

Moteurs plein texte sous MySQL et PostgreSQL pour la gestion de connaissances

- Technologies for Knowledge Management
- Integration of Full-text Search and Relational Database
- Knowledge Management meets ERP

The logo for ERP5, featuring the text "ERP5" in a white, serif font on a blue rectangular background.The logo for Nexedi, featuring a stylized infinity symbol followed by the word "nexedi" in a lowercase, italicized sans-serif font, all on a white background.

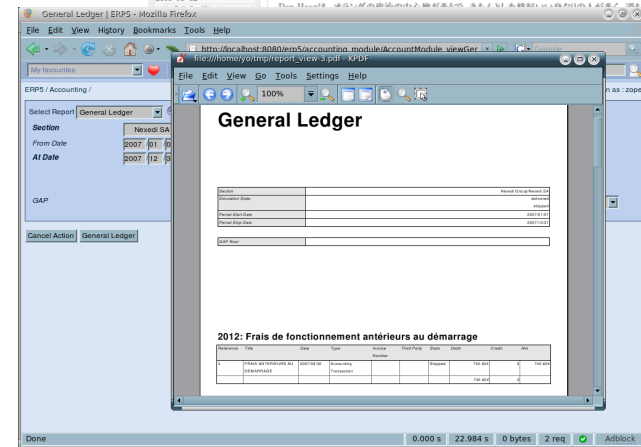
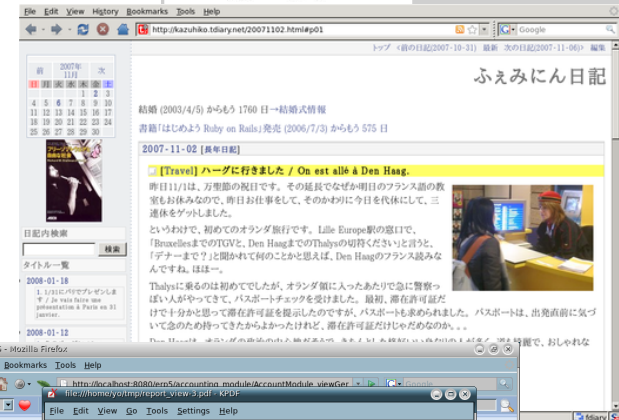
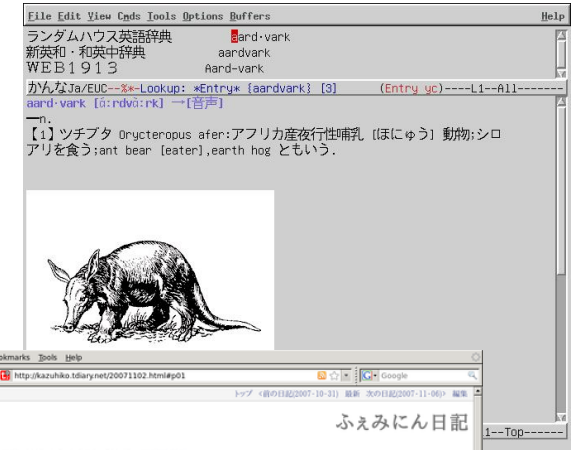
Kazuhiko Shiozaki
Senior Consultant, Nexedi
<kazuhiko@nexedi.com>



