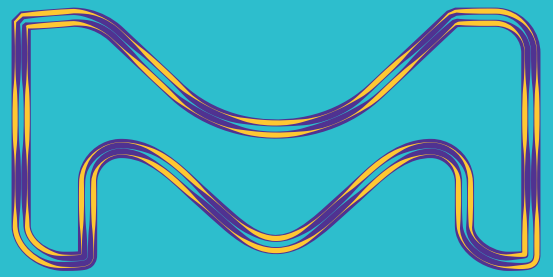


chemAxon@Merck

Migration to ChemAxon Technology
Our way into a new world

Elke Hofmann, Michael Hofmann
Research Informatics, Merck

ChemAxon User Meeting Budapest 2017

MERCK

Agenda

we are MERCK

We live in a world of possibilities.

A world where exploration and discovery are celebrated. Our meticulous and research-driven businesses deliver diverse, high-quality products that enrich lives and enable us to share business success with our customers.

Founded in Darmstadt, Germany, in **1668** by Friedrich Jacob Merck, we are the world's oldest pharmaceutical and chemical company. Today, the Merck family remains the majority owner of the company.

Over the course of nearly 350 years, we have become a truly global company. Our approximately 50,000 people work in 66 countries and are united by their passion for new ideas, the possibilities of technology, and the potential to make a difference in the world.

We are known as Merck internationally. In the United States and Canada we operate as EMD Serono in the Biopharma business, as MilliporeSigma in the Life Science business, and as EMD Performance Materials in the materials business.

What we do

- **Prescription drugs** and solutions to treat cancer, multiple sclerosis, infertility, cardiovascular and metabolic diseases
- **Over-the-counter** products for a healthy lifestyle
- **Allergy** products and **biosimilars**

- Innovative **tools** and **laboratory supplies** for the life science industry that make **research** and **biotech** production better, faster and safer
- **Broad and in-depth portfolio** of 300,000 products
- **Industry leading e-commerce platform**, SigmaAldrich.com
- Award-winning **innovation**

- A wide range of **high-tech chemicals**, such as:
  - **Liquid crystals** and **OLED** materials for displays & lighting
  - **Effect pigments** for coatings and cosmetic products
  - **Specialty chemicals** for the semiconductor industry
  - **Functional materials** for solar panels

Research Informatics
Who we are and what we do

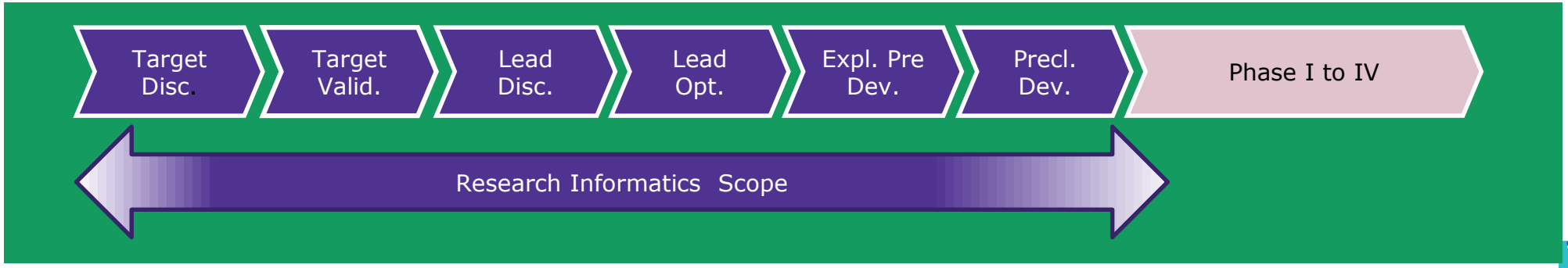
Research Informatics is part of the R&D Organization within Healthcare > Prescription Drugs

High-level RI objectives

- Enable data driven decision making by providing state-of-the-art research knowledge management
- Manage research data from internal & external sources
  - data capture
  - processing and analysis
  - integration
  - visualization

Our Mission:

- Enable project teams and R&D activities by providing state of the art scientific informatics solutions and bioinformatics project contributions



Migration to ChemAxon technology – Our way into a new world
Objectives

- **Consolidation** of small molecule software landscape
  - Too many systems with partly overlapping functionality

    -> reduction of systems and number of system providers
  - Concentration of resources and expertise
- **Future-Orientation**: Implementation of technology with high innovation potential and good support
  - Large molecule and ADC representation via HELM standard
  - Enrichment of Sharepoint with chemical intelligence
  - Advanced science and technologies (chemical text mining, macrocycle naming, Markush structure, …)
- **Scope of the project**
  - Database cartridge for chemistry representation
  - Drawing tool for chemical structures
  - Office Add-Ins for structure visualization and manipulation

Migration to ChemAxon technology

Vendor selection – how will our new world look like?

