

A Dataspace Odyssey: The iMeMex Personal DataSpace Management System

Lukas Blunschi, <u>Jens Dittrich</u>, Olivier Girard, Shant Karakashian, Marcos Salles ETH Zurich, iMeMex.org



What is a DataSpace Management System?

- Vision paper by Mike Franklin, Alon Halevy and David Maier
 From Databases to Dataspaces: a New Abstraction for Information Management.
 SIGMOD Record, December 2005
- Idea: ALL data belonging to a particular person, task, or company gets assigned to a logical dataspace, no upfront investments for schemas.
- Dataspace independent from formats, devices, locations, and systems owning the data.
- What we need is a new kind of system able to manage dataspaces across devices and formats.

From Database Systems to DataSpace Systems

- DataSpace management systems can be seen as a hybrid of
 - a search engine,
 - a databases management system,
 - an information integration system,
 - and a data sharing system.
- iMeMex is one of the first DataSpace Management Systems
- iMeMex is designed as a highly extensible framework (Eclipse-like, OSGi service-oriented architecture)
- Current focus of iMeMex: personal dataspaces
- However, we are also targeting other applications including scientific dataspaces.

iMeMex Project: Integrated MeMex

- Project started 2 years ago
- Our goal: Build the first open source dataspace management system
- All Java, service-oriented (OSGi, cmp. Eclipse project)
- Plus platform specific extensions in C (currently for Windows and Mac)
- iMeMex is an extensible framework for dataspace management
- Supported by SNF (Swiss National Science Foundation)
- Demo at VLDB '05, papers at VLDB '06 and SIGIR PIM '06
- State of the system, milestones:
 - December '06: 0.42.0, RC1a developer version released/open source



- Beg of Feb '07: 0.43.0, RC1b developer version
- End of Feb '07: 0.43.x, RC2 developer version + RC1 AJAX client
- End of May '07: 0.44.x, RC3 developer version + RC2 AJAX client
- End of July '07: 0.45.x, RC1 end-user version

