



Implementation of ChemAxon in a SOA environment

Shane Weaver

Agenda

GSK R&D IT have, over the past 10 years, worked to establish a Services Oriented Architecture environment. Various chemistry toolkits and data cartridges underpinned our chemistry web services which support key chemistry processes. Recent changes to these services to rationalise vendors have been the biggest we have ever experienced. SOA allowed us to proceed with minimal impact to the business.

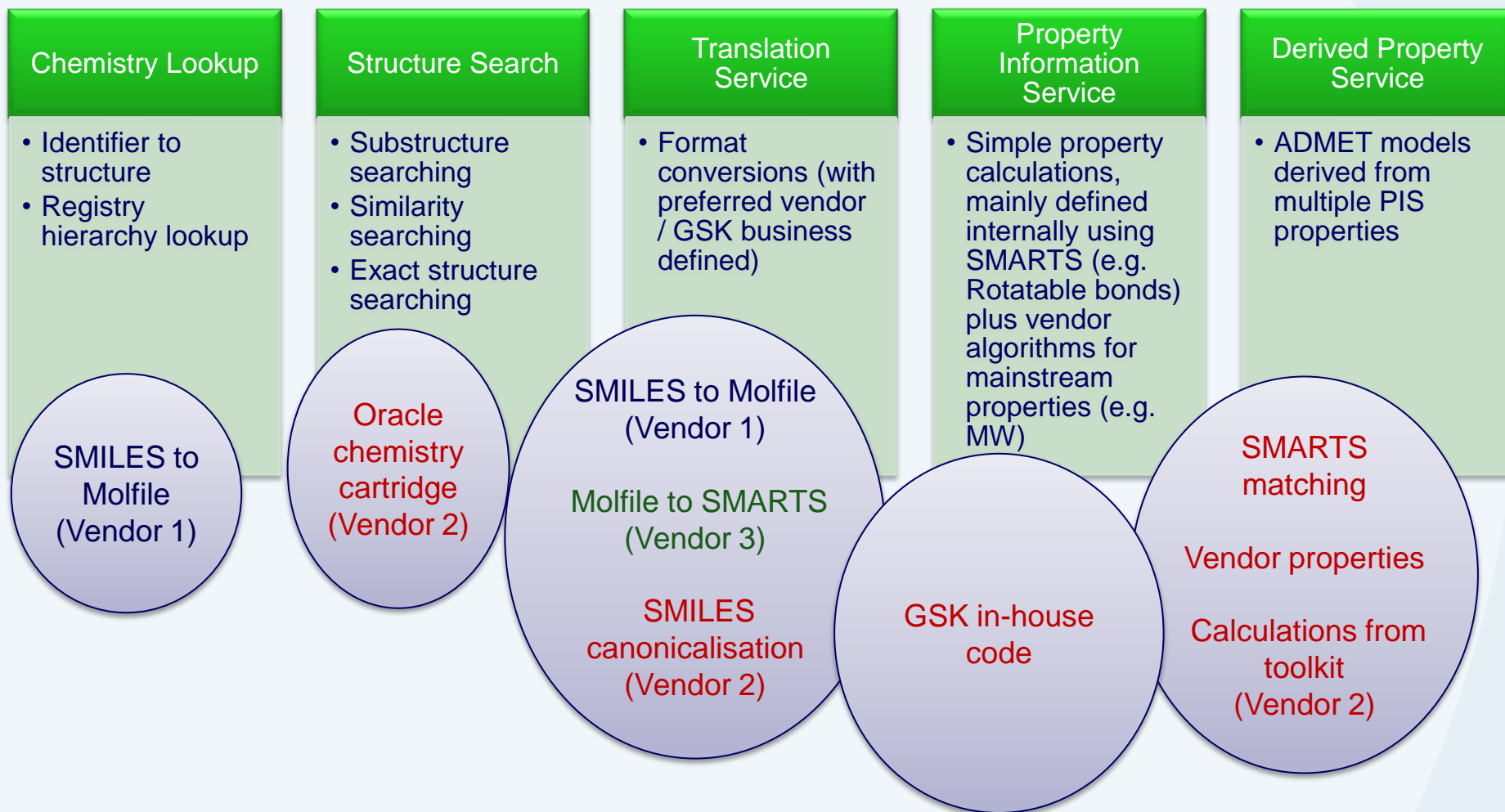
- We will discuss *why*
 - Simplification
 - Cost
- *..and how*
 - Migration Plan
 - Minimal Impact
 - Business Needs

Why...

