



# **Software Product Description**

AMI Enterprise Intelligence version 4.0

SPD-AMIEI-40-ENG v1.0 February 2008

# Table of Contents

1	Description .....	5
2	Approval.....	5
3	Principal Developments .....	6
4	Introduction.....	8
4.1	Issues addressed by AMI Enterprise Intelligence .....	8
4.1.1	Extending the information perimeter.....	8
4.1.2	Develop and Share acquired knowledge .....	9
4.1.3	Exploiting quality information in real time.....	9
4.2	Key Features .....	9
4.2.1	Key functional features .....	9
4.2.2	Key technical features .....	11
5	Functional description.....	13
5.1	Functional architecture .....	13
5.1.1	The Automatic Meaning Interpreter Base Server (AMI) .....	13
5.1.2	Functional applications .....	14
5.2	Example: Market intelligence .....	16
5.3	The AMI E.I. Homepage .....	17
6	Administration Functions.....	18
6.1	Physical Installation Overview.....	18
6.2	Administration interface .....	18
6.2.1	Management of Information sources.....	18
6.2.2	Administration of Users and Groups.....	20
6.2.3	Knowledgebase Management .....	21
7	Accessing Information : Collect, Search, Contribute .....	22
7.1	Principles of Collect and Search .....	22
7.2	The "Collect" Application .....	23
7.2.1	Collect plan .....	23
7.2.2	Defining scenarios .....	24
7.2.3	Execution of the Collect Process .....	25
7.2.4	Email Notification .....	26
7.3	The "Search" application .....	27
	User interface.....	27
7.3.1	Conducting a Search.....	28
7.3.2	Search results and manual population of the Publishing space.....	28
7.4	"Contribute" Application .....	29
8	Developing Information: Publish, Analyse.....	31
8.1	Main principles of Publish and Analysis .....	31
8.2	The "Publish" application .....	31
8.2.1	Function: "New Document" .....	31
8.2.2	Function: "Document Validation" (Publication).....	32

8.2.3	Synthesis.....	34
8.2.4	Function "Search Documents".....	35
8.2.5	Function: Classification Plan.....	35
8.2.6	Function: Document Rules.....	36
8.2.7	Function: Document Administration.....	37
8.3	The "Analyze" application.....	38
8.3.1	Measuring contribution by source.....	38
8.3.2	"Volume and Centering".....	40
8.3.3	Transversal Analysis of the Classification Plan.....	40
8.3.4	Analysis Tools: "Concepts" (detection of tendencies).....	41
8.3.5	Analysis Tools "Ontologies".....	42
8.3.6	Analysis Tools "Cartography".....	42
9	Sharing and Distributing Information.....	45
9.1	MyAMI Application.....	45
9.2	"Share" Application.....	47
9.2.1	Viewing Documents.....	48
9.2.2	Generation of RSS feeds.....	49
10	Advanced functions.....	51
10.1	Advanced linguistic functions.....	51
10.1.1	Signature concept.....	51
10.1.2	Relevance Calculation.....	51
10.1.3	Identification of the Query language.....	51
10.1.4	Spelling approximation.....	52
10.1.5	Synonym management.....	52
10.1.6	Stop word management.....	52
10.1.7	Filtering out outdated documents.....	52
10.1.8	De-duplication.....	52
10.1.9	Extraction of Citations.....	52
10.1.10	Trend detection.....	53
10.1.11	Automatic learning.....	53
10.2	Generic connector.....	53
11	Compatibility with other AMI applications.....	54
11.1	Source management compatibility.....	54
11.2	Knowledgebase compatibility.....	54
12	Compatibility with other applications.....	55
13	Documentation, contact and support.....	56
14	Appendices.....	57
14.1	Technical Appendices.....	57
14.1.1	Physical architecture.....	57
14.1.2	Security.....	57
14.1.3	Scripting language.....	57
14.2	Legal aspects on the use of information sources.....	58
15	Index.....	59

AMI (Automatic Meaning Interpreter™) and "What you mean is what you get" are registered trademarks of Go Albert SA. All the other product names are registered trademarks of their respective owners.

## 1 Description

This document constitutes the functional description of the product:

**AMI Enterprise Intelligence v4.0 (reference SPD-AMIEI-40-ENG).**

## 2 Approval

<b>Approval</b>		<b>Date</b>
Alain Beauvieux	CEO	1 <sup>st</sup> Feb 2007
Eric Fourboul	CTO	1 <sup>st</sup> Feb 2007

## 3 Principal Developments

AMI Enterprise Intelligence version 4.0 includes a significant number of application and functional developments compared to version 3.2 which was released in September 2007. This document describes the new applications and functions delivered with AMI Enterprise Intelligence version 4.0

### ***Look and Feel***

**AMI Enterprise Intelligence v4.0** features new interface designs throughout. Key functions have been made more intuitive along with improvements in navigation and the way in which information is presented.

### ***Distribution of information***

**RSS Publication:** Documents issued by an Administrator can be published in the form of RSS feeds allowing information to be pushed either to a wide range of recipients, for example via integration of the feed into a portal or to individual users via subscription to the feed through an RSS reader available in e-mail clients like MS Outlook.

Feeds contain a link to the document within AMI, to the original page or, can optionally integrate the entire contents of the document. They can also be password protected.

**Export as Microsoft Word** and other popular formats: Export in Word XML format is now extended to dashboards created in the MyAMI application allowing conversion of an entire dashboard and its subsequent distribution directly via email as a Word document or other format such as PDF retaining all points of interactivity and maintaining ease of navigation across the AMI Enterprise Intelligence suite.

