



do more
feel better
live longer



PhUSE SDE Netherlands: “Oh API Days”

February 16th 2017
SDD/LSAF Interoperability

Xavier Deflorenne (GSK)
Wouter Van de Weghe (SAS Institute)



Interoperability, a Real-Life Case

Overview



- **Life Science Analytics Framework (Recap)**
- Overview of SDD/LSAF at GSK
- The APIs
 - The need for APIs
 - SAS Native APIs
 - SAS Custom APIs
 - GSK Custom APIs
- Use cases
 - User management
 - Object management
 - Synchronization of data archives
- SAS future (Viya)

Interoperability, a Real-Life Case

SAS® Life Science Analytics Framework (1/3) - Recap



current status



- **Managing, analyzing, reporting, and reviewing (clinical) research information in the same analytical environment**
 - Repository provides seamless integration with the SAS program development (authoring environment) and execution environment, along with workflow capabilities.
 - *NEW : Authoring Environment compared to SAS DRUG DEVELOPMENT (SDD)*
 - *NEW : Business Process Model and Notation (BPMN) 2.0 standard support compared to SAS DRUG DEVELOPMENT (SDD)*
- **Cloud based (Private Cloud @ SAS)**
 - Provides more flexibility, lower costs, greater scalability, ease of use and, if done correctly, increased security and disaster recovery.
 - Opportunity for IT to innovate and support the rest of the company to achieve its goals.

Interoperability, a Real-Life Case

SAS® Life Science Analytics Framework (1/3) - Recap



Interactive Sessions (example – LSAF 4.7)

The screenshot displays the SAS Life Science Analytics Framework (LSAF) Interactive Sessions interface. The top bar indicates the session is 'Session 6'. The left pane shows a 'Repository' view with a tree structure under 'Phuse2016-Workshop (Compound)'. The tree includes 'Files', 'Phuse2016-Workshop (Study)', and 'Files'. Under 'Files', there is a 'CSR_01' folder containing 'analysis', 'output', and 'programs'. The 'programs' folder contains several files, including 'adds_table.sas', 'Create_ADAM.job', 'Create_ADAM.log', 'Create_ADAM.lst', 'Create_ADAM.mnf', 'Create_adds_table.job', 'Create_adds_table.log', 'Create_adds_table.lst', 'Create_adds_table.mnf', 'createADBE.sas', 'createADDS.sas', 'createADSU.sas', and 'setup.sas'. The 'setup.sas' file is selected and highlighted in blue.

The central pane shows the code editor for 'setup.sas'. The code includes a DATA statement, a LENGTH statement, a STRING variable, a SYMGET function, a PRXCHANGE function, a CALL SYMPUT function, and a RUN statement. The code also includes a %put statement, a libname statement, and a %let statement.

