

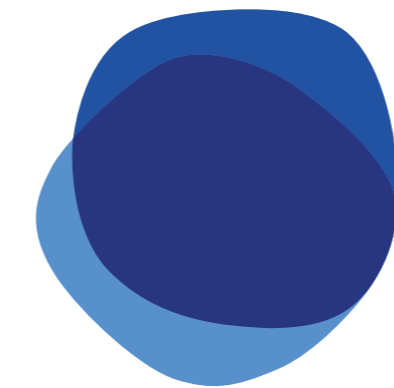
# CombineArchiveToolkit

## facilitating the transfer of research results

Universität  
Rostock



Traditio et Innovatio



SYSTEMS BIOLOGY  
BIOINFORMATICS  
ROSTOCK

SEMS  
simulation experiment management system

### The COMBINE Idea

The 'COmputational Modeling in Bology' Network (COMBINE) is an initiative to coordinate the development of the various community standards and formats for computational models (BioPax, SBGN, SBML, SED-ML, etc.) [1]. One of the major goals of COMBINE is to improve the interoperability of these standards, and to support fledging efforts aimed at filling gaps or new needs.

The steadily increasing size and complexity of models and derived data poses the challenge of sharing reproducible results. Today, these results typically consist of multiple model files, simulation descriptions, publications, and meta data. The question how to provide all relevant files and modelling results, in a reliable and reproducible manner, remains.

In 2011 the COMBINE community [2] proposed the COMBINE archive format [3] which is a container that bundles all files related to a project into a single file. Typically, it comprises the model files needed to run a particular set of experiments. In addition, it contains all associated files that are needed to reproduce the experiments such as simulation experiment descriptions (SED-ML), semantic annotations, or graphical representations in SBGN-ML. All files can be equipped with meta-information such as people attributions and details about the files inside the archive. Generally, a COMBINE archive is encoded using the Open Modelling EXchange format (OMEX).

### What's the gap?

Manually handling COMBINE archives is tedious and error prone. Consequently, computational support is needed to undertake this task. Only then, it will become possible to exchange COMBINE archives seamlessly between different applications and repositories. Such a tool is constrained to provide mechanisms to create, explore and modify files and meta information in a COMBINE archive.