### ***Knowledge Capitalisation***

**Syntheses:** The document synthesis function has evolved considerably. Summaries of documents generated in version 4.0 are direct composites incorporating the content of the documents specified in the synthesis process.

**Reading Tools:** AMI automatically highlights key elements of a document, one of the functions performed by AMI's automatic textual analysis algorithms.

### ***Knowledge Modelling***

Objects can be defined which represent cornerstones of the organisation such as people, products or articles and the relationships between these objects can be used to implement the knowledge modelling process. Cross-referencing of documents collected with those already published is done automatically making it possible to link existing structured information to information managed by AMI.

Knowledge is represented using the RDF Schema allowing the import of external knowledge represented by the same format. Additional tools allow browsing of the knowledge network and help create "syntheses" that directly integrate the contents of those documents referenced.

- ✓ Navigation within the knowledge network
- ✓ Ability to synthesize elements of knowledge

### ***Analysis Engine***

AMI Analyze has been further developed in terms of both its capabilities and form

- ✓ New features are included: concept clouds, new document viewer, better synonymy management.
- ✓ All charting and graph presentation elements have been re-designed and improved

### ***Search Engine***

- ✓ A new search interface is available in AMI Publish and AMI Share providing the user with a greater range of options to pre-define search criteria allowing faster and more accurate queries of the knowledgebase to be conducted.

### ***Source Management***

- ✓ Certain source management functions, to date only accessible via the back office, have been made available via the Administration interface giving the user broader abilities to administrate things such as query expansion parameters, restriction by file type, http authentication etc.
- ✓ A test mode has been added.
- ✓ The processing of Flash format files (swf) is optimised.

### ***New transversal functions***

- ✓ Ability to capture a manual search within a scenario (the automatic collect) i.e. providing a "bridge" between AMI Search and AMI Collect.
- ✓ Ability to access knowledgebase management functions directly from AMI Analyze in order to include or change terms and named entities directly from results generated through analysis processes.

### ***Back-Office***

Log Manager: The logger has been completely revised, expanded and improved.

## 4 Introduction

**AMI Enterprise Intelligence™** is a software solution made up of independent *applications* and based on an innovative patented technology designed to collect, analyse and distribute targeted content from information sources which can reside inside or outside an Organisation. AMI Enterprise Intelligence is built on the AMI technology kernel, which can be acquired separately in the form of AMI Enterprise Intelligence Base Server.

Given its modularity **AMI Enterprise Intelligence™** provides the basis for various types of applications and solutions ranging from strategic intelligence initiatives through to Call-Centre Knowledgebase Management.

### 4.1 Issues addressed by AMI Enterprise Intelligence

Appreciative of the fact that all Organisations have varying reasons for needing to collect and analyse information they often fall into the following categories:

#### 4.1.1 Extending the information perimeter

AMI Enterprise Intelligence™ allows collection of information from the widest possible range of internal and external sources with the extraction of relevant information.

#### **Enhanced decision making**

The growth of information networks such as the Internet has given rise to the need for a new type of activity which consists of collecting and analysing information for strategic *decision making* purposes. The high speed at which information can be published and circulated has created the requirement to be alerted to the emergence of new information in support of effective and timely decision making processes.

#### **Promote information sharing**

The exponential growth in the number of online information sources calls for efficient tools that can automatically detect what's useful amongst information which is often highly unstructured, subject to rapid change and overall becoming harder to access.

#### 4.1.2 Develop and Share acquired knowledge

AMI Enterprise Intelligence provides an efficient means by which to share information and develop knowledge across an organisation.

##### **Capitalise and Develop**

While millions of documents can be stored on an intranet or portal, relevant documents are often hard to identify. AMI Enterprise Intelligence ensures that the relevant documents are made available to the right people at the right time turning an information “safe” into an information asset.

##### **Analysis and enhanced decision making**

The majority of decision support tools handle only structured information such as lists of customer addresses or financial reports. Purely textual information is seldom structured and yet conceals a hidden potential in terms of enhancing the decision making process which is still largely unexploited. AMI Enterprise Intelligence’s text analysis capabilities can, for example, help to better understand a customer’s expectations, detect changes in a competitor’s product development or technical capabilities or capture market trends key to improving the decision making process.

#### 4.1.3 Exploiting quality information in real time

Generally, in terms of fully addressing the problems above regarding the emergence of new material, its relevance, and flexibility in distribution it has become necessary to be able to apply rules and management processes to this task.

### 4.2 Key Features

**AMI Enterprise Intelligence**, and its embedded kernel, the **AMI Base Server**, include advanced features which promote the intelligent management and use of information.

#### 4.2.1 Key functional features

##### **Kernel and Web interface**

**AMI E.I.** is composed of a web interface which provides a single point of access to all applications and a *kernel* which collects and processes documents.

##### **Modular and cross-functional**

**AMI E.I.** applications are independent, complementary, and interacting. They are packaged and can be assembled to create a process chain dedicated very precisely to an organisation’s requirements for managing and using information. Together they encompass all functions necessary for processing information from collection to delivery.

## **Dynamic, Personalised Dashboards**

MyAMI is a dynamic configurable workspace allowing presentation of multiple streams of information reflecting very accurately a users defined centres of interest. MyAMI allows the creation of templated newsletters and personalised dashboards from both user generated content and third party sources.

## **Ontology Management**

AMI I.E. allows the management of ontologies whose contents can be associated, by matching, with documents collected and published. Support for RDFS allows synchronization between AMI and external knowledge modeled according to the RDF schema.

## **RSS Publication**

Documents published by Administrators are exportable in RSS feed format and can hence be made available to users via tools in every day use such as web browsers and e-mail clients.