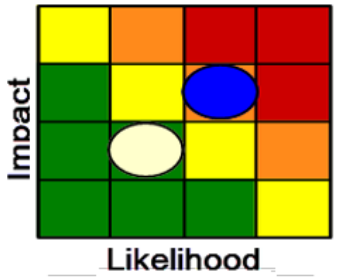
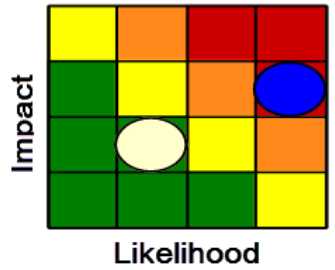
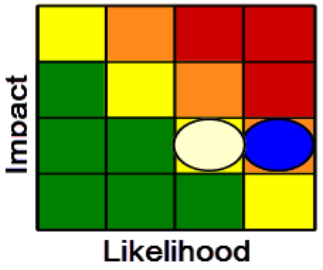
- Market assessment to identify suitable candidates
- Request for Proposal (RfP)
  - Several vendors were contacted and asked for detailed information about their systems

- Decision to proceed with ChemAxon based on
  - Better coverage of functionality
  - Support option during migration

- Before the final decision
  - In-house installation of ChemAxon technology to evaluate potential risks
  - Extended pilot with Marvin and JChem for Office

Migration to ChemAxon technology

Risk assessment and mitigation plan

Risk	Mitigation	Risk Matrix Blue: Risk w/o mitigation White: Risk after mitigation
Different chemical representation of the cartridge e.g. stereochemistry	• In-house evaluation installation: detailed assessment with our research structures • Training • If required, adjustment of registration guideline	
Change of drawing tool different and maybe reduced functionality, quality aspects	• In-House assessment with key users • Training • If required, agreement to amend drawing tool in close collaboration with vendor	
Chemical structures in legacy documents (several thousands) Can structures added with old drawing tool be modified with new drawing tool? Conversion possible?	• In-House assessment with key users • Conversion tool mandatory	

Migration to ChemAxon technology
Risk evaluation and mitigation based on in-house installation

Chemical Representation

- Analysis of differences between old and ChemAxon cartridge
  - Testcase: Research Database with 1.5 M chemical structures
  - Potential pain points:
    - Check on duplicates, isomers, tautomers during NCE registration
    - Workflows including structure searches

Duplicate and isomer recognition very good, no further action required

Legacy Documents

- Thousands of legacy documents with OLE objects of old drawing tool
- Conversion to Marvin OLE on-the-fly possible, but room for improvement regarding stability & speed
- OLE only supported with Java 1.6
- Strategy
  - Active documents to be converted
  - Non-active documents can be kept as is