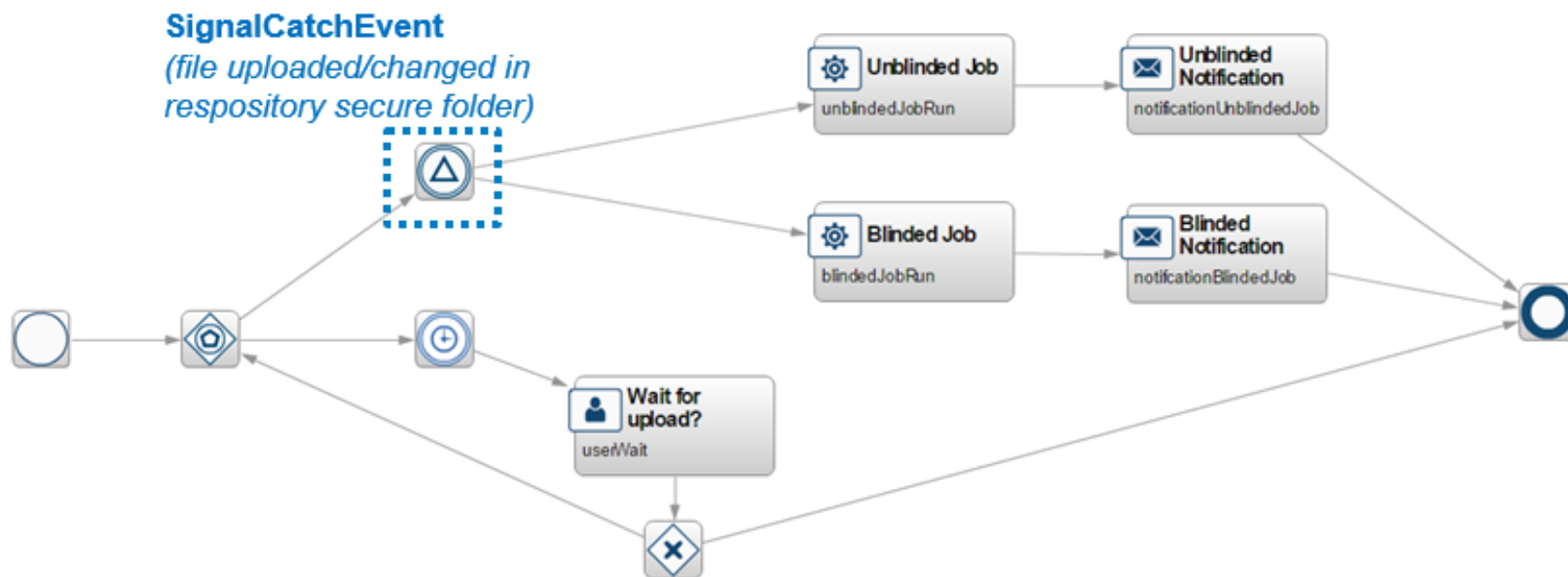
The bottom pane shows the 'Session Log' for 'Session 6'. The log includes a URL, a message about the SAS news file, and a NOTE about SAS Initialization used (Total process time).

Interoperability, a Real-Life Case

SAS® Life Science Analytics Framework (1/3) - Recap



Workflow BPMN standard (example – LSAF 4.7)



Interoperability, a Real-Life Case

SAS® Life Science Analytics Framework (2/3) - Recap



▪ Collaboration & Accessibility

- Internally and externally (geographically remote) i.e. sponsors, development partners, contract research organizations (CRO), as well as compliance and regulatory stakeholders.
- Direct access to research content, clearly organized and searchable.

▪ Configurable User Interface

- Highly flexible (configurable) permission model → privileges & permissions.

▪ High-performance (Remote) Application Programming Interface (extend & automate)

- Java API and SAS Macro API ... moving to RESTful API

▪ Complete audit history and robust version control system managing change history of all files

▪ Job concept and Manifest file

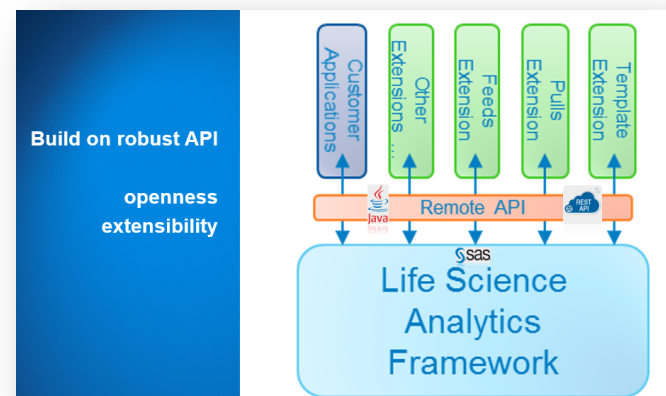
- Job Concept → reproducibility and usability (for non-technical savvy users)
- Manifest File → transparency and traceability

Interoperability, a Real-Life Case

SAS® Life Science Analytics Framework (3/3) - Recap



- **Standards metadata management** (driving automation)
 - Standard (Global) Level (SDTM & SEND) → “*Data Standards*” Module
 - Study Level → **Study Specification**: Empty Datasets + Define XML 2.0 creation → “*Studies*” Module
- **Support Visual Analytics** (integration) -- *NEW compared to SAS DRUG DEVELOPMENT (SDD)*
- **Expanding Suite of Extensions** (optional modules build on robust LSAF API)
 - ‘Feeds’ to remote locations (via file share or SFTP)
 - ‘Pulls’ from remote locations (example: Lab; IxR; ... vendors)
 - ‘Templates’ to automate creation of folder structure within repository incl. administration
- **Last but not least, some SDD/LSAF numbers ...**
 - + 35 customers
 - + 15 years of R&D roadmap
 - User connection events organized frequently (last one in Brussels/Belgium Oct 2016 @ Janssen / J&J Campus)