## **Source Coverage**

Information accessed with **AMI E.I.** can reside outside (www, competitor's websites, forums, real-time newsfeeds, newsgroups etc.), or inside the organisation (Portals, Content Management Systems, E-mail archives, Databases, etc.).

Virtually all information sources that require a specific interface such as search engines and online forums can be accessed with AMI's *Generic connector*, an application driven by AMI's innovative patented technology.

## **Ease of Integration**

**AMI E.I.** connectors also provide access to most third party Document, Content Management and Database applications. The data flows (i.e. queries, results...) are generated in XML format and can easily be utilised by other platforms and applications.

## **Intelligent collection of information**

During the document collection process, **AMI E.I.** automatically identifies and *de-duplicates* documents. It also extracts key expressions, i.e. expressions which are related to the query or which characterise the document.

This data, along with other data such as author names, information sources, scores or trends, are added to the original document and can be delivered to users in e-mail messages or news feeds.

De-duplication is a major feature which considerably enhances productivity and performance.

## **Publish**

AMI Publish has many features aimed at simplifying the tasks associated with reviewing and publishing documents such as:

- ✓ Multiple document selection.
- ✓ Multiple document validation.
- ✓ Facilitated reading functions.
- ✓ Visualisation of documents and their properties in a variety of ways.
- ✓ User selectable rules to automate all or part of this work.

## **Automatic learning**

The **AMI E.I. Base Server** kernel automatically identifies and learns the user's vocabulary and rules. This enhances the accuracy of the information search and retrieval processes over a given period of time and use.

This knowledge is stored in a dedicated *Knowledge base* , which can also be manually enhanced.

## **Natural language and Multi-lingual applications**

The **AMI E.I. Base Server** kernel embeds language algorithms for all the major European languages (English, Dutch, French, German, Italian, Portuguese and Spanish). The statistics applications of the **AMI E.I. Base Server** are language independent and can perform analysis of further languages.

## **Access rights of users and groups of users**

The administration application of **AMI E.I.** provides as standard for up to five levels of management and access rights to be applied to users and groups.

### **4.2.2 Key technical features**

#### **Customisable user interface**

**AMI E.I.** application interfaces can be redesigned to be transparently integrated to a website or company portal.

#### **Patent technology**

The **AMI Enterprise Intelligence™** kernel relies on the **Automatic Meaning Interpreter™** (AMI), a fully automated, patented technology designed for the indexing and analysis of information. The underlying concept behind this technology is one of *document signature* .

The performance of AMI surpasses normal Boolean technology. It includes linguistic applications for all major European languages and can be enhanced by a thesaurus or dictionary if required.

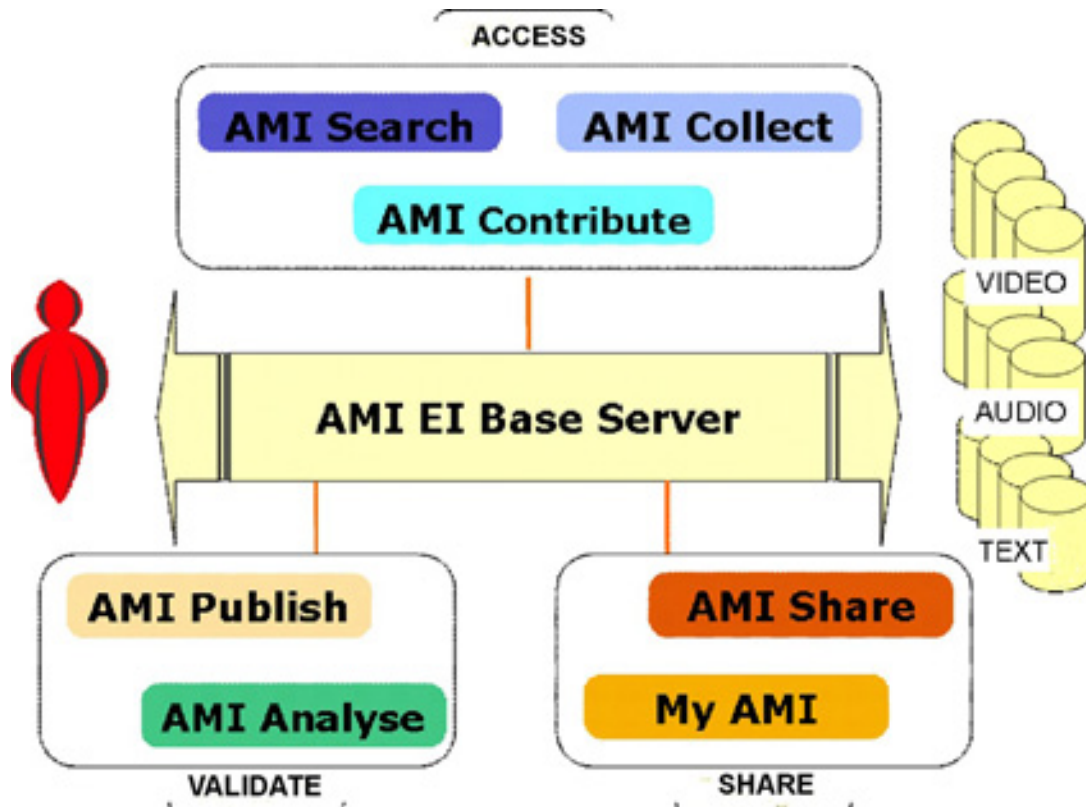
### **Personalised User interfaces**

AMI EI user interfaces can be customised for integration into a company portal, intranet or website

## 5 Functional description

### 5.1 Functional architecture

The following diagram illustrates the functional architecture of AMI Enterprise Intelligence.