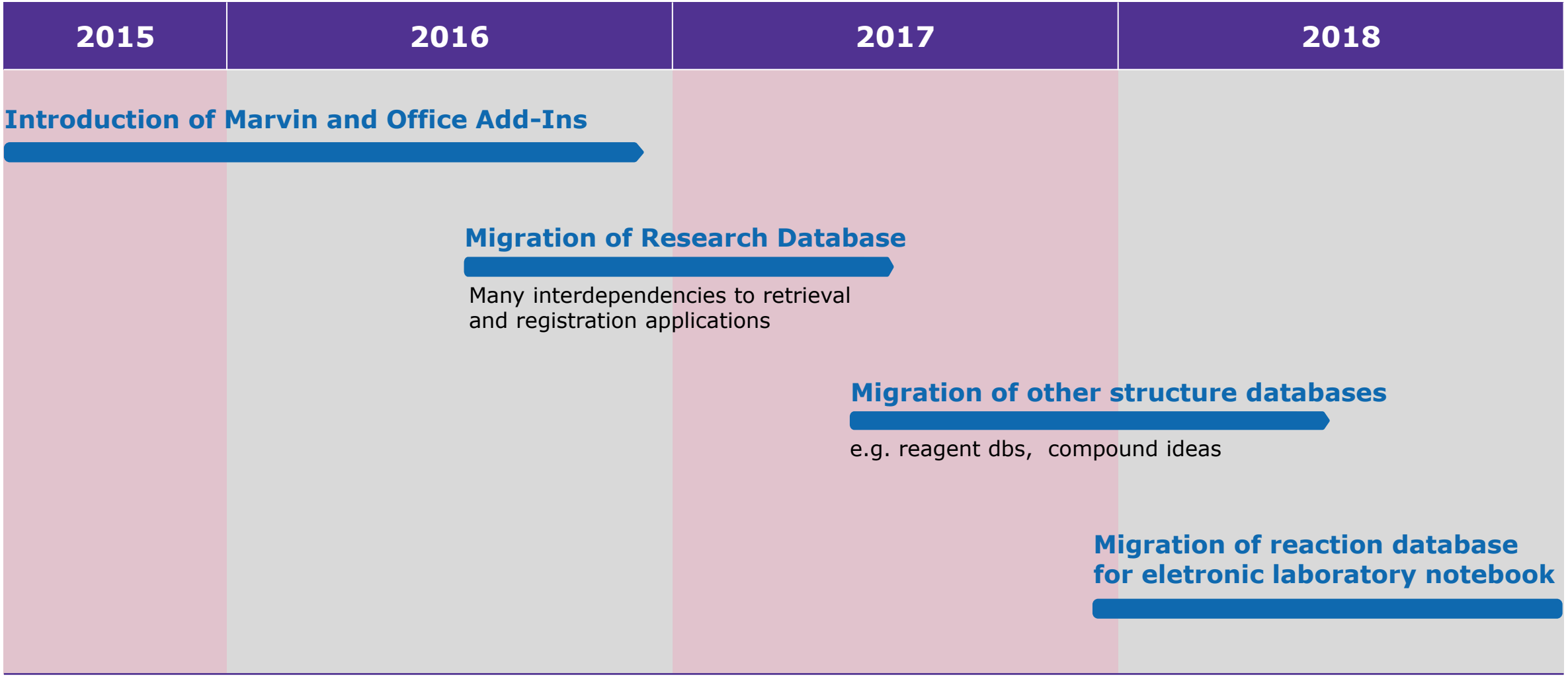
Migration to ChemAxon Technology
Risk evaluation and mitigation based on in-house installation

Drawing Tool

- Main issues with V14
  - Configurable settings like standard bond lengths
  - Journal settings
  - Publication-ready quality of drawn structures
  - Handling of no-structures
  - Complex reaction schemata
- Close collaboration with ChemAxon to mitigate most of the issues
- Marvin integrated as „working horse" in our applications

