaboration between Statistics Sweden, AI Sweden, and Örebro University
- Goal: Accelerating machine learning beyond innovation
- Supporting ML and regular statistical production
- Overcoming challenges and ensuring scalability

Scaling Machine Learning



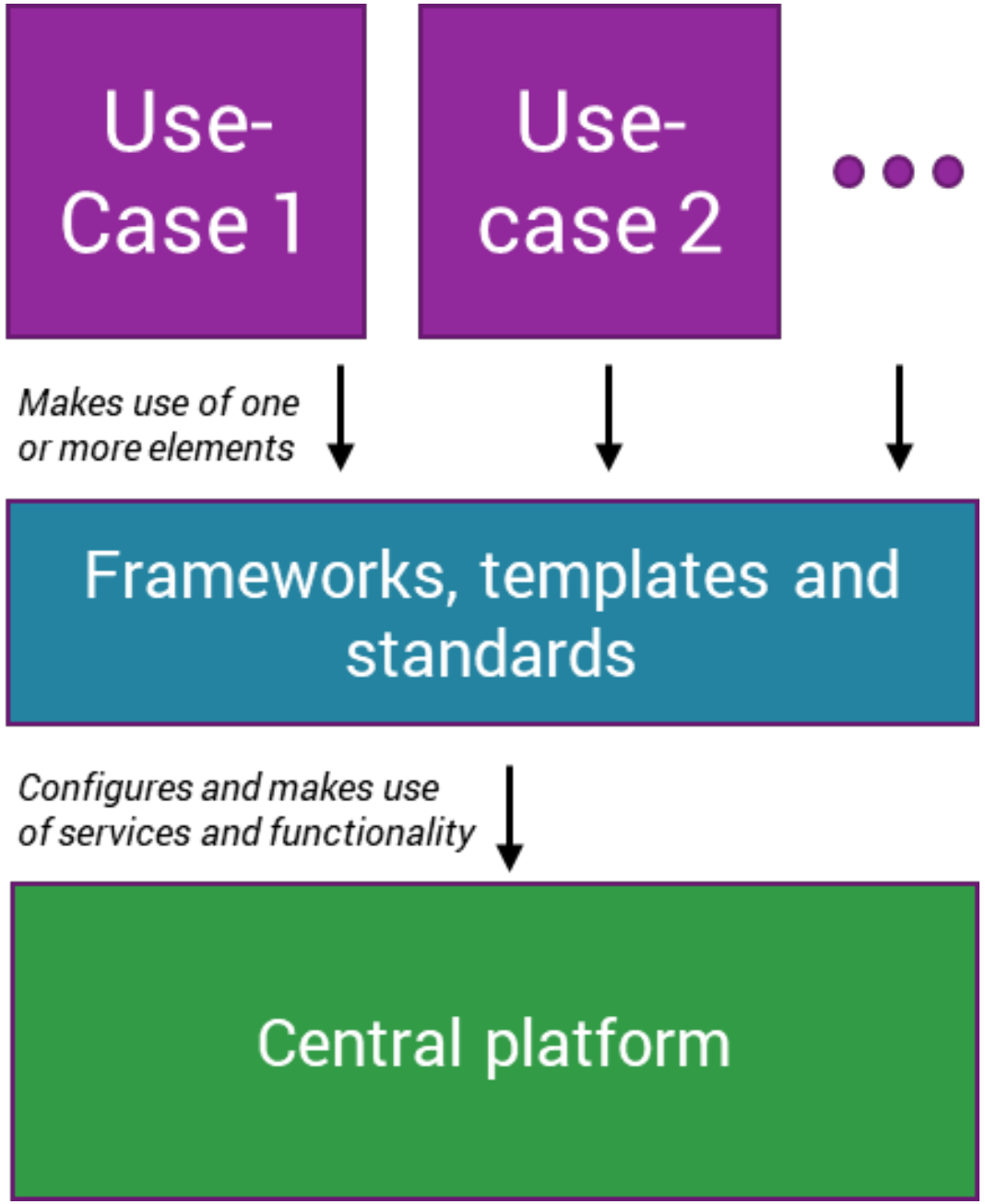
Similar Requirements for Statistical Production



