



Migrating chemical information

2007

Agenda

- Business Drivers for change
- Architecture and technology
- Approaching migration
- Design considerations
- Migration use cases
- Migration methods
- Conclusion

Migrating chemical information to new architectures...

© The Edge Software
Consultancy Ltd

Business Drivers

- Costs
 - Operational
 - Capital expenditure
 - Staff and support
 - Deployment
 - Buy vs Build
- Licensing models
- **Value for money**



Things need not be as they have always been...

© The Edge Software
Consultancy Ltd

Business Drivers

- Support
 - Issue tracking vs help
 - Documentation access
 - Response time vs resolution time
 - Accessing bug fixes
 - Release cycles



When things go wrong...



© The Edge Software
Consultancy Ltd

Business Drivers

- Architecture and technology
 - Proprietary software
 - Dependency
 - Risk balancing
 - Mainstream technologies
 - Advantages
 - Knowledge base
 - Accessibility of training and staff

Keeping up with new approaches...



© The Edge Software
Consultancy Ltd

Technology advances

- Oracle Data Cartridges
 - Encapsulates chemistry data types
 - Integration in the database tier
 - “Standardised” interface
- “Standard” Business objects
 - Applets for visualisation and editing
 - Business objects for specialist types

New ways of treating old problems...

© The Edge Software
Consultancy Ltd

Approaching migration

- Business analysis
 - Base line assessment survey
 - Understand the current information landscape
 - Define integration requirements
 - Identify information source differences
 - Formulate system map and requirements
 - Define information and work flow

How to give an old dog some new tricks...



© The Edge Software
Consultancy Ltd

Chemical Information Systems

- Content databases
 - Synthetic methods, reagents, compounds etc
- Reagent Inventory
- ELN
 - Library design, Synthesis planning, analytics
- Chemical Registration
- Compound Inventory
- Virtual compounds

More than just compound registration...

© The Edge Software
Consultancy Ltd

Selecting a partner

- Evaluation
 - Vendor capability (size and development and test team)
 - Synergy with internal staff
 - Chemical representation
 - Performance
 - Business relationship
- Set criteria for evaluation
- RFPs
 - Remember the answer is always yes!

You may not be able to pick your family ...



© The Edge Software
Consultancy Ltd

Partitioning

- Separate the application tiers
- Define and maintain interface boundaries
- Add chemistry to standard application architectures and technology
 - Not build standard applications from chemistry specific technology...

Design to make next time easier...

© The Edge Software
Consultancy Ltd

Migrating chemical structures

- Format conversion
- Chemical representation
- Business logic
 - Business rules
 - Processing

The challenges...



© The Edge Software
Consultancy Ltd

Format considerations

- Stay agnostic on format
- Keep source format
 - Future proofing
- Be aware of information loss
 - Stereochemistry
 - Coordinates

Translating apples and pears to fruit ...



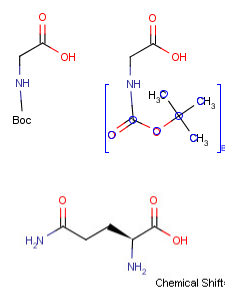
© The Edge Software
Consultancy Ltd

Chemical representation

- Stereochemistry
- Molecular mass!
- Molecular formula
- Aromaticity
- Explicit Hydrogens
- Abbreviated groups
- Data Sgroups

English or Franglais?...

© The Edge Software
Consultancy Ltd



Business logic

- Chemical business rules
 - Salt stripping
 - Charge neutralisation
 - Standardisation
- Duplicate checking
- Search logic
- Enumeration

No surprises...

© The Edge Software
Consultancy Ltd



Migration methods

- In database
 - ETL
- Export and import
 - Pipelining PP, Knime, etc
- Siphoning
 - Identify use cases
 - Implement business rule
 - Test
 - Migrate

Moving the mountain rock by rock...



© The Edge Software
Consultancy Ltd

Conclusion

- Migrating large databases is
 - Feasible
 - Tractable
- Tools and help is available
- Chemaxon components facilitate use of new architectures
- Migration can deliver real benefits

andrew@edgesoftwareconsultancy.com
<http://www.edgesoftwareconsultancy.com/>



© The Edge Software
Consultancy Ltd