The Add-In (.NET) links a semantic chemistry engine (.NUMBO) through a command interface (CID) to a chemistry zone. Chemistry zones are textual or graphic renderings within a Word document (DOCX). All content and relationships are bound to CML in the DOCX package.

The Chemistry Add-In for Word, is an open source program that allows chemists to create, edit and manipulate chemistry (labels and 2D structures) in the Word environment. The on screen representation is backed by semantic data in Chemical Markup Language (CML). Combined with domain aware libraries we enable novel functionality in data checking during the authoring process, chemistry-centric article reading support and data-mining applications.

University of Cambridge
Dr Joe Townsend
Collaborative R&D between the Unilever Centre for Molecular Science Informatics, University of Cambridge and Microsoft Research, Redmond.

Semantically Editing Chemistry in Word
People (UoC) Peter Murray-Rust, Joe Townsend, Jim Downing. (Microsoft) Lee Dirks, Alex Wade, Oscar Naim, Mike Galos, Tim Haughton.

Current publishing practice
... is broken for data-rich science

- Non-semantic data
- Data publication difficult and unsupported
- Insufficient data to fully support research

With Chemistry Add-In for Word
... the cycle is closed

- Semantic data inside
- Data available for e-science and reuse
- Data directly mined from document
- Data preparation integrated into user workflow
- Open Standards promote Open Semantic Science

The Add-In (.NET) links a semantic chemistry engine (.NUMBO) through a command interface (CID) to a chemistry zone. Chemistry zones are textual or graphic renderings within a Word document (DOCX). All content and relationships are bound to CML in the DOCX package.

The navigator gives an overall view of the chemistry in the document and allows users to insert linked or copied data.

Different representations of the same data – all linked to the same backing CML.

A complete record of the changes to the data is retained in the customXML.