Interoperability, a Real-Life Case

Overview



- Life Science Analytics Framework (Recap)
- **Overview of SDD/LSAF at GSK**
- The APIs
 - The need for APIs
 - SAS Native APIs
 - SAS Custom APIs
 - GSK Custom APIs
- Use cases
 - User management
 - Object management
 - Synchronization of data archives
- SAS future (Viya)

SDD/LSAF Interoperability

Overview of SDD/LSAF at GSK



History

- SDD 3.5 Launched in November 2009; migrated from SAS Windows Desktop.
- Migrated to SDD 4.3 in October 2015
- Novartis integration (Novartis SDD 3.5-> GSK SDD 4.3) ended in early 2017.

Facts and figures

- SDD is hosted in Cary (NC, US).
- Holds about **1M folders and 7.5M files** for a total of 6+TB
- About **950 active users** in SDD (incl. GSK users, CROs, and 3rd parties).
- 200 projects, 2000+ studies and 6000+ analysis have restricted accesses.
- Strict security model is applied down to the study level. → **34K+ groups**.
- SDD is used mainly by Data Management and Bio-Statisticians, but also as an exchange platform between GSK and 3rd parties.
- High turnover in access for some profiles.

Integration

- SDD coexists with SAS Windows platform Server and Workstations (legacy use and backup).
- SDD is integrated with other systems (source data, user management)

Interoperability, a Real-Life Case

Overview



- Life Science Analytics Framework (Recap)
- Overview of SDD/LSAF at GSK
- **The APIs**
 - The need for APIs
 - SAS Native APIs
 - SAS Custom APIs
 - GSK Custom APIs
- Use cases
 - User management
 - Object management
 - Synchronization of data archives
- SAS future (Viya)

Needs

- High volume of objects to manage (users, files, folders, permissions)
- Integration and synchronization with other systems: user management, data source, cross-platform integrations (programs)

► Solutions for automation and integration

Solutions

- Using the SDD APIs for automation and integration
 - Automation tasks to facilitate the management of the platform: creation of users, group, management access rights (ACLs ...)
 - Integration tasks between SDD and the GSK environment: login/logout, data transfers
- 3 APIs families in use:
 - SAS Native APIs
 - SAS Custom APIs
 - GSK Custom APIs

SAS Native APIs: provided by SAS along with SDD/LSAF.

- GSK is currently using:
 - Java API release 1.6
 - SAS Macro API release 1.3

- Usage:
 - Creation/deletion of Users, Groups, Containers, Roles and Privileges
 - Management of groups
 - Update of permissions (ACLs) on folders and files.
 - Login/logout
 - ...

SAS custom APIs: SAS developed 13 extra custom APIs/Macros for GSK

- Purpose: provide additional functionalities not available with the native APIs.
- Usage:
 - Permanent Deletion of an object from the recycle bin
 - Permanently delete an object from the repository (not going thru the recycle bin)
 - Get all groups of a SDD user
 - Get the list of scheduled jobs
 - Submit or schedule an existing job
 - ...

GSK custom APIs: set of Java APIs and macros developed by GSK

- Purpose: automation of tasks using Java functionalities and native SDD APIs.
- Usage:
 - Compute the checksum of a file
 - Tar, Zip & Gzip utilities (incl. unTar, unZip & unGzip).
 - Transfer a file from/to the repository to/from the workspace (download/upload)
 - Query the audit trail (header and details).
 - Update the attributes of a user
 - ...