IT Simplification Drivers – Reduce TCO

- Reduce number of different technologies used in the environment
 - .NET
 - Desktop thick clients
 - Web applications
 - Oracle
 - Databases
 - Java
 - TBD
 - C++
 - Phased out
- Reduce software architecture components
 - Analyse, Migrate, Retire
- Reduce available vendor software
 - Consolidate and reduce license and support costs

Web Services at GSK *as it was*



Why we originally choose these vendors

- Vendor 1
 - Generates good molfile coordinates from SMILES
- Vendor 2
 - Commercial chemistry toolkit and Oracle cartridge integration
- Vendor 3
 - Generates best quality molfile to SMARTS conversions
- GSK in-house
 - Algorithms not supported natively by vendors

Consolidating to ChemAxon as a key Vendor

- Not trying to change the world – implementing most common chemistry features in a single package using current technologies
 - ChemAxon provide both a Toolkit and Cartridge product
 - ChemAxon products provide opportunities for further consolidation
 - such as IJC and JC4XL which will now be used to replace further vendor software
 - Has support for our currently preferred developer languages (Java & .NET) plus support for SMILES/SMARTS with close enough adherence to the standard
 - Continuous development enabling product improvements
 - Contains as standard items custom written by GSK which could be replaced
 - Must have good performance – User perception of performance key

Web Services at GSK to be

Chemistry Lookup

- Identifier to structure
- Registry hierarchy lookup

Structure Search

- Substructure searching
- Similarity searching
- Exact structure searching

Translation Service

- Format conversions (with preferred vendor / GSK business defined)

Property Information Service

- Simple property calculations, mainly defined internally using SMARTS (e.g. Rotatable bonds) plus vendor algorithms for mainstream properties (e.g. MW)

Derived Property Service

- ADMET models derived from multiple PIS properties

SMILES to Molfile

Molfile to SMARTS

SMILES canonicalisation

SMARTS matching

Vendor properties

Calculations from toolkit

Oracle chemistry cartridge

(ChemAxon)

GSK in-house code

How...

