Adopting Grid in Knowledge Management area

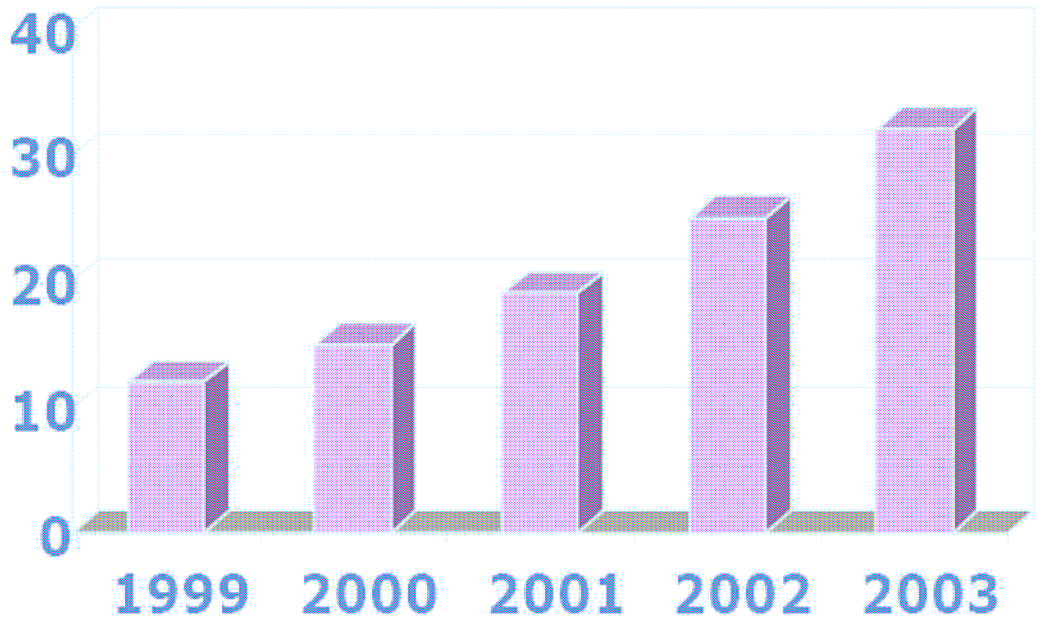
Lorenzo Gianoli
CEO - Gempliss

"Knowledge has became the main economical resource and a major, perhaps eventually the major source of competitive advantage "

Peter Drucker

Annual cost of knowledge deficit for US Fortune 500

Amounts in $ Billion



IDC estimates the cost of knowledge deficit at Fortune 500 companies to work out to about $5000 per knowledge worker per year. This is money wasted on repeated efforts, time taken to search out information, etc.

Fonte : IDC

"What we cannot speak of we must pass over in silence."

Ludwig Wittgenstein, Tractatus, 7

Grid Infrastructures are born to solve problems and problem solving is deeply related to creation and use of knowledge.