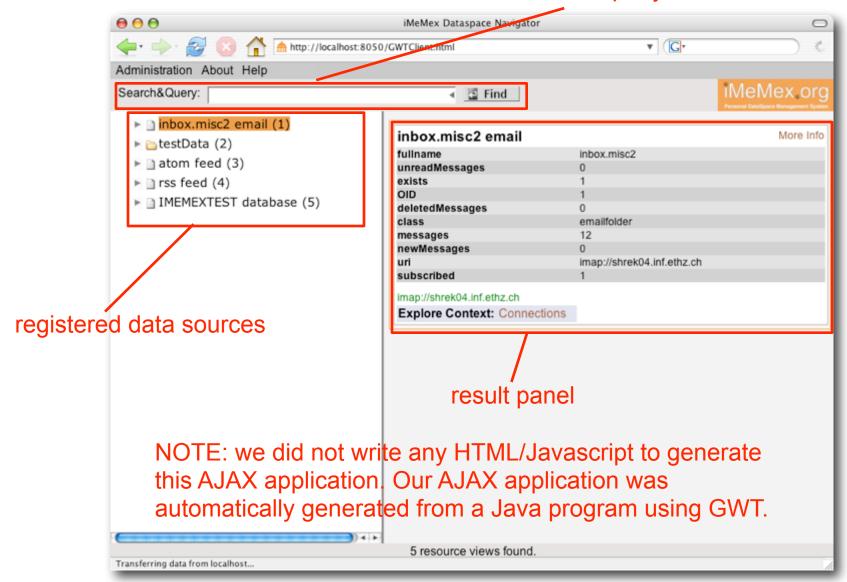
Literature on PIM and Dataspaces

- J. Dittrich, M. Salles, D. Kossmann, L. Blunschi iMeMex: Escapes from the Personal Information Jungle (Demo Paper) In VLDB 2005, September 2005.
- M. Franklin, A. Halevy, D. Maier
 From Databases to Dataspaces: A New Abstraction for Information Management SIGMOD Record, 34(4):27–33, December 2005.
- J. Dittrich iMeMex: A Platform for Personal DataSpace Management SIGIR PIM 2006, August 2006.
- J. Dittrich, M. Salles iDM: A Unified and Versatile Data Model for Personal Dataspace Management VLDB 2006 (IIS Track): September 2006.
- L. Blunschi, J. Dittrich, O. Girad, S. Karakashian, M. Salles
 A Dataspace Odyssey: The iMeMex Personal Dataspace Management System
 CIDR 2007 (Biennial Conference on Innovative Data Systems Research), Jan. 2007
- J. Dittrich, L. Blunschi, O. Girad, S. Karakashian, M. Salles, M. Färber From Personal Desktops to Personal Dataspaces: A Report on Building the iMeMex Personal Dataspace Management System BTW 2007 (Biennial German Database Conference), March 2007