### Our Approach

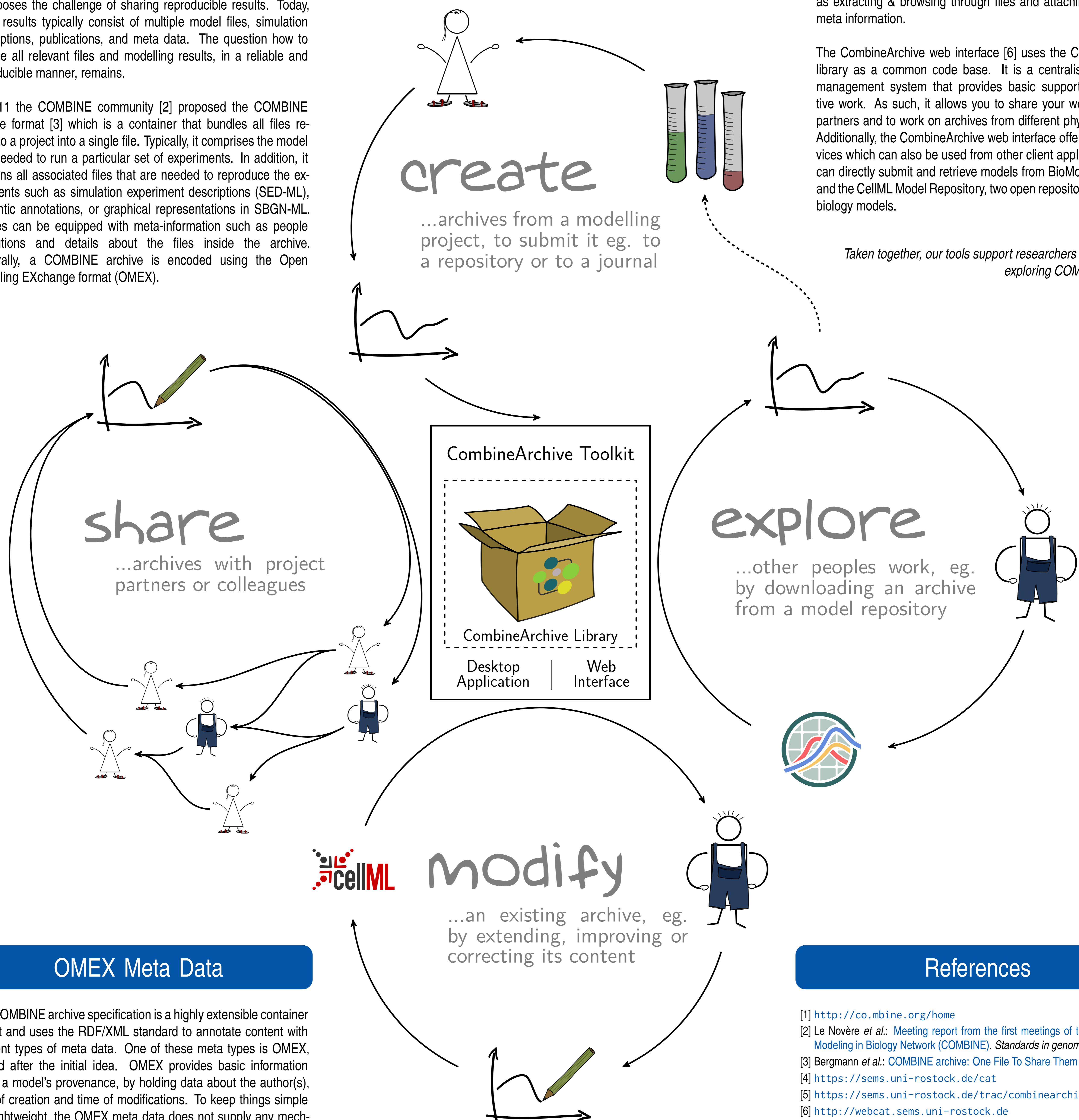
To provide the needed computational support, we developed the CombineArchiveToolkit [4]. It consists of a core library, a desktop application, and a web based interface. The CombineArchiveLibrary [5] was implemented using latest Java technologies.



It offers all necessary methods to handle COMBINE archives, such as extracting & browsing through files and attaching & retrieving meta information.

The CombineArchive web interface [6] uses the CombineArchive library as a common code base. It is a centralised cloud data management system that provides basic support for collaborative work. As such, it allows you to share your workspaces with partners and to work on archives from different physical locations. Additionally, the CombineArchive web interface offers RESTful services which can also be used from other client applications. Users can directly submit and retrieve models from BioModels Database and the CellML Model Repository, two open repositories of systems biology models.

Taken together, our tools support researchers in creating and exploring COMBINE archives.



### OMEX Meta Data

The COMBINE archive specification is a highly extensible container format and uses the RDF/XML standard to annotate content with different types of meta data. One of these meta types is OMEX, named after the initial idea. OMEX provides basic information about a model's provenance, by holding data about the author(s), time of creation and time of modifications. To keep things simple and lightweight, the OMEX meta data does not supply any mechanism for version control, although this can be easily archived by using a version control system. [7]

### References

- [1] <http://co.mbine.org/home>
- [2] Le Novère *et al.*: Meeting report from the first meetings of the Computational Modeling in Biology Network (COMBINE). *Standards in genomic sciences*, 2011.
- [3] Bergmann *et al.*: COMBINE archive: One File To Share Them All. *arXiv*, 2014.
- [4] <https://sems.uni-rostock.de/cat>
- [5] <https://sems.uni-rostock.de/trac/combinearchive>
- [6] <http://webcat.sems.uni-rostock.de>
- [7] Waltemath *et al.*: *Bioinformatics*, 2013.



Martin Scharm, Florian Wendland,  
Martin Peters, Markus Wolfien,  
Tom Theile, Dagmar Waltmath  
<http://www.sbi.uni-rostock.de>