The architecture is based on:

- ✓ The AMI E.I. Kernel / Base Server
- ✓ The functional applications of the AMI Enterprise Intelligence solution

Video and Audio processing is delivered via technology integration partnerships

#### 5.1.1 The Automatic Meaning Interpreter Base Server (AMI)

The **AMI** Base Server/kernel is composed of a set of programs which can be accessed through the **AMI E.I.** Web interface. The kernel embeds the core AMI technology:

- ✓ The advanced linguistic functions, especially the signature calculation and search relevance functions.
- ✓ The indexing programs and the Connector and Descriptor capabilities.
- ✓ The albScript language interpreter.
- ✓ These programs can be accessed via a standard interface (http/XML) or via albScript.

The **AMI** kernel is distributed as an independent package called the **AMI E.I. Base Server**.

#### 5.1.1.1 The Searcher

The Searcher is the service in charge of the interface with the corresponding client application. It provides federated search, query and result analysis. It constitutes a single point of access to heterogeneous sources and provides advanced query capabilities.

#### 5.1.1.2 The Indexer

The Indexer is the indexing and retrieving system. It handles everyday language queries.

The *Indexer* and *Synchroniser* are the components in charge of creating and updating the index tables.

- ✓ The *Indexer* is the **active** indexing mechanism. It can cover multiple addresses and update a reverse index.
- ✓ The *Synchroniser* is the **passive** element of AMI. This service is located on an http server. It allows the index to be easily and quickly updated by submitting one or more documents for indexing.

#### 5.1.1.3 albScript

albScript is AMI's script language based on the ECMA-262 (JavaScript) language. It is open (XML) and allows exchange with web services, internal programmes or external libraries, as well as with ActiveX (COM) applications.

The albScript language provides all the functions of AMI which can be used to build any variance of application. It provides the functions necessary to handle AMI objects and their properties.

### 5.1.2 Functional applications

The functional applications and kernel of **AMI Enterprise Intelligence™** communicate via a web services application bus (HTTP/XML).

Each application has its own user interface (PHP/HTML). Particular attention has been paid to the ergonomic design of each interface in terms of the ease with which each function can be identified.

User interfaces can be customised to match the look & feel of specific websites or portals.

The functional applications of the **AMI Enterprise Intelligence™** solution are as follows:

#### **Collect Engine: AMI Collect**

AMI Collect automates the process of information collection from complex arrays of sources both internal and external. It provides the ability to modify the frequency of collection, the definition of the subject to be tracked and the management of information sources. It also ensures the attribution of time stamps and the de-duplication of information.

#### **Search Engine: AMI Search**

AMI Search is the multiple source (internal and external) Search and Retrieve application. It handles everyday language and meta-data queries. It provides extraction of concepts and abstracts and a comprehensive post-ranking, filtering and navigation environment.

#### **Publishing: AMI Publish**

AMI Publish allows the user to select, validate, enrich and organise collected information in order to make it available to a relevant community.

#### **Distribution: AMI Share**

AMI Share is a portal in which the user can search and navigate through shared information in accordance with their pre-defined rights.

#### **Analysis: AMI Analyze**

AMI Analyze provides a variety of tools for the analysis and visualisation of unstructured data based on various criteria such as time, date and source. It also provides trend analysis and concept generation based on Names, Locations, Organisations etc.

#### **Contribute: AMI Contribute**

Functions for including “unpublished” information direct into a classification plan and the knowledgebase by way of creating a memo note or new document on-demand for inclusion in processes of distribution and analysis.

#### **Personalisation: MyAMI**

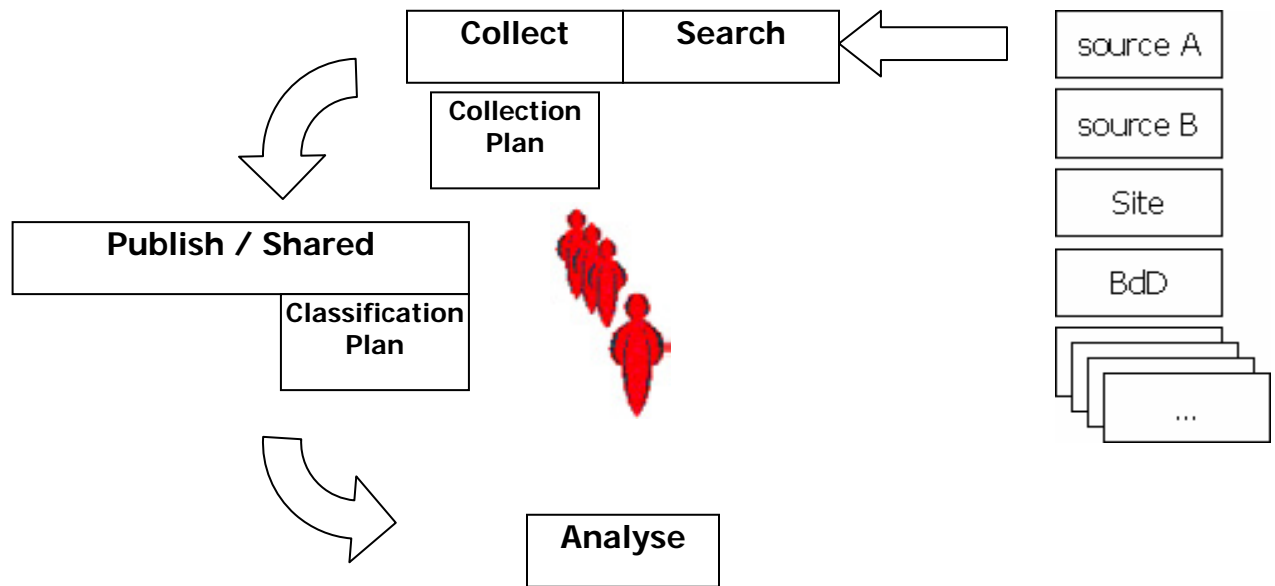
Application for dynamic presentation of information and processes reflecting individual user requirements for specific information.