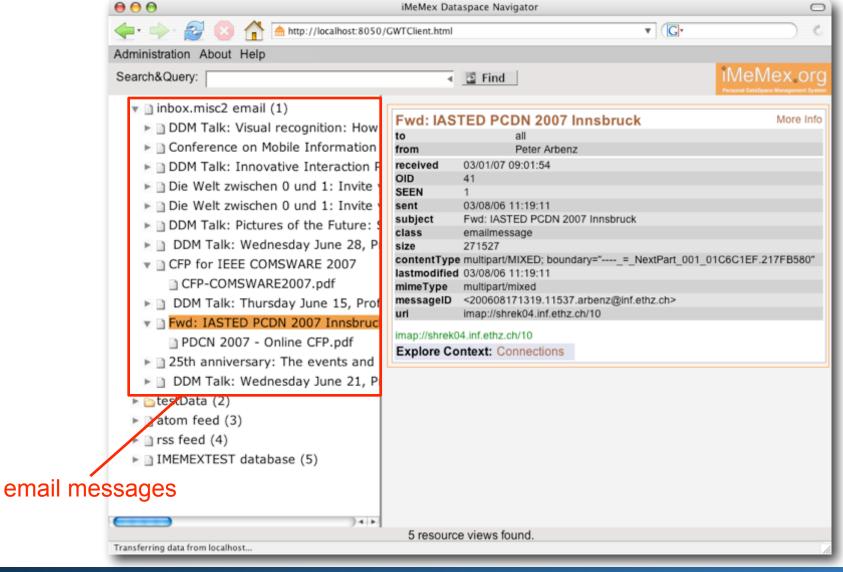
Demo Screenshots

iMeMex AJAX Interface

search&query box

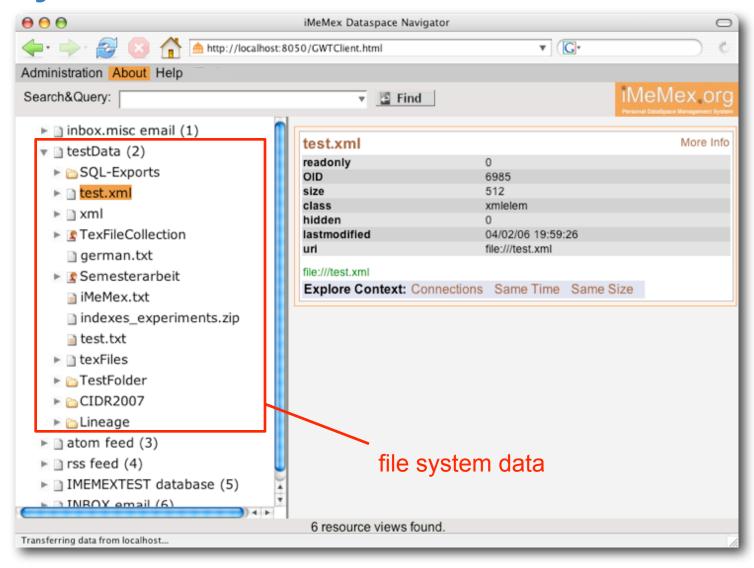


IMAP Mail Data Source



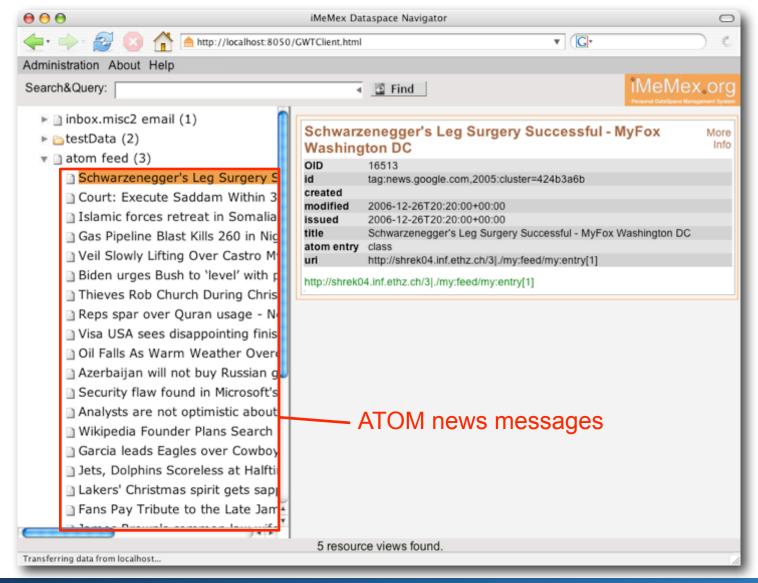