SCB

BALSAM Architecture: A High-Level Overview

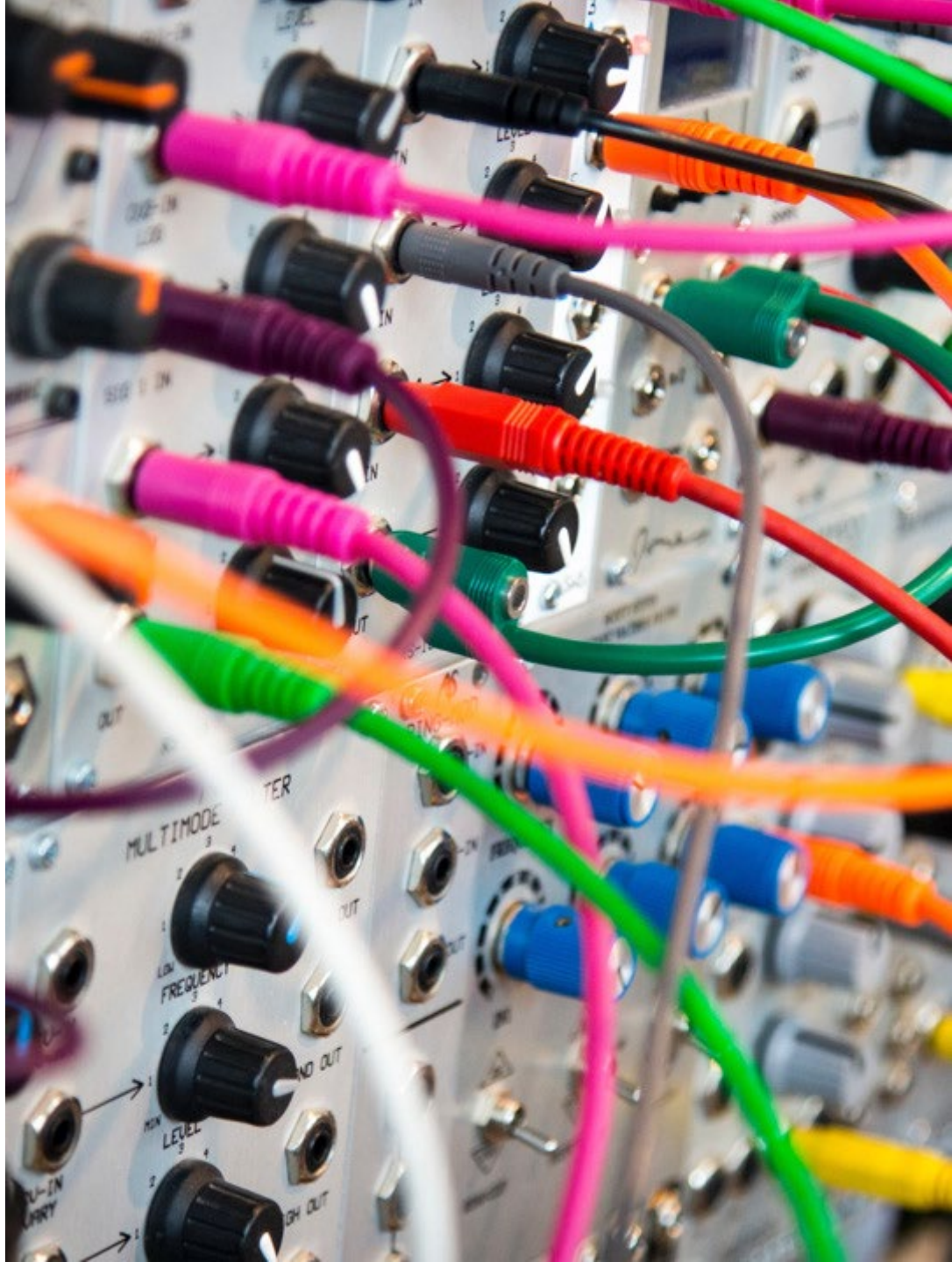
- Layered architecture for the BALSAM platform
- Focus on loose coupling and flexibility
- Importance of APIs for integration and automation
- Leveraging open-source services and adapters



SCB

Use-Case Level: Customization and Adaptability

- Specific setup for each project/use-case
- Common structure and services across use-cases
- Extensive use of Jupyter Notebooks (Elyra AI)
- Supporting both ML projects and regular statistical production
- Versioning and naming standards for tracking changes