## 5.2 Example: Market intelligence

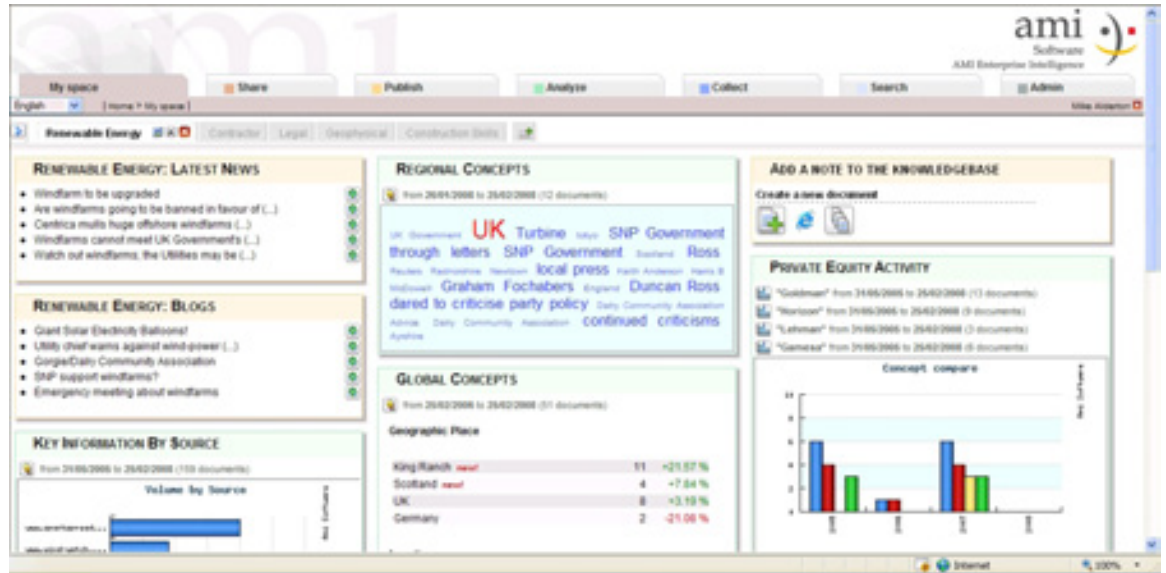
Market Intelligence is a good illustration of AMI Enterprise Intelligence's capabilities.

The Market intelligence process follows three steps:

- ✓ Collect: **collect** and **search** information in compliance with a Collect Map
- ✓ Validate and publish: Information is **validated**, **published** and **shared** among users.
- ✓ Analyze: the information is analysed to help decision making.



## 5.3 The AMI E.I. Homepage



The AMI I.E. homepage is a web interface accessed from any standard browser

After authentication by login ID and password, the first page allows access to all the applications via an “always visible” menu, whose dynamic contents adapt to the rights assigned to the user.

By default, the MySpace page (MyAMI), described in section 9.1, is loaded if this application is installed but any other application can just as easily be set to load by default as chosen by the administrator.

Each menu option gives access to an AMI application, or its functions. These applications and functions are detailed in the following chapters.

## 6 Administration Functions

### 6.1 Physical Installation Overview

Physical Installation of **AMI E.I.** consists of two stages

- ✓ Installation of the AMI Base Server, the technology kernel which provides the functional foundation for all the applications;
- ✓ Installation of the required applications depending on specific requirements: AMI Collect, AMI Search, AMI Publish, Contribute, AMI Analyze, MyAMI and AMI Share.

### 6.2 Administration interface

AMI Enterprise Intelligence has a web based administration interface which allows the user to manage the installation and configuration.

This interface also provides a function to import databases from another AMI applications such as *AMI Market Intelligence* .

The following functions are available from the Administration interface:

- ✓ information sources management;
- ✓ users and group administration;
- ✓ knowledgebase administration.

#### 6.2.1 Management of Information sources

Information sources can be external or internat. External sources can mainly be accessed on the World Wide Web. These sources can be:

- ⇒ Databases
- ⇒ Mailing-lists
- ⇒ RSS Feeds and other information flows.
- ⇒ Forums, Newsgroups, Blogs
- ⇒ External or third party Search Engines
- ⇒ Portals, websites or other information systems

Internal sources are typically intranets, file systems and company data bases. They can contain various types of documents: internal mail messages, customer list etc.

Information sources are managed “globally” and each source can be applied to one, several, or all users. Information sources are made available to the users as a list menu. From a user point of view, each item on the list is a unique information source while, from a technical standpoint, each item can point to several information resources.

### 6.2.1.1 Classification of Sources

There are two types of information source:

- ✓ **information sources indexed by AMI:** in this case, AMI analyses the information source (e.g. a web site) and regularly detects the new or modified documents within that site.
- ✓ **information sources queried by AMI:** AMI references the search function provided by the information source, when available, to collect relevant documents. In this case, the relevant documents are first chosen by the information source, then filtered by AMI and assigned a relevance level.