Interoperability, a Real-Life Case

Overview



- Life Science Analytics Framework (Recap)
- Overview of SDD/LSAF at GSK
- The APIs
 - The need for APIs
 - SAS Native APIs
 - SAS Custom APIs
 - GSK Custom APIs
- **Use cases**
 - User management
 - Object management
 - Synchronization of data archives
- SAS future (Viya)

SDD/LSAF Interoperability

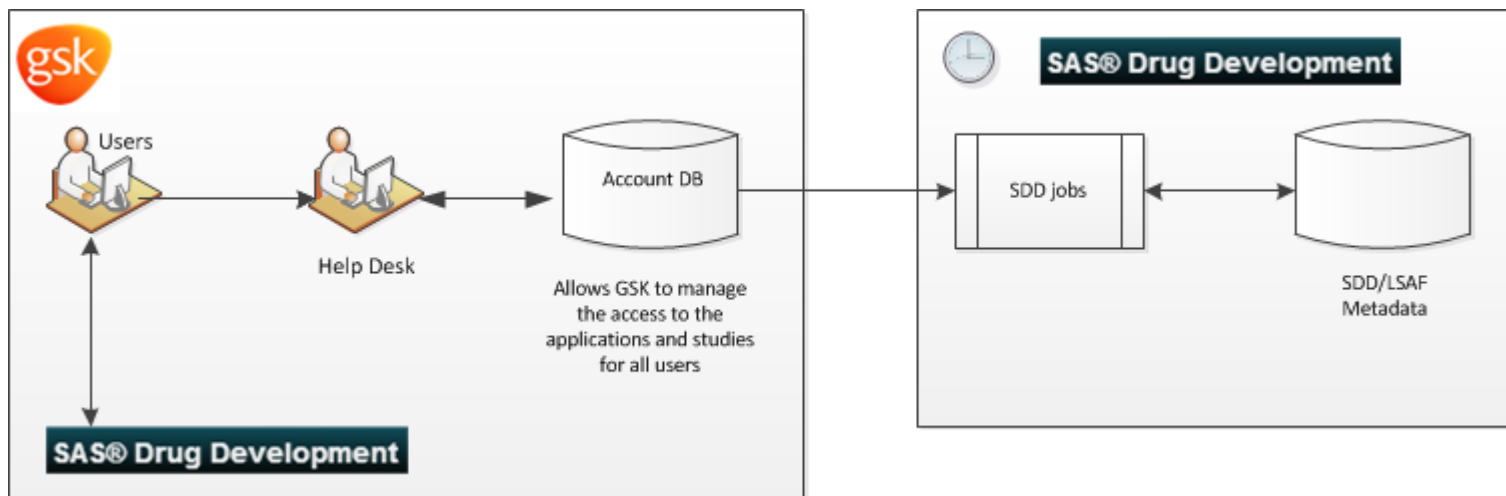
Use cases: user management



User Management: uses SAS Native and Custom APIs/macros.

Process:

- Access to GSK application and studies are managed in Account (GSK App. with Oracle DB back-end).
- Every 20 minutes an SDD job queries the Account DB for changes and applies the changes on SDD:
 - creation of new users and update of their attributes (first name, last name mail address,...),
 - de-activation, re-activation, un-lock,
 - reset of password,
 - grant/revoke access to projects and studies objects (assigning uses to access groups).