File System Data Source



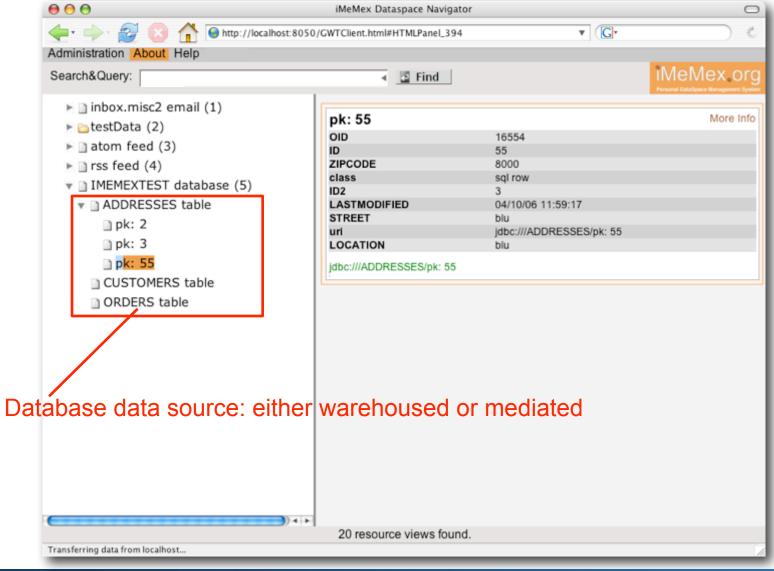


ATOM Feed Data Source



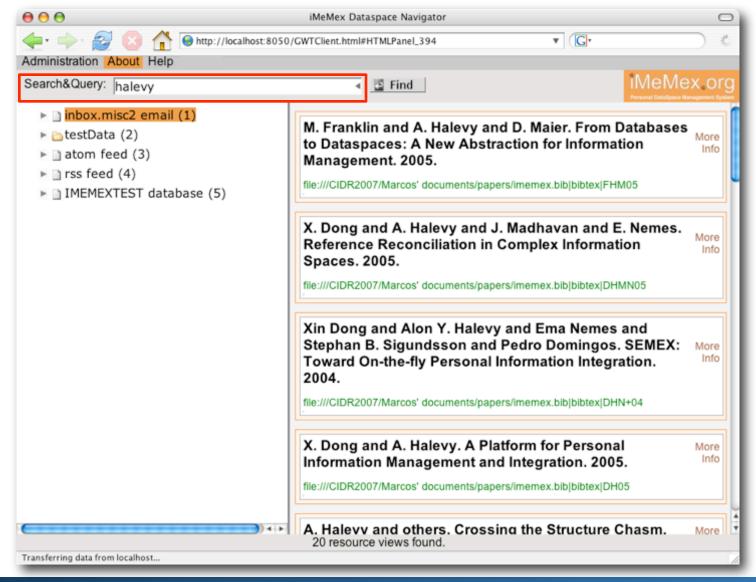


Oracle Database Data Source



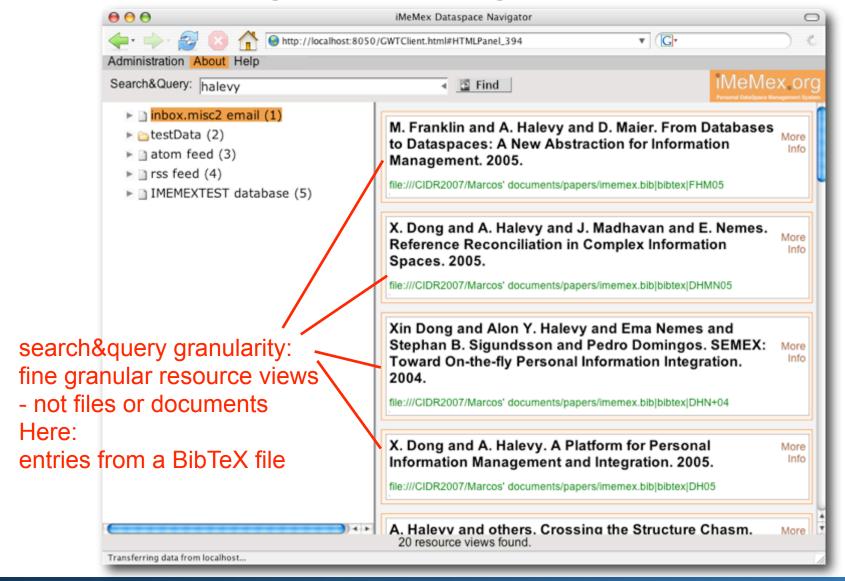


Keyword Search