AMI E.I. provides a single point of access for source interrogation based on the concepts of a **site connector** and a **site descriptor**

#### Tags

The sight by source is often necessary but sometimes constraining. So certain research projects take account of the source (e.g. I seek certain information on such and such website), others can relate to sets of sources in which it is preferable to describe the research terms functionally (e.g. I seek certain information on institutional sites and other information from engineering databases, etc)

AMI E.I. makes it possible to group sources by logical sets, a source being able to belong to one or more sets, the principle being the identification of each source by one or more labels (tags)

#### Source Properties

Online sources have different characteristics related to various web technologies; protocols, conventions, languages, etc. AMI EI can be connected to any type of source and offers an interface which allows the user to specify any number of these parameters for optimal source and site analysis.

#### Site Descriptors

When a source is particularly complex further directives can be defined in more detail through site descriptors.

The mechanism of a site descriptor makes it possible to address questions that other systems cannot address such as access or denial to "pop-up" pages, the way in which Javascript is treated and optimisation of "crawl" processes.

The descriptor also makes it possible to specify the way in which AMI EI operates regarding how links found in web pages are managed (e.g. to follow the links or not, to index the page or not etc).

Descriptors can also be used to specify where, within a page, the useful information is and to extract from it additional information which then becomes descriptive metadata (such as the author of an article for example).

The *topological descriptor* is an important feature of AMI EI making it possible to automatically retain only the most useful part or parts of a page.

The descriptor is an optimisation process. In many cases it is not necessary and in these instances AMI will, by default, index the whole site.

### Connectors

Connectors describe the way in which AMI E.I. connects to other search engines. Connectors provided with AMI E.I. can typically be:

- ✓ Third party search engines such as Yahoo!
- ✓ the AMI engine: AMI Indexer ;
- ✓ the *generic connector* (see section 10.2 below), which gives access to all the sources that can be queried (online forms, languages, protocols...).

In both cases, the AMI Search and AMI Collect applications retrieve information by querying information sources be it through the source's native interface or the one defined by AMI.

Some sources require specific procedures that are defined in advance to capture and index documents which could not otherwise be accessed.

### 6.2.2 Administration of Users and Groups

The typical categories of user are as follows:

- ✓ Supervisor: Manages the entire AMI E.I. environment
- ✓ Administrators: Administrators
- ✓ Group Manager: Manages access rights for groups of users
- ✓ Knowledge Manager: Manages collection and publishing processes
- ✓ Contributors: Users with the rights to add documents
- ✓ Visitors: Users with rights to read documents.

Supervisors can assign groups of users access rights to the main functions of **AMI E.I.** depending on the applications which are implemented. Access rights can be one or several of read/right/update.

Access rights to information can therefore be very precisely defined according to organisational requirements.



Admin UK

Login :	adminuk
Name :	Admin UK
Password :	*****
Group	Administrators
Can log in :	X
Company	AMIEI UK

Module	access	view	add	edit	delete	admin
Collect	X	X	X	X	X	X
Publish	X	X	X	X	X	X
Knowledge Manager	X	X	X	X	X	X
Analyze	X	X	X	X	X	X
Share	X	X	X	X	X	X
Sources	X	X	X	X	X	X
My space	X	X	X	X	X	X
Search	X	X	X	X	X	X
Users & Groups	X	X	0	0	0	0

Module	Feature
Publish	X Manage document rules
Main	X Enable preferences edition

### 6.2.3 Knowledgebase Management

The knowledge base can be enhanced by an automatic learning process. The Administration interface provides several functions to control its content:

- ✓ manage synonyms
- ✓ import external thesauri, when available (external language tools are not a pre-requisite)
- ✓ control the behaviour of the search functions.

Base (17\_analyse\_kdb) > Named Entities > Geographic Place (173)

Search:

Select a letter/digit: **A B C D E F G H I J K L M N O P Q R S T U V W X Y Z**

Named Entities starting with 'N' (11): 1 | 2

<input type="checkbox"/> Naimshire	English	Geographic Place
<input type="checkbox"/> New Caledonia	English	Geographic Place
<input type="checkbox"/> New Zealand	English	Geographic Place
<input type="checkbox"/> Niger	English	Geographic Place
<input type="checkbox"/> Nigeria	English	Geographic Place
<input type="checkbox"/> Norfolk	English	Geographic Place
<input type="checkbox"/> North America	English	Geographic Place
<input type="checkbox"/> Northamptonshire	English	Geographic Place
<input type="checkbox"/> Northumberland	English	Geographic Place
<input type="checkbox"/> Norway	English	Geographic Place

## 7 Accessing Information : Collect, Search, Contribute

### 7.1 Principles of Collect and Search

*The collection* is an automated process which consists of systematically scanning information sources at predefined time intervals.

- ✓ The Collection process is controlled through a Collect Map – or Intelligence Map, which defines the type of operations to be performed by the AMI Collect application
- ✓ The collection of information is performed by the AMI Collect application from which knowledge workers can define the Collect plan.
- ✓ This application automatically performs the following operations:
  - ⇒ the collection of information according to the Collect plan
  - ⇒ detect new information
  - ⇒ de-duplicate documents by analysing their contents
  - ⇒ generate an XML flow containing the results
  - ⇒ send email alerts to subscribers

The XML flow is sent to AMI Publish. Subscribers are defined in the Collect plan.