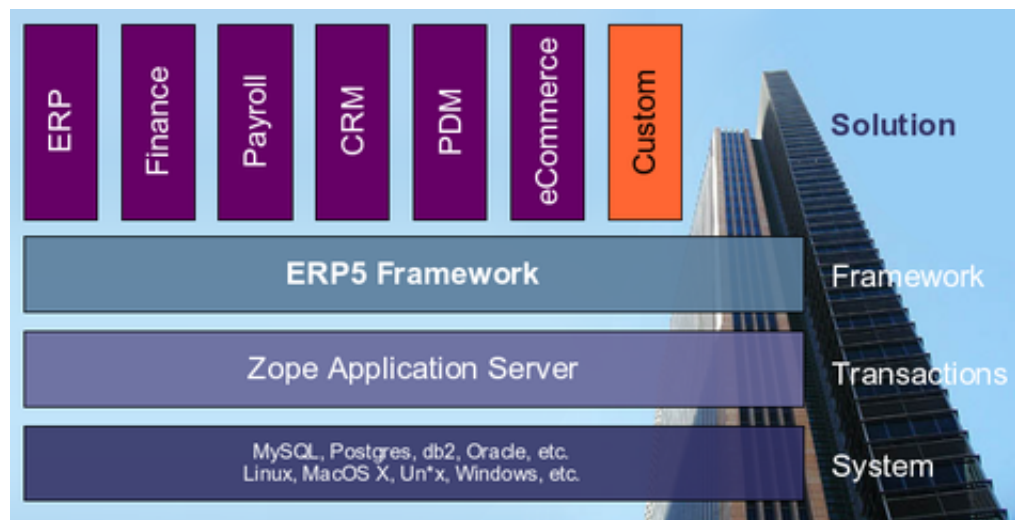
Self Introduction

ERP5

- Open Source Programmer
 - electronic dictionary
 - blog with full-text search
 - wiki with full-text search
- Developer of ERP5



- Open Source ERP (GPL)
- CRM, PDM, eCommerce, Finance, Payroll, ...
- and **Knowledge Management**