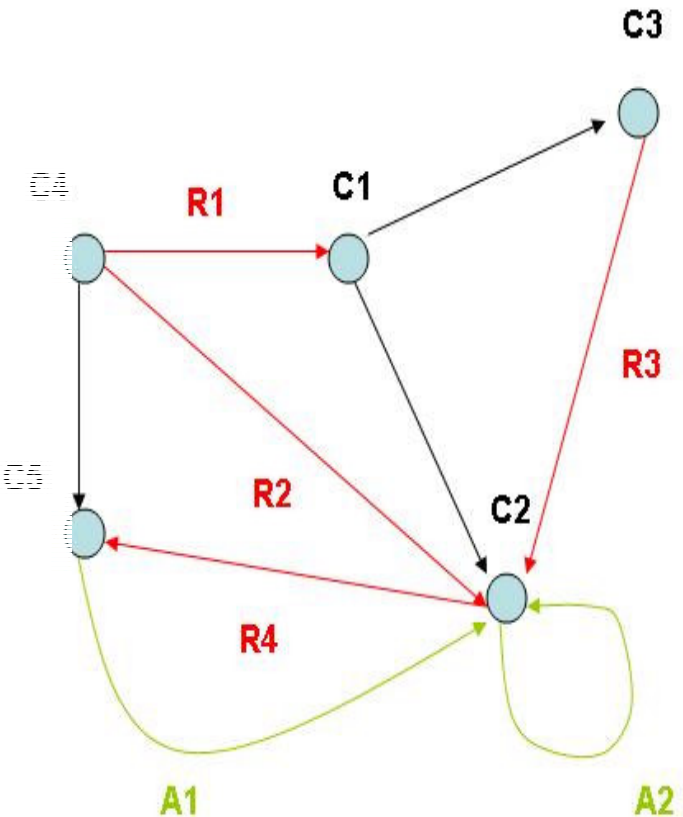
HOW TO CREATE KNOWLEDGE

- The only possibility to really create knowledge is "to read" information and analyse it semantically, understanding lexical structures end extracting concepts and relations to represent knowledge
- Through Semantic, information and services are readable both to humans and machines making possible inferences and dynamic discovery

ONTOLOGIES

After analysis Knowledge representation can be made through ontologies, that are structure composed by:

- **Taxonomies:** Concepts trees (hierarchical)
- **Properties**: Relations inside taxonomies
- **Axioms:** Rules related to concepts and relations



BUT…

1. Semantic Analysis is an heavy calculus job
2. Inside KM nobody is able to say how much information we deal with
3. Social processes are very strong (Is Knowledge power?)
4. Information is loosely integrated
5. Implicit knowledge is a problem
6. Flexibility and low impact from a structural point of view

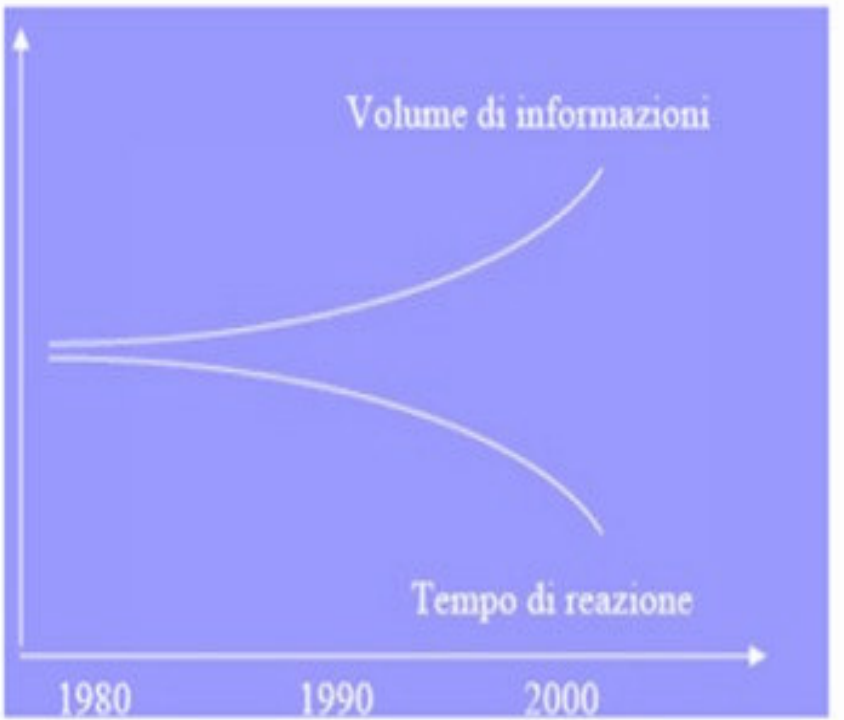
1. Semantic Analysis is an heavy calculus job

Grid is able to supply flexible calculus power, that is already inside company

2. Inside KM nobody is able to say how much information we deal with

Grid let scalability, high QoS and independency from underlying infrastructure (Operating system, HW,…).