SDD/LSAF Interoperability

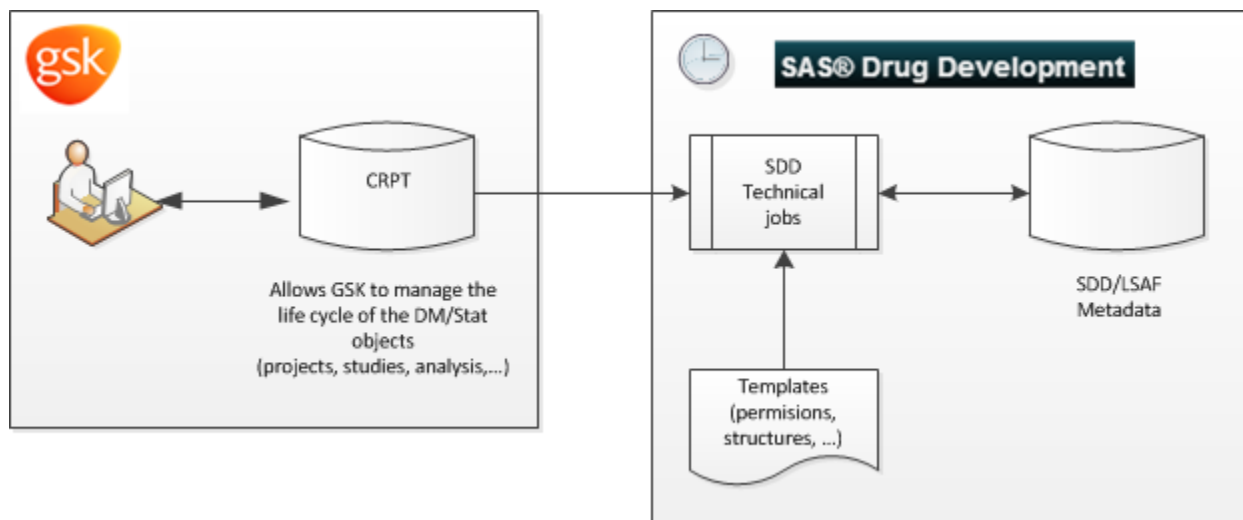
Use cases: object management



Object Management: use SAS Native APIs/macros.

Process:

- Life cycle of DM and Stat objects is managed in CRPT (GSK App. with an Oracle DB back-end).
- Project, studies, study analysis, project analysis are created, cancelled, completed, locked in the App.
- A set of SDD Jobs are scheduled to monitor the DB for changes and apply the changes on SDD:
 - creation of new files and folders, of access control groups, assignation of access control groups to objects,
 - templates describe the files and folder structures, access group and permissions.



SDD/LSAF Interoperability

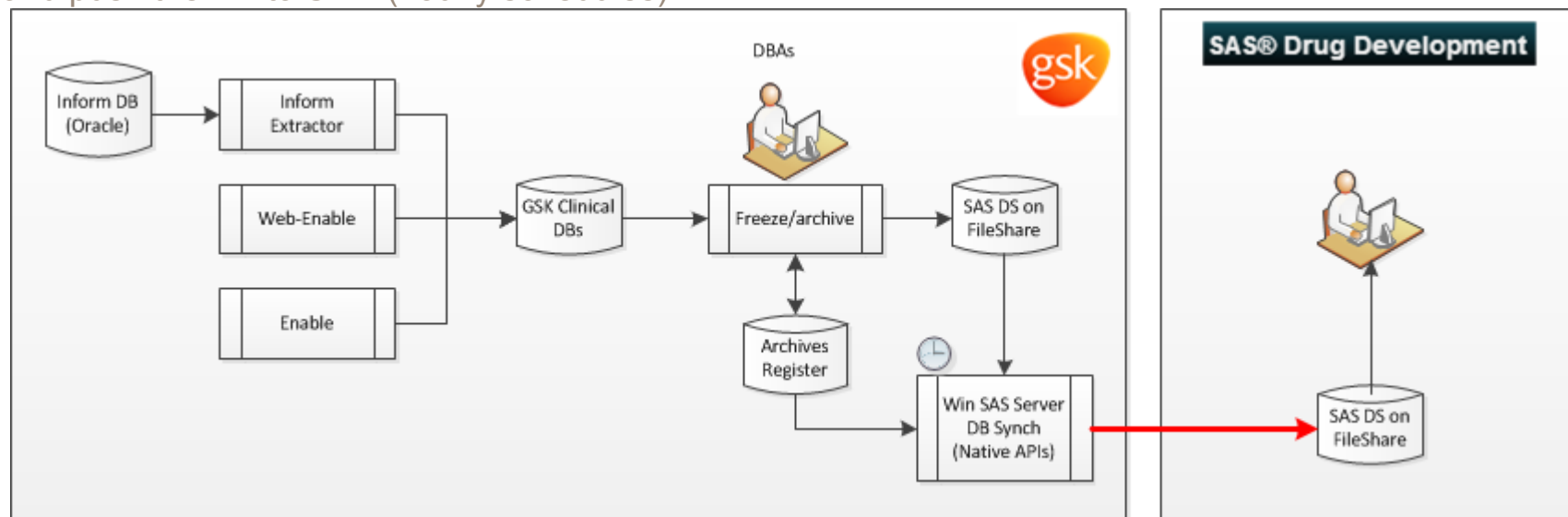


Use cases: synchronisation of data archives

Archives Synchronisation: use SAS Native APIs/macros.