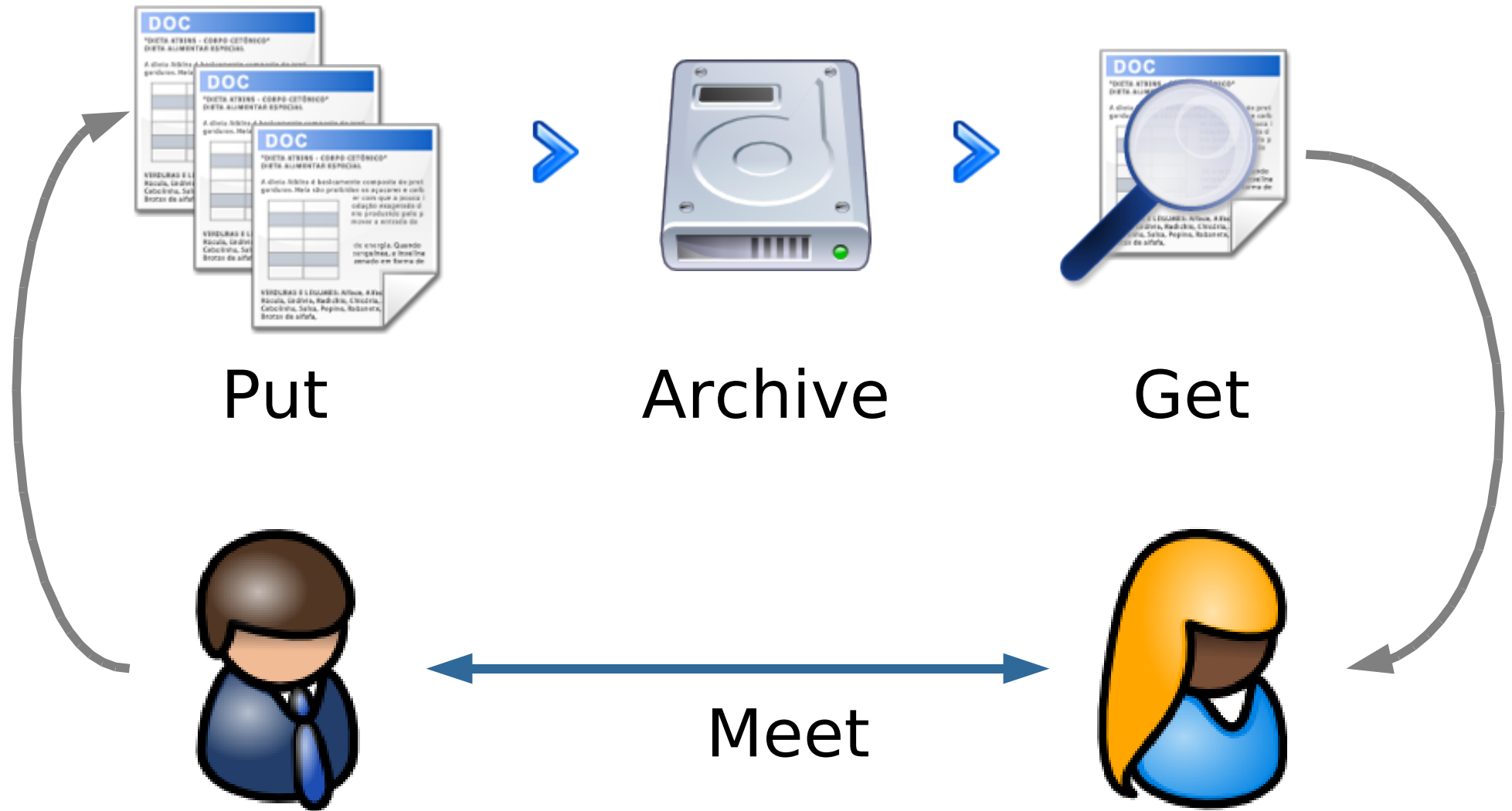


The right information
to the right **person**
at the right time