KM could require a scaling of order of magnitude in a short time

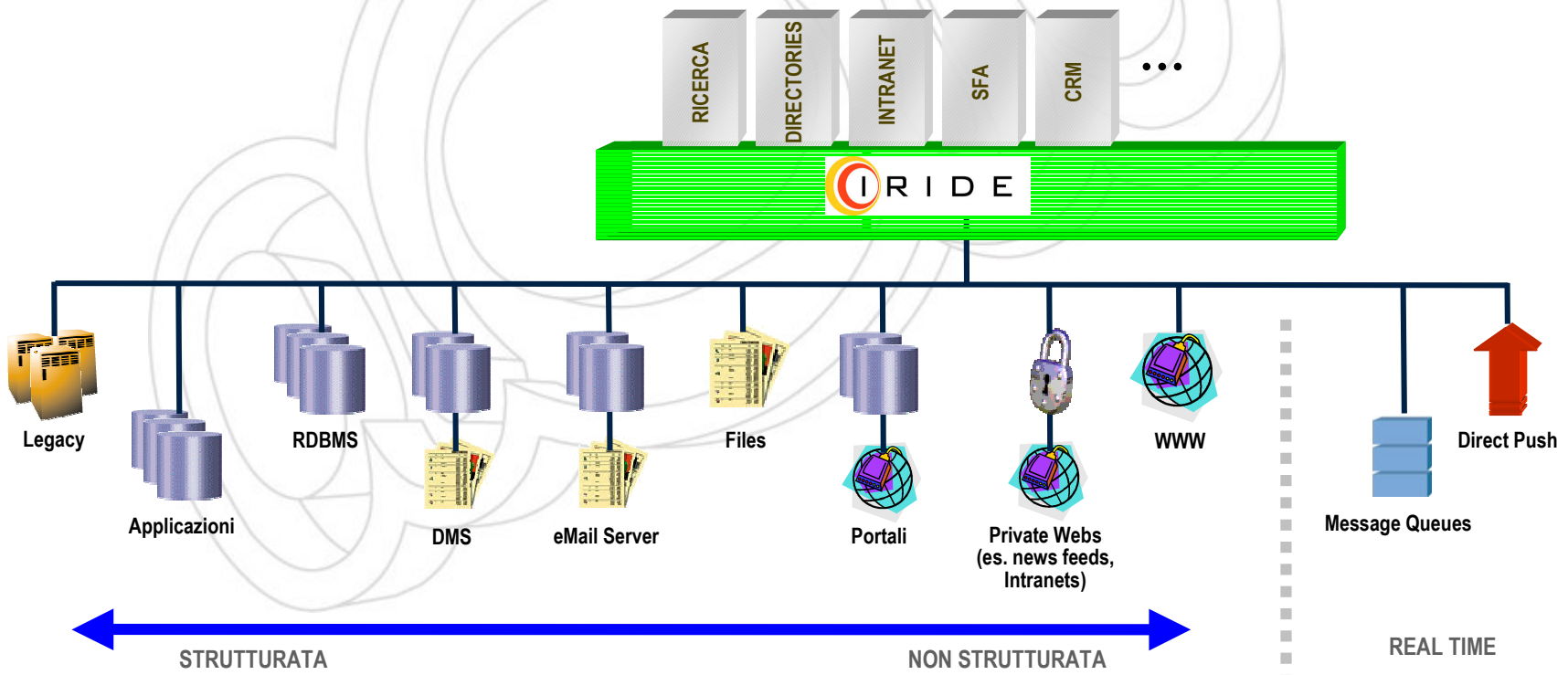


3. Social processes are very strong (Is Knowledge power?)

Grid let us to remove pysical barriers from local to network resources. Local element could be calculus resource, storage area, object of analysis. User look at a logical environment where information is loosely relkated to physical elements

4. Information is loosely integrated

Due to calculus power and distributed paltform systems can be globally analysed and reached through a single point of access



5. Implicit knowledge is a problem

A Knowledge Driven organization, where explicit knowledge is shown is nearly to emphasize competences. Dynamic KM make possible to put in evidence ensemble of implicit knowledge.

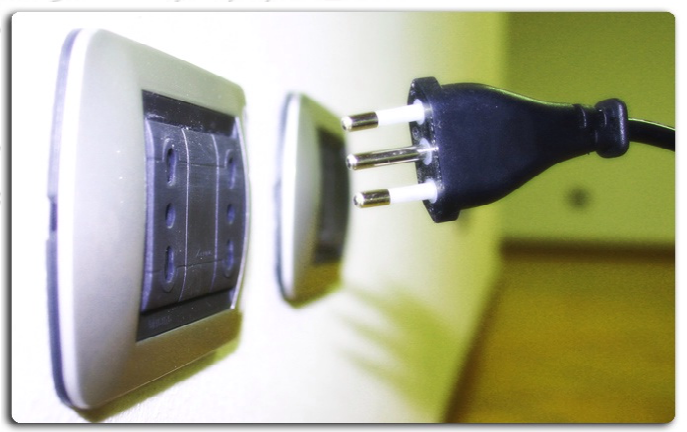
6. Flexibility and low impact from a structural point of view