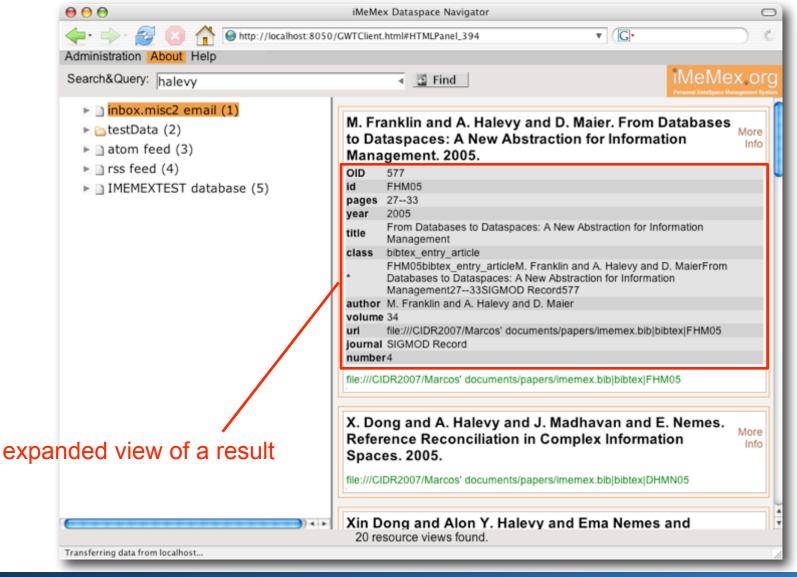




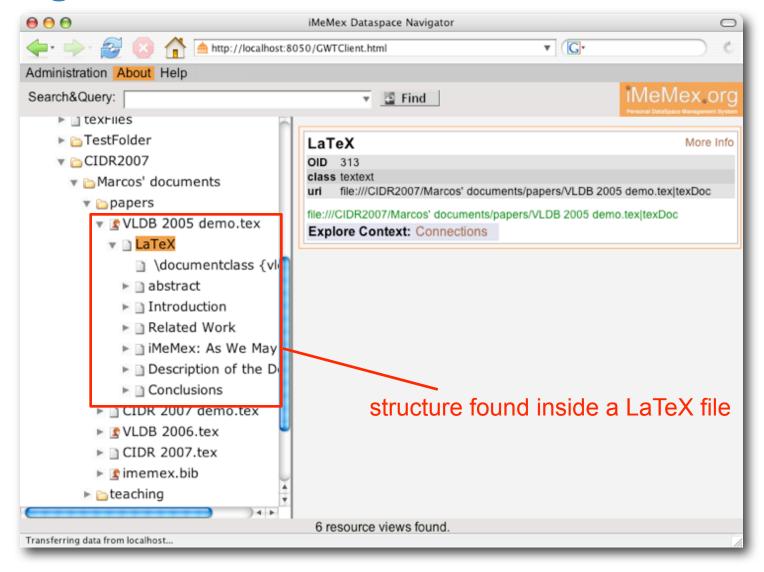
Search&Query Granularity: Resource Views



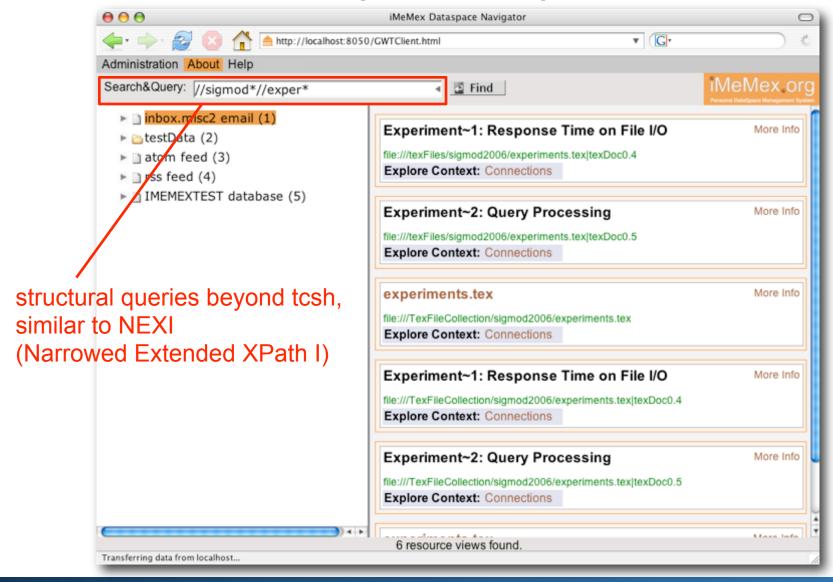
Search&Query Granularity: Resource Views