(by IBM's definition)

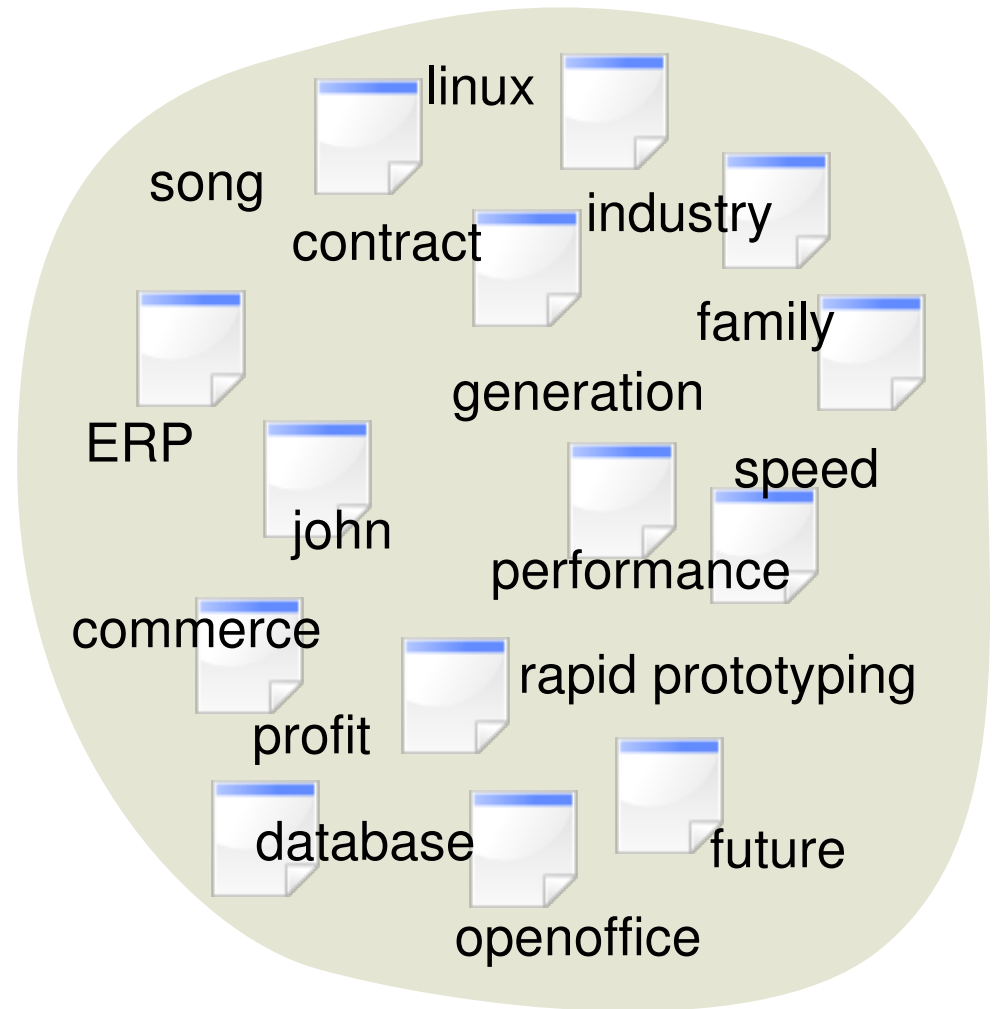
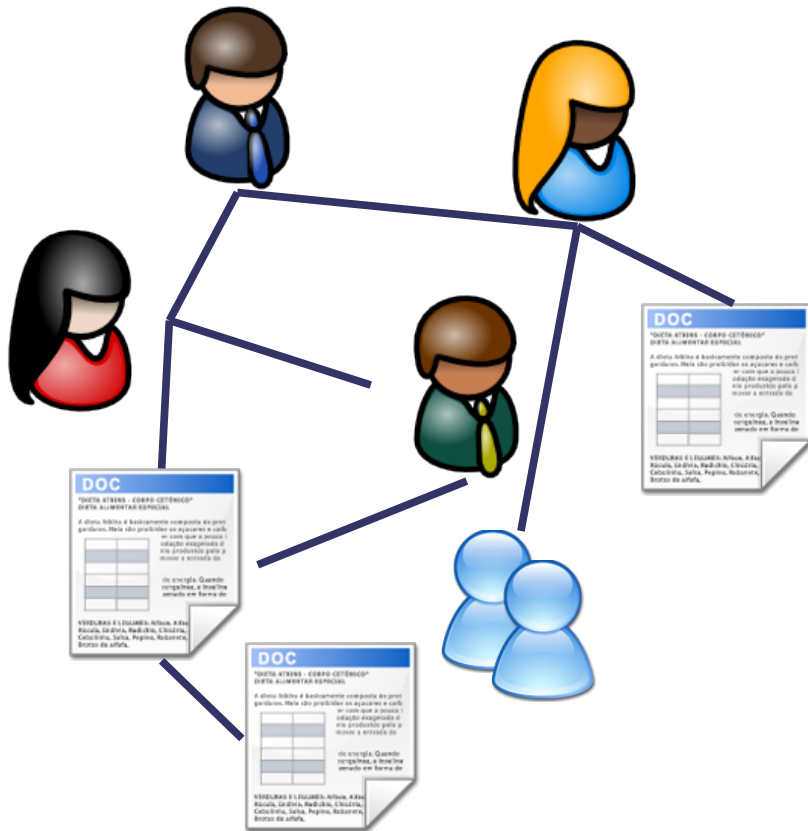