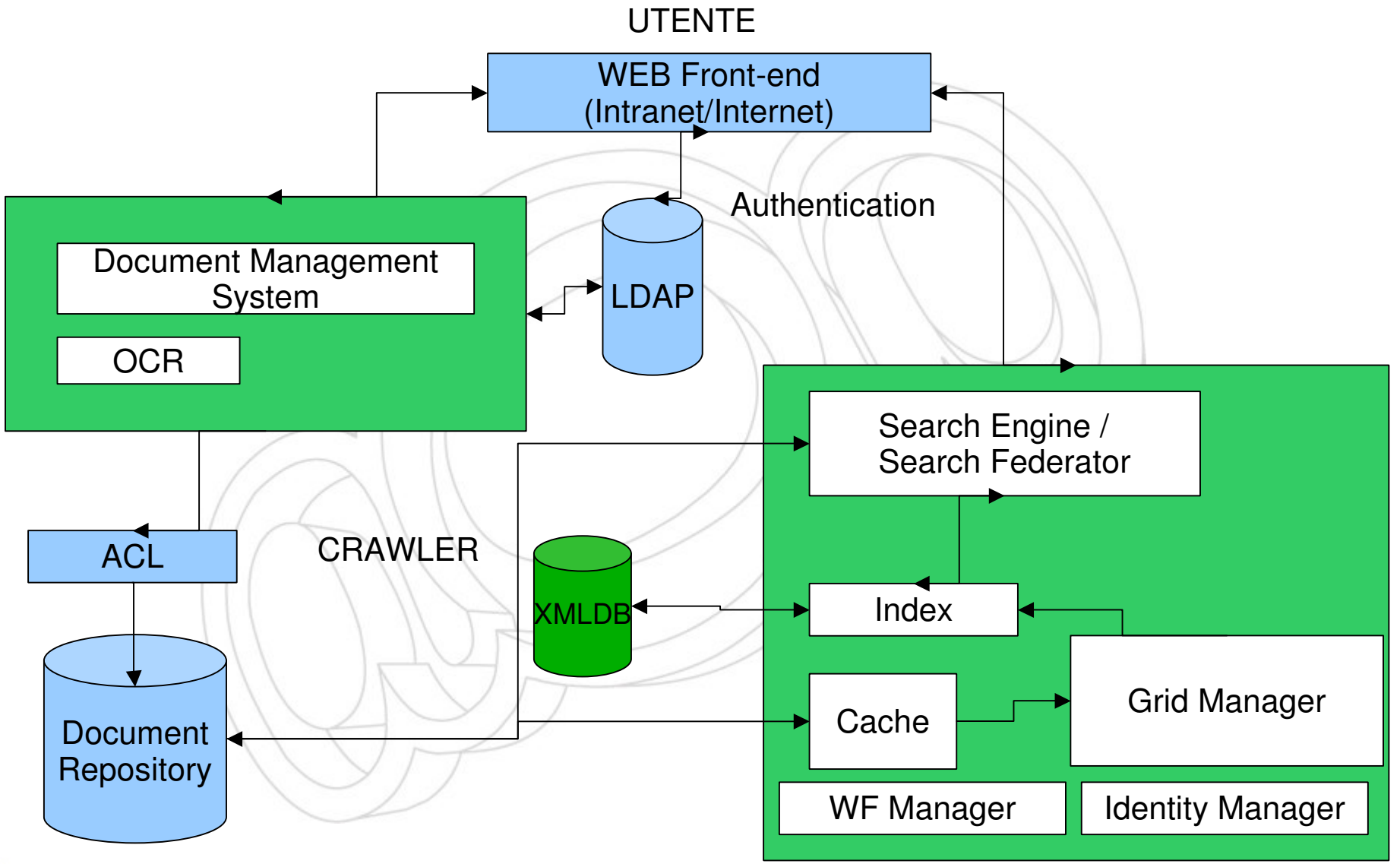
Grid let to change the logic of analysis making possible the realisation of flexible, non invasive platforms. SOA architectures are important to define really modular services.