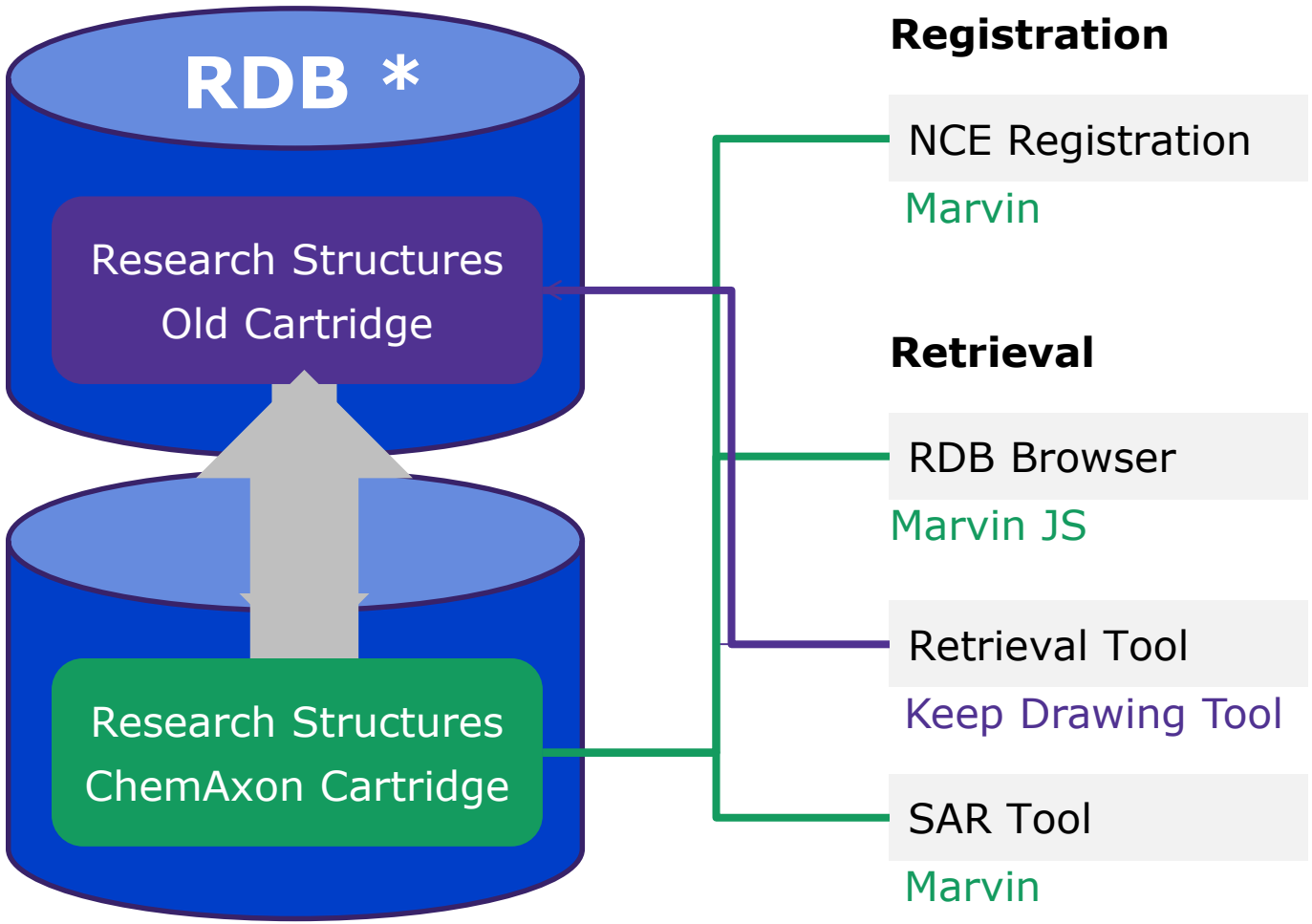
- However there is still the need for a drawing tool with support for ambitious publication sketches (more graphical elements, more design elements…)

Migration to ChemAxon technology
High level implementation plan – 4 main phases



Migration to ChemAxon technology
Details of implementation plan for RDB* – stepwise approach

1. Set-up secondary structure database in ChemAxon cartridge format and synchronize
2. Integrate Marvin / Marvin JS into retrieval and registration applications
3. Migrate retrieval applications to use ChemAxon cartridge
4. Migrate registration applications to use ChemAxon cartridge
5. Use ChemAxon cartridge as primary structure cartridge
6. Retrieval tool will be replaced. In the meantime old set-up will be kept and synchronized



*** RDB: Research Database**

Controlled Substance Compliance

What is a controlled substance?

- Controlled substances are subject of legislative control either on local, national or international level
- There are restrictions about the production, import/export, cross-border shipment, supply, use and possession

1 Selection of vendor and tool

- Three vendors contacted
- Test case (SDFile) provided
- Decision for Compliance Checker from ChemAxon
  - Best fit to our landscape
  - Good experience with support/collaboration with ChemAxon

2 Is our research substance pool in compliance with applicable regulations?

- Our research substances are managed in a central dispensary and shipped between the two major research sites in Germany and the US
- It is mandatory to be able to ship cross-border
- **Regularly check research substance pool to cover new legislations**
- **Make shipment restrictions available to compound logistics system**
- **Integrate in NCE registration as background process**

3 What about our idea generation?

- Pro-active checks to prevent that we spent time to design compounds which are not compliant
- **Integration in compound design workflow**