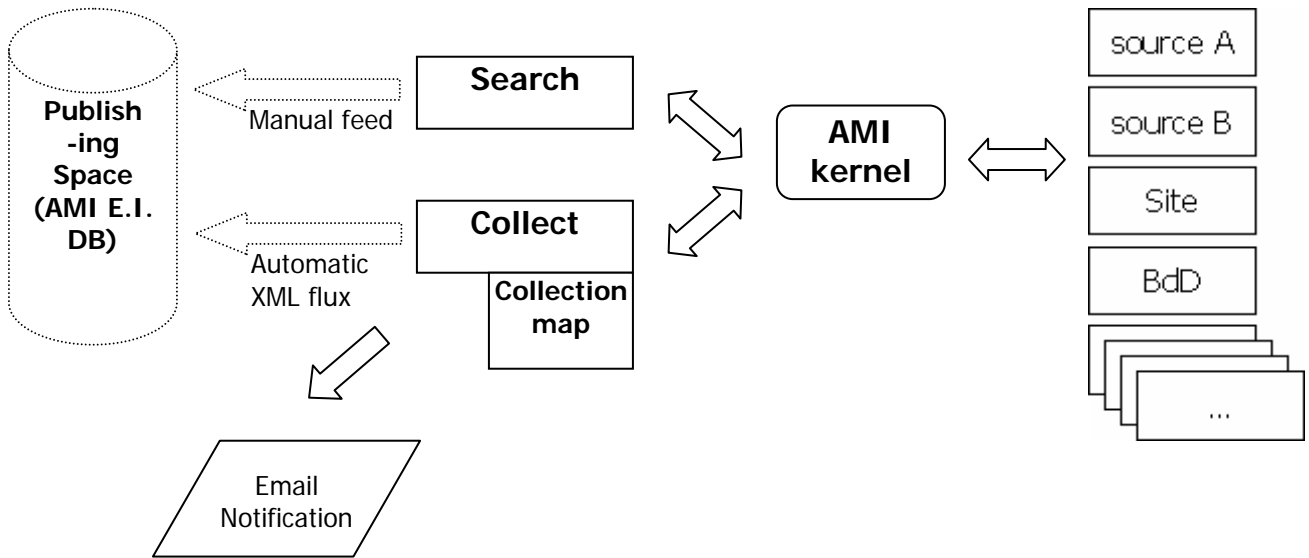
- ✓ The *Search* is performed by users through the **AMI** meta-search engine. It therefore benefits from the kernel's advanced functions.

Search functions can be accessed via the Search application.

The search results are returned to the users as a list of documents enhanced with various kinds of information.

These results, or part of them, can then be selected to populate the Publishing space.

The Collect/Search process:



## 7.2 The "Collect" Application

### 7.2.1 Collect plan

The Collect plan is shown as a hierarchical tree from which user can:

Select the information sources to be automatically interrogated by AMI

- ⇒ select individual sources by name
- ⇒ select groups of sources by tag

for each subject created define scenarios which describe the scanning schedule and the search criteria for retrieving documents.

The Collect plan is hierarchical comprising a minimum of three levels:

The **Subscription** level

- ⇒ E-mail alerts are defined at this level
- ⇒ A subscription can aggregate the results of several Subjects (see below) according to organisational requirements.

The **Subject** level

- ⇒ A Subject can aggregate several scenarios (see below) relating to a same subject. The scanning schedule is defined at this level (i.e. the frequency of collection - daily, weekly, monthly or personalised) on each of the Scenarios

- ✓ The **Scenario** level

- ⇒ A scenario defines a search query for a specific information source or sources as explained below.

It is important to note that the Collect Plan is global to the application. The structure of the Collect Plan is standardised by AMI in the form of DTD (Document Type Definition) to standard XML. The Collect Plan can be independent of how information is actually presented. The distribution and sharing of information is managed via a further hierarchical folder structure called the Classification Plan.

## 7.2.2 Defining scenarios

Defining a scenario consist of the following actions:

- ✓ Selection the information sources to be queried from the ones which have been made available by the administrator (see below)
- ✓ this operation is carried out by informing a list starting from the list of the sources available and their tags.
- ✓ Specifying the activity of this Scenario (only active Scenarios).
- ✓ Optional, to formulate a request.

The frequency of collection of the Scenario is not specified in the parameter settings of the Scenario itself but in the Subject to which it belongs (see above).

The formulation of the search request is carried out by defining a strict “must contain” constraint, a lower “similar to” constraint, or a combination of both:

- ✓ a strict constraint (Mandatory or “Must Contain”) is a search term made using Boolean or natural language. It is particularly useful when the user already has some knowledge of the specific terms relating their subject of interest. The name of a company could, for example, constitute a strict constraint.

AMI’s Boolean language capability includes/understands the instructions AND, OR and NOT, as well as quotation marks, parenthesis and truncation. It can add to a request an “AMI” operator (the question mark) which instructs the system to generate variations of the same term referencing first at any existing knowledgebase

A simple example:

AMI OR "AMI Software"