What is KM?



Attribute Filtering

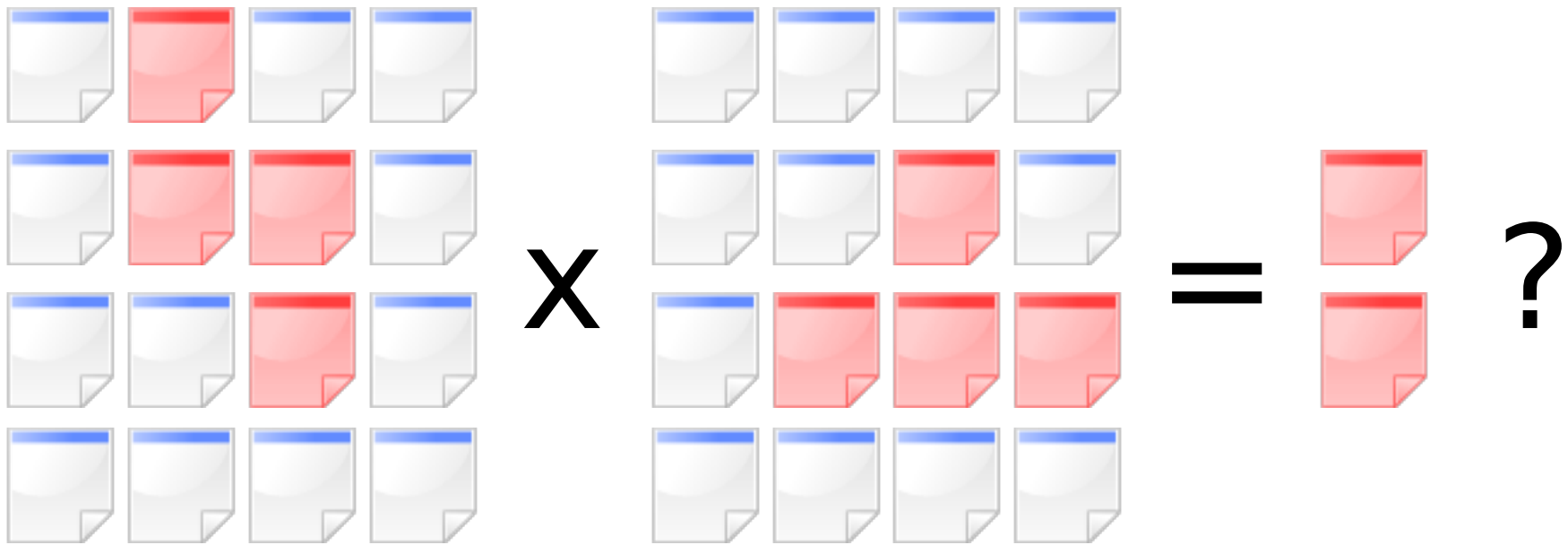
MySQL, PostgreSQL



Full-text Search

Lucene, Senna, Sphinx, Hyper Estraier etc.

- How can you get?
 - 10 documents ordered by modification time, that contain 'ERP5' and belong to Project X and are viewable for me



- Get the final result by only 1 query



- Both have full-text search but...
- No M17N support (eg. Japanese)
- No sub-string search
- Slow “phrase search”



PostgreSQL



Sphinx

- Sphinx Storage Engine for MySQL

senna

- Tritonn (MySQL + Senna)
- Ludia (PostgreSQL + Senna)

- Supports MySQL/PostgreSQL natively as source storage
- Supports distributed searching
- Fast indexing and fast search
- Slow updates
- <http://www.sphinxsearch.com/>



- A storage engine for full-text search using Sphinx
- You can use “JOIN” to get results for full-text search and other queries
- Indexing is not integrated



- Supports M17N search
- Fast “phrase search”
- Fast search without read lock
- Fast updates
- Supports related search
- <http://qwik.jp/senna/>



- Replaces MySQL's full-text search function
- You can get better full-text search **WITHOUT** modifying your application
- <http://qwik.jp/tritonn/>



- Add full-text search function to PostgreSQL
- <http://sourceforge.jp/projects/ludia/>



	Hyper Estraier	Lucene	Senna	Sphinx
Independent				
Connector	PGEstraier			Sphinx SE
Embedded			Tritonn, Ludia	
Distributed				
P2P				
Incremental				
Ngram				
M17N				

- 100,000 entries from Wikipedia (en)

	MySQL	Sphinx	Tritonn
Phrase Search	44.09s	0.31s	0.42s
Create index after inserting records	442s	148s	520s
Insert records with index	3645s	N/A	938s

- 100,000 entries from Wikipedia (en)

	Tsearch2	Ludia
Phrase Search	12.46s	1.61s
Create index after inserting records	635s	377s
Insert records with index	993s	1730s

KM meets ERP

ERP5

The screenshot shows the UNECA Knowledge Management System (KMS) web interface. The browser address bar displays the URL: `http://10.8.0.246:9280/erp5/web_site_module/uneca/WebSite_viewAsSCALE?portal_status_message=`. The page header features the UNECA logo and the text "UNECA Knowledge Management System". Navigation links include Home, Knowledge Pad, COPs, My Documents, My Tasks, My News, Preferences, and Logout. A search bar is present with a dropdown menu set to "All Documents" and a "Go" button. Below the search bar, there are tabs for "Search", "Browse", and "Query". A "Contribute" section includes dropdown menus for "Select Document Type", "Select Classification", and "Select Programme", along with "Make Template" and "Contribute" buttons. The main content area is divided into three columns: "Act", "Learn", and "Evaluate". The "Learn" column contains a list of recent documents with their titles and dates. The "Evaluate" column features a "Google Calendar" widget showing the month of April 2007. At the bottom, there is a "Google Map Search" widget displaying a map of a city area.



- Integration of full-text search and relational model is required for Knowledge Management
- Tritonn / Sphinx SE for MySQL
- Ludia for PostgreSQL

www.erp5.com