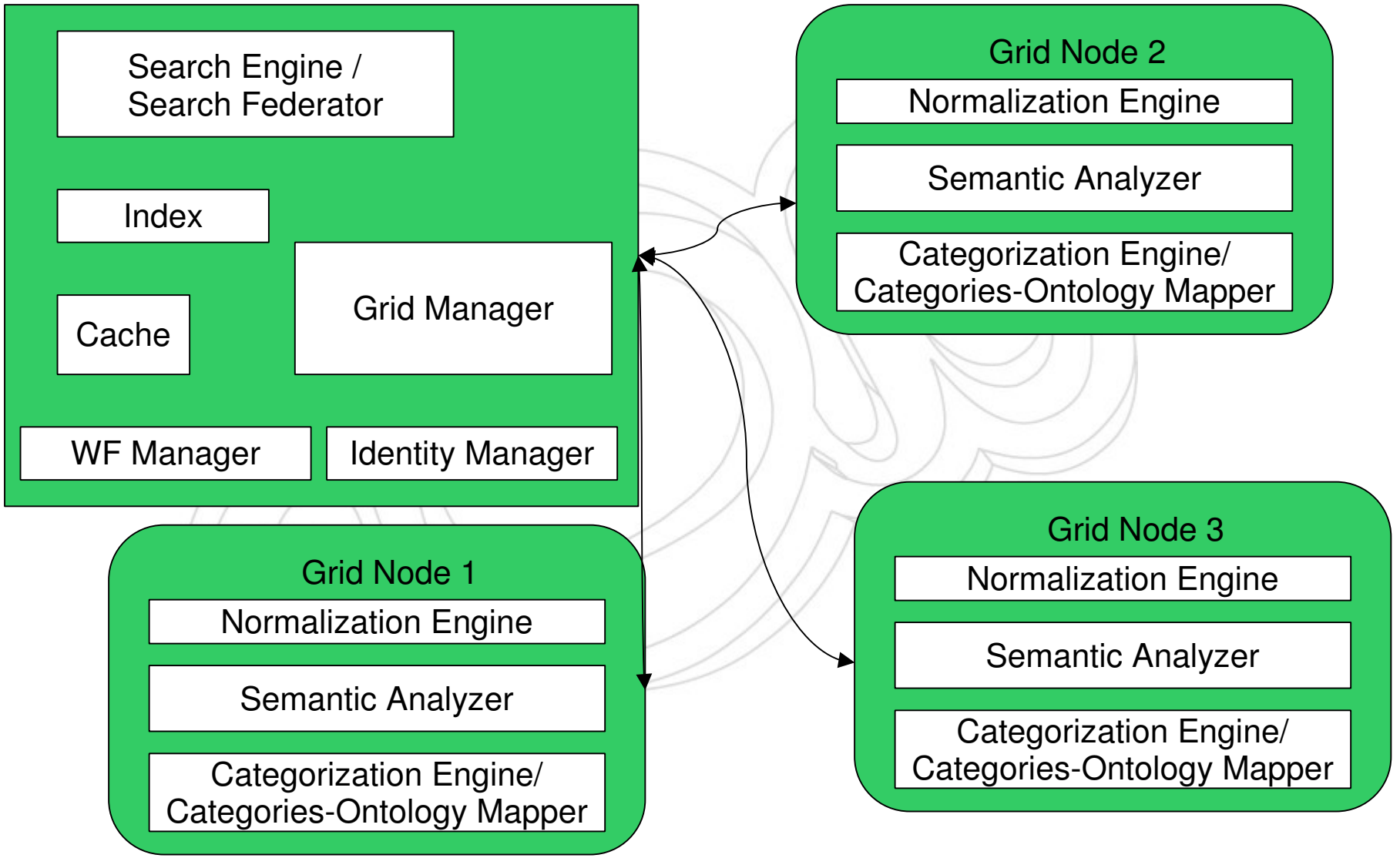
Distributed Knowledge Management

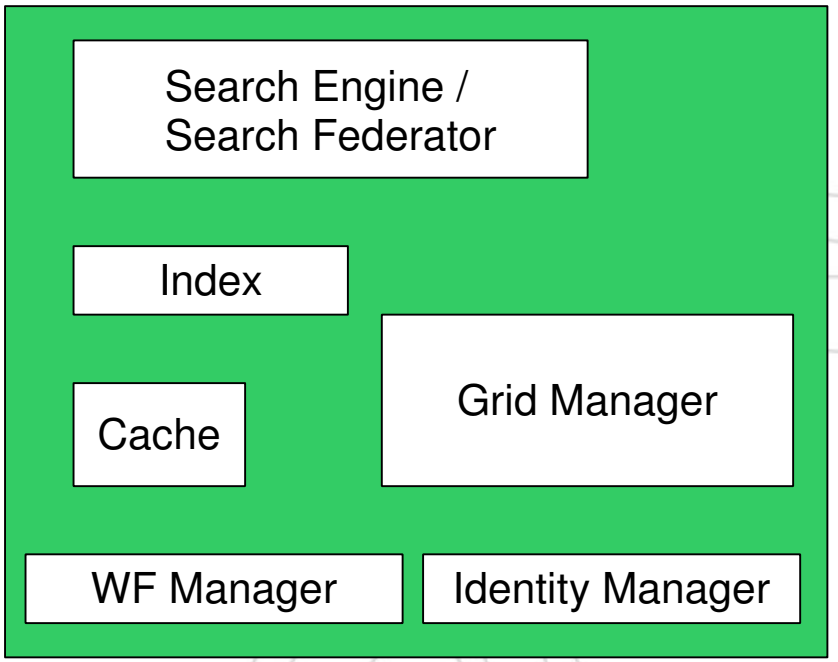
Gempliss is the first company able to realize a distributed knowledge management platform entirely SOA and GRID based.

For us GRID is not only a way to federate services connected with infrastructure but is the infrastructure able to deliver power to services.




UTENTE
WEB Front-end
(Intranet/Internet)
Authentication
Document Management
System
OCR
LDAP
Search Engine /
Search Federator
ACL
CRAWLER
XMLDB
Index
Document
Repository
Cache
Grid Manager
WF Manager
Identity Manager


Search Engine /
Search Federator
Index
Cache
Grid Manager
WF Manager
Identity Manager
Grid Node 2
Normalization Engine
Semantic Analyzer
Categorization Engine/
Categories-Ontology Mapper
Grid Node 1
Normalization Engine
Semantic Analyzer
Categorization Engine/
Categories-Ontology Mapper
Grid Node 3
Normalization Engine
Semantic Analyzer
Categorization Engine/
Categories-Ontology Mapper



- Globus Toolkit 4
- Java WS
- CAS
- Reliable file Transfer
- Condor Pool

Central Manager

- Globus Toolkit 4.0
- Condor 6.6.10
- Linux kernel 2.4/2.6
- Java 1.4.2

Executor

- Condor 6.6.10
- Java 1.4.2
- OS Independent (Linux,Windows, Mac OSX or BSD)