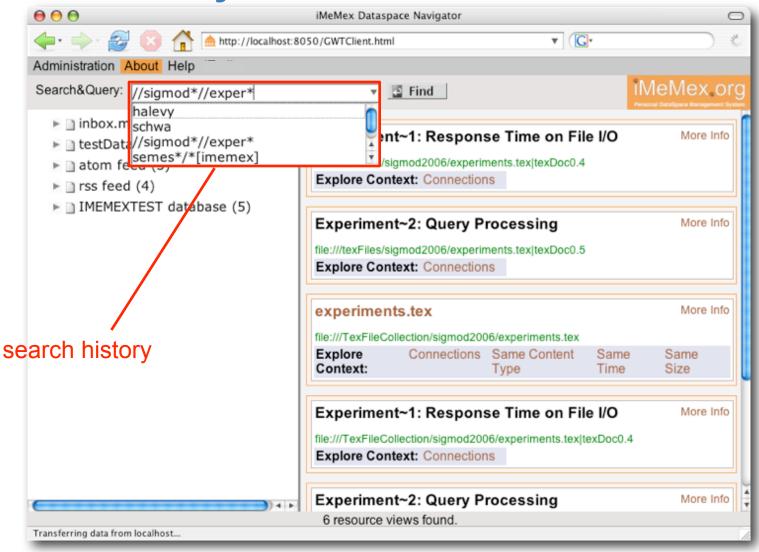
Navigate Inside Files



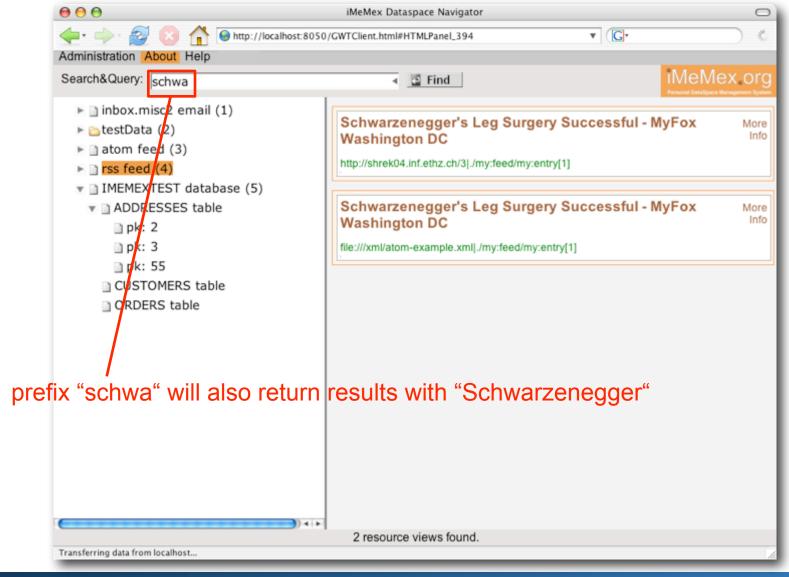
Structural Search (NEXI-like)



Search History

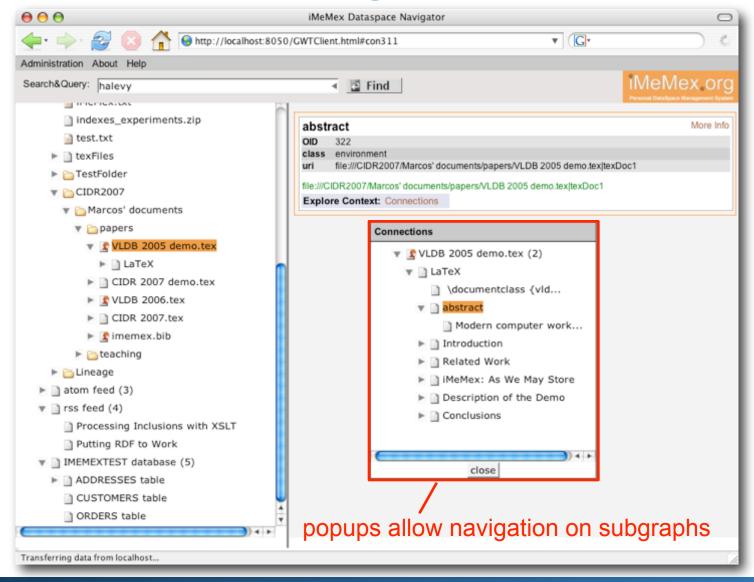


Keyword Prefix Search





Popups for Graph Navigation



Upcoming Features

- Declarative Pay-as-you-go Information Integration Framework
- Data Sharing Framework using distributed iMeMex instances
- Cost-based optimization
- <your plugin here>
- Details:

see our web-site: http://www.imemex.org