Frameworks, Templates, and Standards: Harmonizing Usage

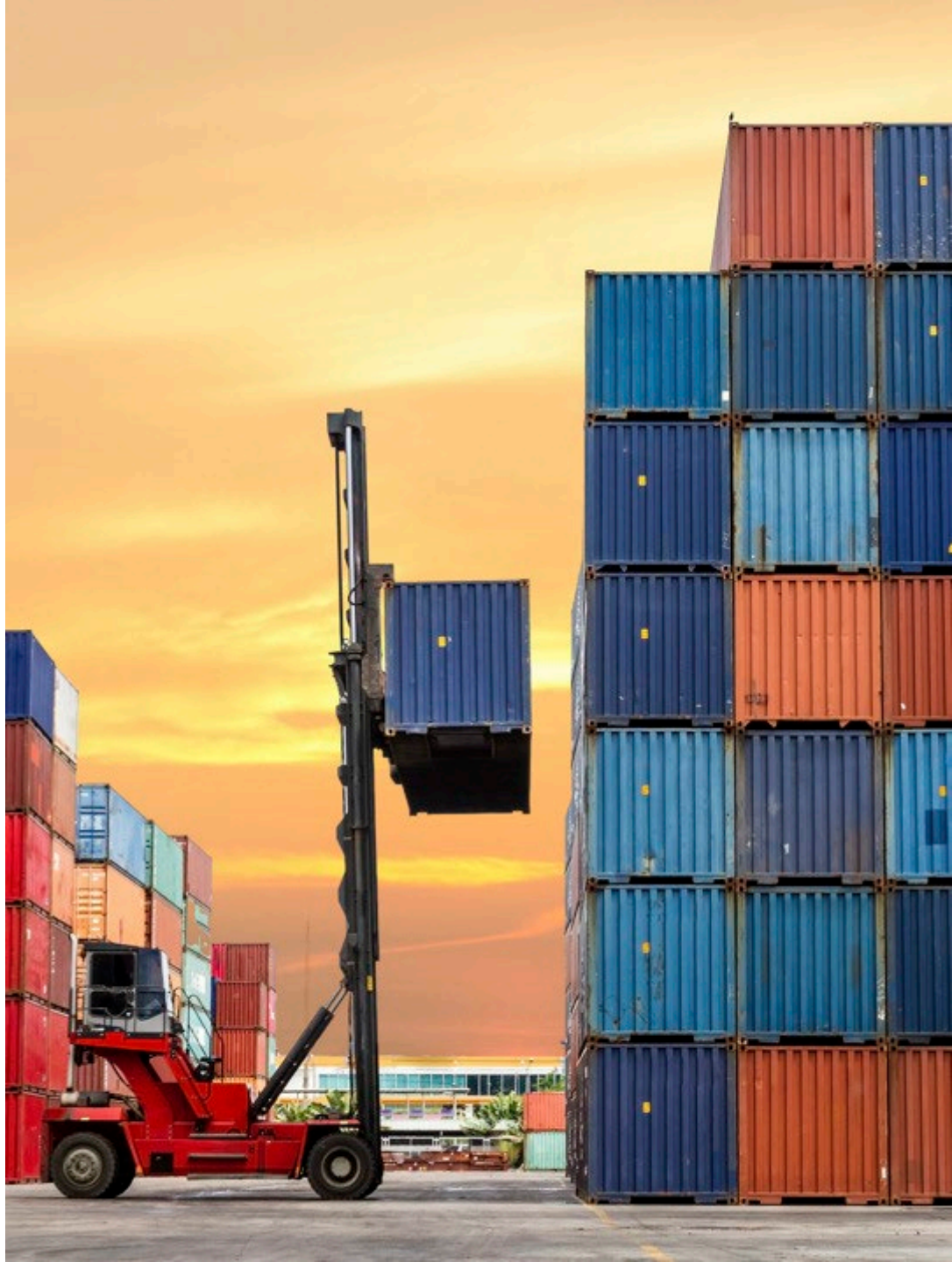
- Pre-configured project templates for different use-cases
- Harmonizing services, configurations, and access
- Avoiding unnecessary variance between projects
- Simplifying adherence to policies and standards
- Federated knowledge base

SCB

Central Platform: Shared Functionality

- Functionality shared by all use-cases and projects
- Project management and resource connectivity
- Containerization environment for service creation
- Versioning of code, documents, and metadata (Gitlab)
- Versioning of data (Minio and future integration)
- Collaboration platform

SCB



Current Status and Future Directions

- Continue development until end of year
- Expand the functionality of the "Frameworks and Standards" layer
  - Project templates
  - Federated knowledge base
- Integrate BALSAM into new data center
- Outreach to the community

Questions and comments!