



# Abstract

## Reusing open-access content using authoring tools

Lieven Droogmans, C.E.O. @MIRE,  
<[lieven@mire.be](mailto:lieven@mire.be)>

### INTRODUCTION

@mire, with the support of the Flemish government, has developed a prototype in order to investigate the possibilities and usability of using open access repositories in the entire life-cycle of academic research, authoring as well as publishing. Several experiments have been conducted to test the feasibility of integrating open access repositories with authoring software tools and workflows. This leads to answers on the question: How can a repository aid a scholar in writing a (research) paper or prepare slides for a presentation? @mire had created a prototype setup and conducted usability testing.

@mire is a repository service provider, mainly active in the DSpace community, dedicated to promoting and enhancing the use of open access repositories in academic & research institutions.

### FEATURES AUTHORIZING TOOL

An array of possible innovative features for an authoring tool has been evaluated for technical feasibility and potential users from a diverse target group have been surveyed. The following features have been studied.

- Searching for relevant research in open access repositories from within the authoring tool.
- Adding bibliographical references by directly searching online open access repositories from within the authoring tool.
- Including paragraphs, images and other decomposed content derived from open access material based upon:
  - Digital object decomposition and storage of linked decomposed objects.
  - Automatically generating descriptive metadata.
- Depositing content into an open access repository from within the authoring tool.
  - Automatically generating and manually improving descriptive metadata.

The presentation will outline the most important conclusions for the different features.

A subset of these features have been carefully selected based on their feasibility and the feedback provided by potential

@mire nv  
Romeinse straat  
18  
3001 Heverlee  
☎ +32-2-  
888.2956  
[info@mire.be](mailto:info@mire.be)  
[www.atmire.com](http://www.atmire.com)

# Abstract

users. These features have been implemented as an add-on for the most popular office platform, Microsoft Office. For this prototype setup Microsoft Office 2007 (Word & PowerPoint) on Windows Vista was the selected platform for the authoring tool.

## DECOMPOSITION

In order for the open-access repository content (e.g. papers, presentations, ...) to be reusable it needs to be decomposed in smaller objects. For example, a research paper can be decomposed into text fragments, images, etc. and for each of these decomposed objects metadata needs to be added automatically in order for an image to be retrievable by a search query. This metadata can be extracted from the context of the decomposed object. For instance for an extracted image metadata can be extracted from the a caption, surrounding paragraphs, and the parent object's metadata.

### File formats

A number of experiments have been conducted to attempt to semantically decompose Microsoft Word documents (OpenXML/OOXML), PDF documents and Microsoft PPT presentations (OpenXML/OOXML).

### Issues

Every file formats has its advantages and its limitations. Some of the difficulties in interpreting and decomposing the different file formats (PDF, docx, pptx) will be outlined. For instance for (most) PDF documents there is little or no structure to be determined without interpreting the visual page layout.

## STORING & EXPOSING DECOMPOSED CONTENT

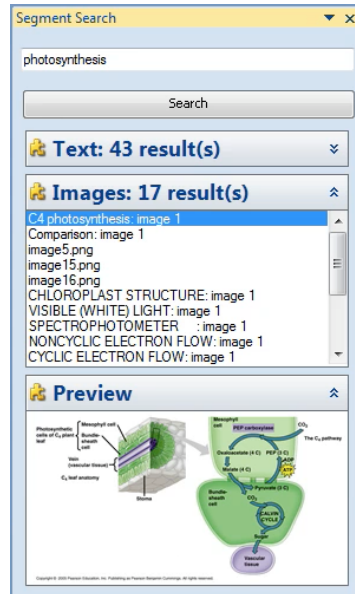
### Architectural overview

In order to offer decomposed repository content ready to be reused immediately in a text editor, content needed to be obtained, decomposed, stored and exposed. This presentation will provide an overview of the architectural prototype setup.

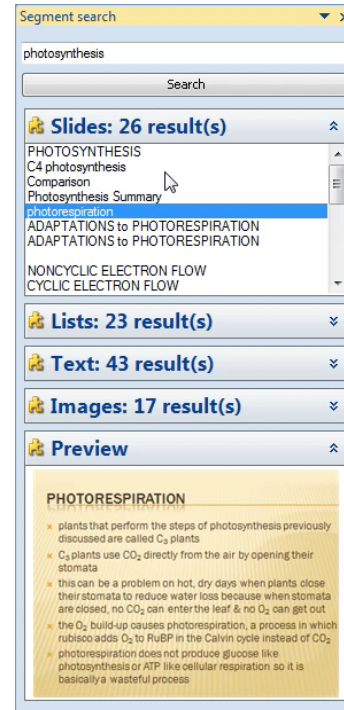
This setup included harvesting open access content into a repository. Decomposing the harvested content and automatically providing the decomposed content with metadata. Storing the decomposed content and its metadata and exposing it through webservices. From within Microsoft word and Powerpoint the decomposed content store can be queried and the content for the query results can be previewed and immediately added to the document.

# Abstract

Task pane in Microsoft Word



in Powerpoint



## USABILITY STUDY

### Introduction

The usability test was mainly targeted at researching the user's preferences and behavioral patterns to allow the pros and cons of the prototype setup to surface.

### The Tests

Based on 4 usage scenarios, the authoring tools have been tested in a renowned high-end usability lab at the K.U. Leuven on seven researchers from different academic backgrounds, and of different age. Before conducting the usability test, the researchers were surveyed to determine their information gathering and publishing profile. Ages varied from 22 to 55 years, 5 males opposed to 2 females from •-, •- and •-science disciplines, ranging from architectural history to materials science and engineering.

During the usability test, the researchers' behavior was recorded on video and synchronized with the computer screen and their eye focus and movement was recorded using an eye tracker. The four scenarios will be outlined and video fragments of the actual test as well as the eye tracker data will be presented.

### The Results

@mire nv  
 Romeinse straat  
 18  
 3001 Heverlee  
 ☐ +32-2-  
 888.2956  
 info@mire.be  
 www.atmire.com



# Abstract

The authoring tool was generally perceived as lightweight and easy to start to work with. The concept behind the tool had not been disclosed nor its purpose, but this was immediately clear to all but one of the researchers.

Everyone was able to complete the scenarios within the available timeframe. And after completing the four scenarios the users were again surveyed and were asked to give suggestions for improving the features they had tried out as well as additional innovative features. And more importantly they were asked to comment on the general setup and value of the authoring tool. In the presentation the feedback will be provided and summarized.

@mire nv  
Romeinse straat  
18  
3001 Heverlee  
☐ +32-2-  
888.2956  
info@mire.be  
www.atmire.com