Where we are today
Cartridge & JChem web services

Cartridge

- Chemical functions are good
- Integration of chemical calculations directly in cartridge functions (e.g. use of chemical terms) opens new possibilities
- Performance under heavy load not yet tested
- Separate middleware & need to give password of user to middleware (for re-connection to Oracle) requires security concept with non-personalized access

JChem web services

- To retrieve chemical structures or additional data based on compound ID from within Office products (e.g. Excel) is like bread and butter for the chemist
- At Merck, direct client/server connection is unwanted and restricted
- Idea was to use JChem web services not only for web applications, but also as data source for JChem for Office (JC4O)
- Unfortunately, JC4O did not support this
- Also JChem web services did not support JChem cartridge
- ChemAxon enhanced both products to support this scenario. As of today, functionality is very good, performance fine

Where we are today
JChem for Office (JC40)

JC4O influences standard behavior of Office so that Office is not working as expected

- ChemAxon aims to ease users life with comfortable automatisms, e.g. interpreting content of clipboard
- Especially in copy/paste scenarios, this leads to unexpected and unwanted behavior, even when no chemical structure is involved
- Unwanted behavior can be minimized by disabling “take over control” of Ctrl+C/V, but then some functions will not work

Suggestions

- **Do not take over any Office functions**
  Office has to work with enabled ChemAxon Add-Ins as well as without
- **Make a clear difference**
  On-demand usage of ChemAxon functions if structures are involved
- **No more automatisms**
  Expose all functionality to the GUI…
  e.g. let the Chemist decide if he wants to insert a MOL or a chemical name in the clipboard as structure or as text
- **Native .NET application for office add-ins**

Where we are today
Marvin

Non-standard behavior

- "Arrow up" moves selected structure down
- Selecting and deleting a bond deletes the atoms as well

Integration of Marvin in DotNet world, OLE and rendering quality

Marvin [Java] → IKVM [Java/.NET] → Marvin.NET → custom application

- 1st challenge: version independency (new version of Marvin must not trigger rollout of all custom applications using Marvin!)
  –> solved by implementing wrapper DLL between Marvin.NET and custom application
- 2nd challenge: unexpected problems somewhere in the chain above; e.g. Ctrl+C in Marvin called from custom application lets custom application crash
  –> search for solution ongoing
- OLE in office Add-ins: provide same rendering quality in drawing tool and office

Standard, consistent behavior makes life easier
- **reduces learning effort**
- **increases acceptance**

Where we are today
ChemAxon's release cycles

ChemAxon releases new versions weekly

Each version contains both: bug fixes and new features which may introduce new bugs

General roll-out policy at Merck

- Stable software: maximum biannually
- Software to be stabilized: maximum 4 new versions per year
- Exceptions: severe bugs which have to be fixed

Internal workflows and resources

- SW packaging and testing is time-consuming, can require up to 3 months
- Limited resources to test new versions
- Users need stability – ever-changing functionality is confusing

What we wish for:

- **One, max two releases per year (further development)**
- **Bug fixes based on this release**

Where we are today
Summary

Chemistry cartridge

- Research Database migrated
- Next steps – migration of:
  - Several smaller structure databases
  - Reaction database for electronic laboratory notebook

Drawing tool, web services, Office add-in

- Marvin integrated in major applications
- Marvin JS & web services integrated into some small web apps
- JC4O still in beta test

We made it
Let us explore the new land

Compliance checker

- Check of research substance pool and integration of results in dispensary inplemented
- Next step: Integration in NCE registration and compound design workflow

Next possible steps?

- Chemical data mining in documents?
- Sharepoint ??
- Marvin Live?
- NBE and ADC?

Acknowledgements

Without the support and engagement of many our scientists and ChemAxon the migration couldn’t have been realized.

Merck KGaA

- Key users from different functions
  - Hans-Michael Eggenweiler
  - Hans-Peter Buchstaller
  - Dieter Dorsch
  - Lars Burgdorf
  - Claudio-Wilfried Ritzert
  - Claudia Blank
- Research Informatics
  - Zhaowen Luo
  - Christophe Betton
  - Achim Herrmann
  - Veit Ulshoefer

… and many more!

ChemAxon

- Akos Papp
- Akos Tarcsay
- Efi Hoffman
- János Fejérvári
- Jozsef David
- Balázs Zaicsek
- Miklos J. Szabo
- Roland Lohner
- Daniel Bonniot

… and many more!