Process:

- 3 EDC applications are used to collect CRF data: enable, Web-enable and Inform.
- During the course of a study, data are centralised in a live operational DB (one schema per study)
- At times, data snapshots must be taken (e.g. interim analysis, final analysis) and the process occurs on the SAS Windows platform (due to technical constraints and for legacy compatibility).
- Batches scheduled on the Windows SAS Server scan the Windows file shares for new data snapshots and push them into SDD (hourly schedules).



Interoperability, a Real-Life Case

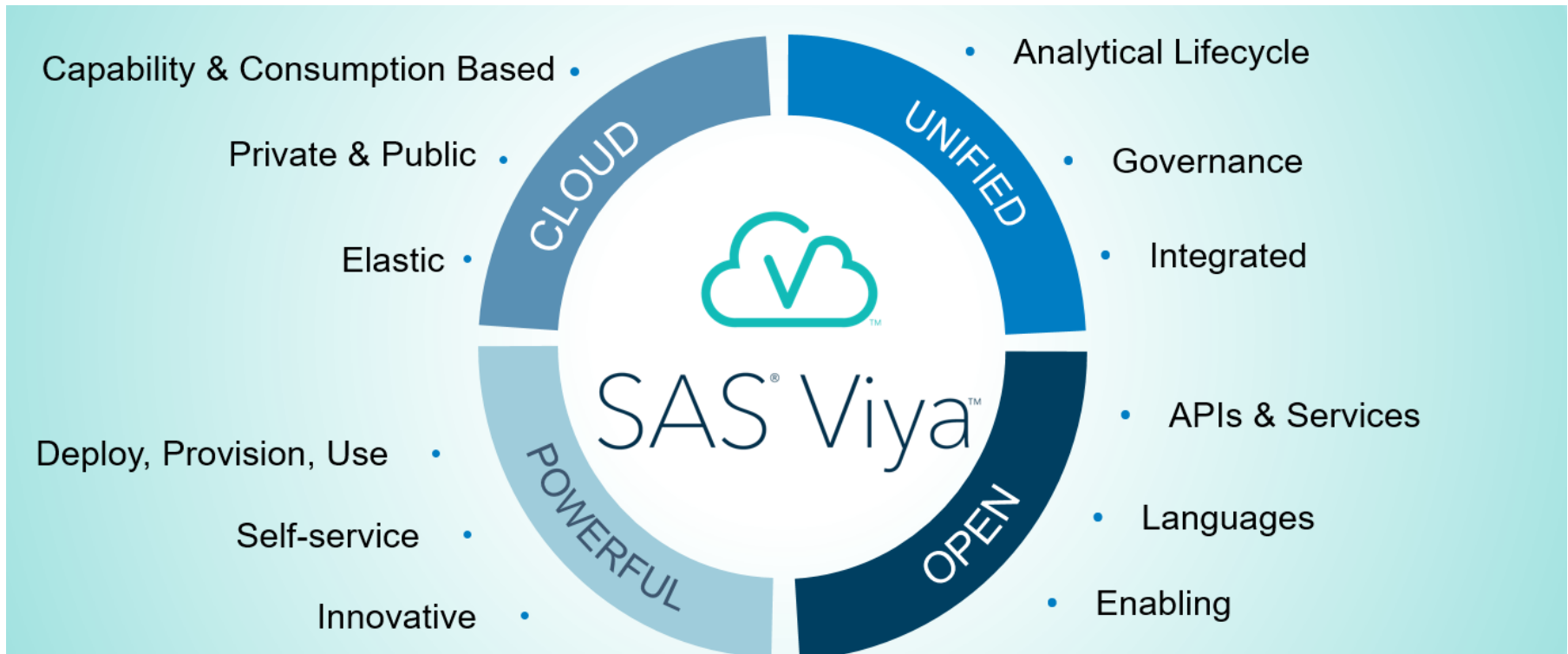
Overview



- Life Science Analytics Framework (Recap)
- Overview of SDD/LSAF at GSK
- The APIs
 - The need for APIs
 - SAS Native APIs
 - SAS Custom APIs
 - GSK Custom APIs
- Use cases
 - User management
 - Object management
 - Synchronization of data archives
- **SAS future (Viya)**