General migration plan

Determine vendor usage



Revoke unused license



Migrate legacy applications to use forward looking SOA



Retire obsolete services/software



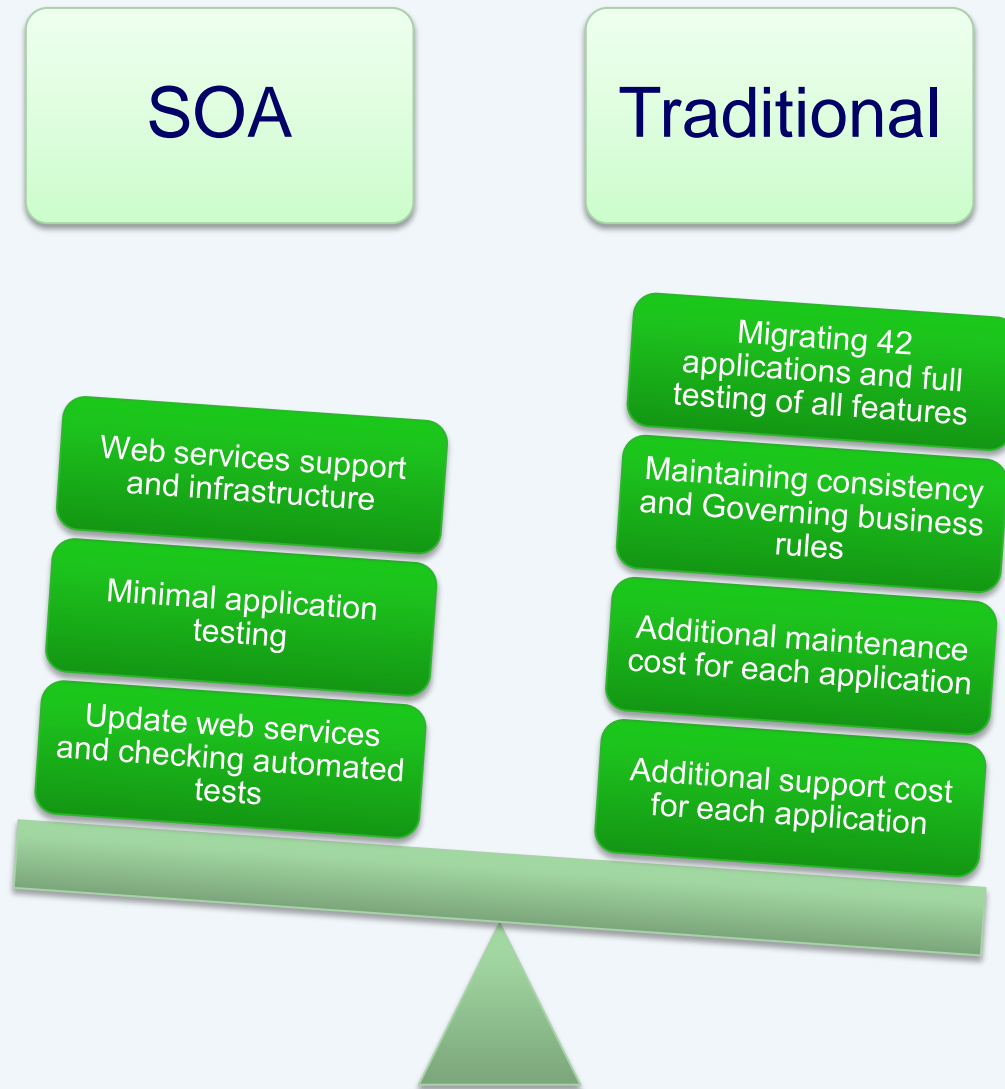
Migrate services and database to ChemAxon toolkit & cartridge

Objectives for minimal impact to business

- Changes transparent for client applications
- Ensure existing business rules are maintained in services
- Zero downtime

SOA can benefit in the delivery of these objectives

SOA saves on Total Cost of Ownership



- After 3 years: 65% more
- After 5 years: 82% more
- After 10 years: 110% more

Case Study – Rendering structures

The Problem

- Web apps require a method to display a rendered image of structures
 - Chemistry toolkit on every desktop is slow and difficult to support/maintain

The Solution

- Use SOA
 - SOAP/HTTP APIs
 - Fast
 - Load balanced
 - Easy to update
 - Simple to use
 - E.g. Via HTML img tag:

```

```

Structure Picture Service

Use case – dynamic alignment

- A dynamic HTML/Javascript page rendering images, requires on-the-fly structure alignment
 - Previous mechanism used background AJAX server script running alignments sequentially
 - Aligning 10 structures ~26s
 - Aligning 100 structures ~200s
- Solution: Implement alignment directly in SPS by passing both target and alignment structure to service as parameters
 - Alignment algorithms already built-in to ChemAxon toolkit stack
 - Simpler client code/architecture, no need for AJAX script
 - Better use of load balancing, so much faster
 - Aligning 10 structures ~1s
 - Aligning 100 structures ~7s

Summary

- **Why** we migrated to ChemAxon
- **How** we migrated to ChemAxon
- How our implementation of a SOA environment has lead to
 - One place to house and easily change business rules
 - Reduced Total Cost of Ownership
- **What next?**
 - Instant JChem
 - JChem for Excel



GlaxoSmithKline