Client

- Web Browser (IE 6.x,FireFox,Opera or Safari)

IRIDE

Multiple sources: File server, DMS, DB, Internet, Open Archive, paper documents (Integrated OCR), audio formats

All formats: XML, HTML, pdf, doc, xls, ppt, txt,……

Fully integrated with security

Semantic analysis, logical views through ontologies

Infinite scalability

Local use of platform for user itself (like desktop search)

ORIDE

Semantic and statistical filters

Rank measure

Abstract creation for every document (doc, pdf,…)

Search agents, shared researches, shared results

Multilingual (English, Italian, French, German, Polish)

Personalised Knowledge Domains

DMS functions (WF Management, Document Warehouse,….)

APPLICATIONS:

ORIDE

- *KM*
- *CRM*
- *SFA*
- *Contact Centers Automation*
- *Concurrency analysis*
- *Human Resource*
- *.....*

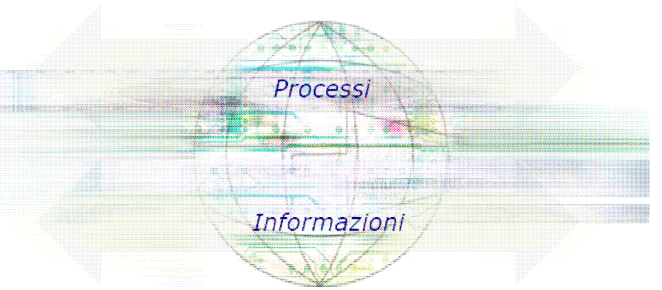
ORIDE

Concepts and relations are useful to find but also to classify, extract specific informations, to aggregate informetions around concepts, to generate automatic reply, to realise dynamic discovery

ORIDE

R&D

Measure of Knowledge Diffusion processes (prototype is in progress)



Navigational 3D Ontologies maps

Images Semantic