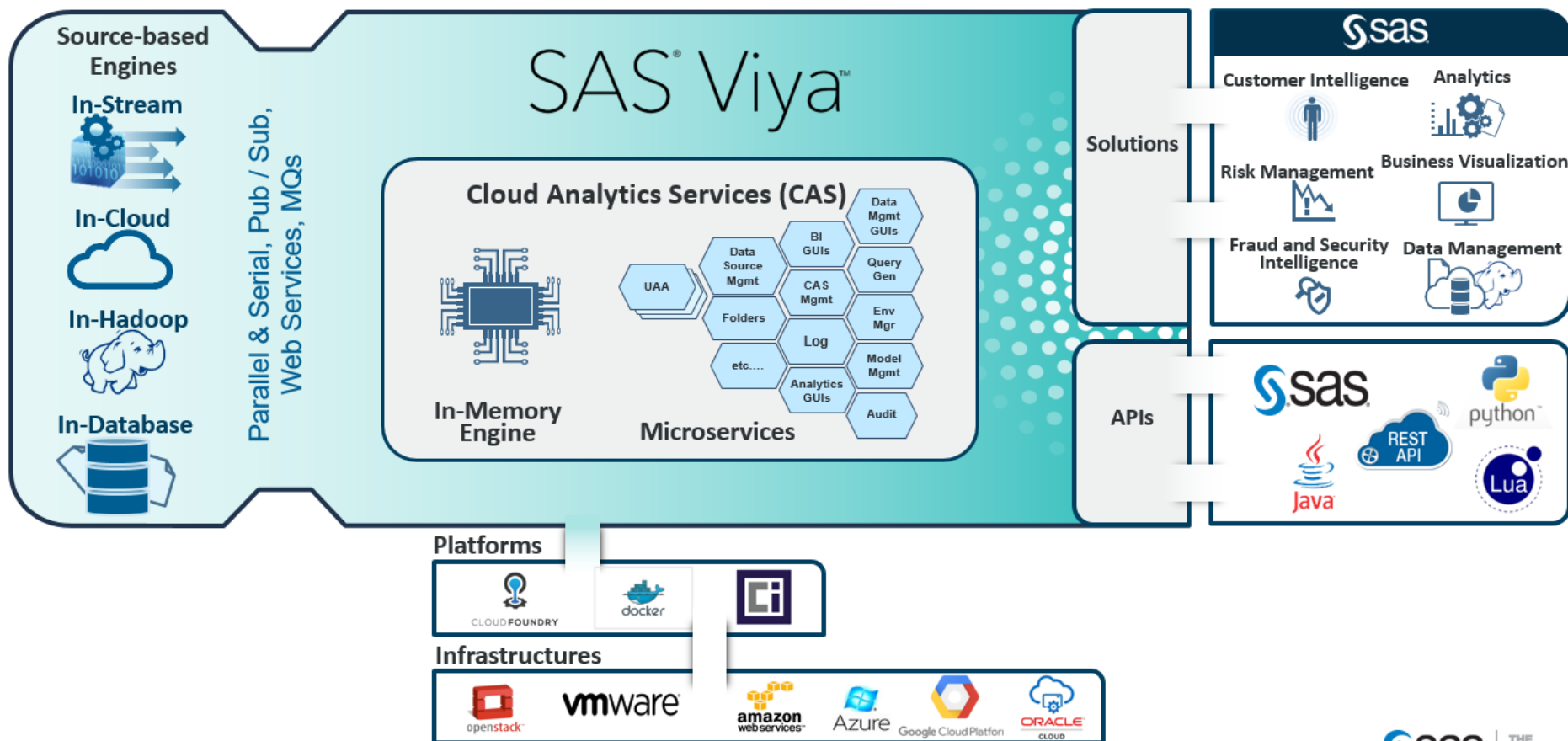
Interoperability, a Real-Life Case

Future - SAS® Viya



Interoperability, a Real-Life Case

Future - SAS® Viya



Thank you !