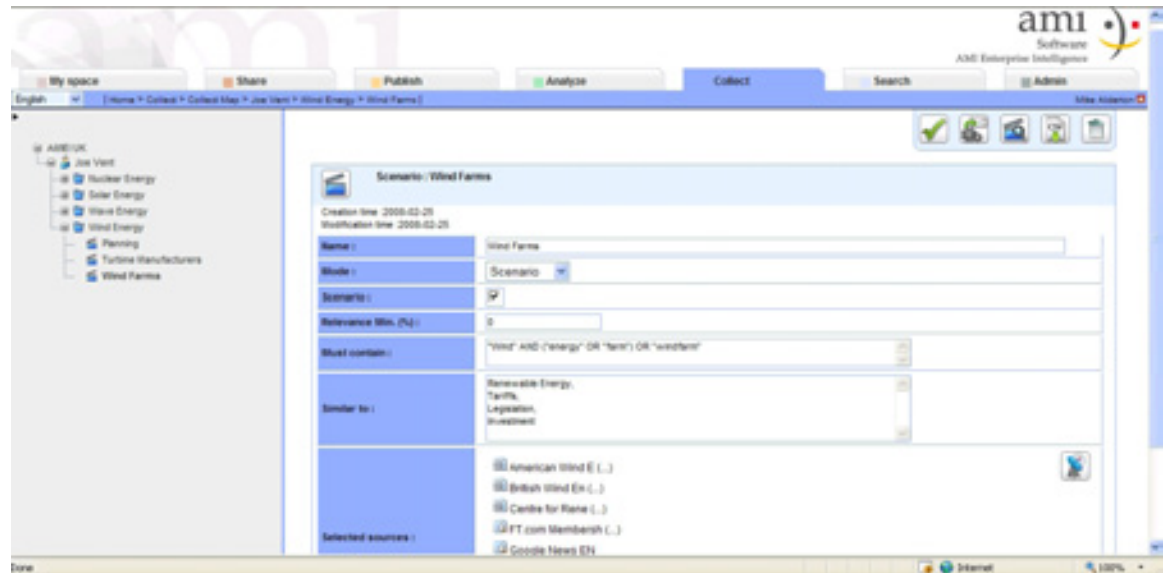
A more complex example:

("Spondylarthritis ankylosante" AND AINS?) NOT (psoriasis OR anorex\*)

- ✓ A weak search term constraint (semantic or “similar to”) is also a request. Permissive (as opposed to restrictive) it can consist of a entire paragraph. AMI will seek similar information. A weak constraint is useful when the user wishes to gather information that is related to their subject of interest.

A weak constraint can also be used to describe the "environment" (i.e. a subset of the information reached through the available sources) of the search. For example, the collection of information about a very common term can be narrowed by means of a weak condition which describes the "environment" of the subject:

A scenario example:



In this example, the strong and weak conditions are described in the "Must contain" and "Similar to" fields, respectively.

### 7.2.3 Execution of the Collect Process

The Collect plan is processed as scheduled in the Topics.

Each scenario is performed by analysing the queries it contains and enhancing them with the semantic analysis provided by the **AMI** kernel (see the section entitled "Enhanced function"), and the application knowledge base. .

The kernel generates *Search hypothesis*, which are derived from the user's specified query. This mechanism enhances the quality of the search which is performed and the relevance of the results. The information sources are then scanned via the corresponding connectors to discover relevant documents.

The results are then processed by the **AMI** kernel to build the output flow. This flow is then used to:

- ⇒ Populate the publish space, the application "AMI Publish", when this last is activated generate alerts to subscribers.
- ✓ It is reprocessed by an automat of the AMI core to manufacture e-mail notification to the Subscribers.

#### 7.2.4 Email Notification

Email alerts are composed of a list of *news* composed of the following information:

- ✓ Address (or URL) of the original document
- ✓ URL of the information source
- ✓ Author's name -when this information can be automatically extracted
- ✓ Key expressions
- ✓ List of the terms which matched the ones contained in the condition. These terms may be highlighted.
- ✓ Name of the scenario(s) which collected the document
- ✓ Meta-data associated with the scenario
- ✓ Concepts automatically extracted from the set of answers

The underlying concepts are detailed in the section entitled Advanced functions.

Example alert email

**Carbon Exchange News** Carbon Exchange

---

**Carbon Exchange Activity Monitor** carbon credit,carbon dioxide,carbon sequestration

**1. Carbon Credit Craze** 97 %

Document source: [www.pigswillfly.com.au](http://www.pigswillfly.com.au) (English Blogs)

Main quotations:

- "Origin promises to verify business' carbon credit programs, that is make sure money spent on carbon credits actually goes to projects which will negate their carbon emissions. Unfortunately Origin hasn't yet announced how exactly it will 'check out' or verify carbon credit schemes. So far in Australia there isn't a government or private body that properly verifies carbon credit spending so businesses who want to enter a carbon credit program do so blindly. ... [Exchange](#) ... [Soil](#) ... [Offset](#) ...

*Author: this information could not be extracted automatically but may be available in the document* [Download...](#)

**2. Carbon Credit Payments for US Forest, No-till Crop, Manure, or Grassland** 96 %

Document source: [transectpoints.blogspot.com](http://transectpoints.blogspot.com) (English Blogs)

Main quotations:

- "The North Dakota Farmers Union has successfully started a Carbon Credit Program. The North Dakota Farmer's union started the carbon-credit program and technically, is the official aggregator - the entity that collects credits and puts them on CCX for sale. To learn more about the carbon credit program, visit the National Farmers Union Web site: [www.nfu.org](http://www.nfu.org) and click on the environment tab, or call the Wisconsin Farmers Union at 715-723-5581 or visit it on the Internet at [www.wisconsinfarmersunion.org](http://www.wisconsinfarmersunion.org). ... [Exchange](#) ... [Soil](#) ... [Offset](#) ...

*Author: this information could not be extracted automatically but may be available in the document* [Download...](#)

**3. Major TV Network Program -The Great Carbon Credit Swindle** 94 %

Document source: [agmates.blogspot.com](http://agmates.blogspot.com) (English Blogs)

Main quotations:

- "From: Steve Truman To: name removed by agmates Sent: Wednesday, 28 March, 2007 5:22:39 PM Subject: Re: The Great Carbon Credit Swindle G'day (name removed) Thanks for your email, The Blog article was a little disjointed, so I've done a chronological summary to make it a little easier for you. these are Soil carbon credit traded on the Chicago Climate Exchange. - March (14) - Major TV Network Program -The Great Carbon Credit. ... [Offset](#) ...

*Author: this information could not be extracted automatically but may be available in the document* [Download...](#)

duplicate (55%): [News about John McCain](#)

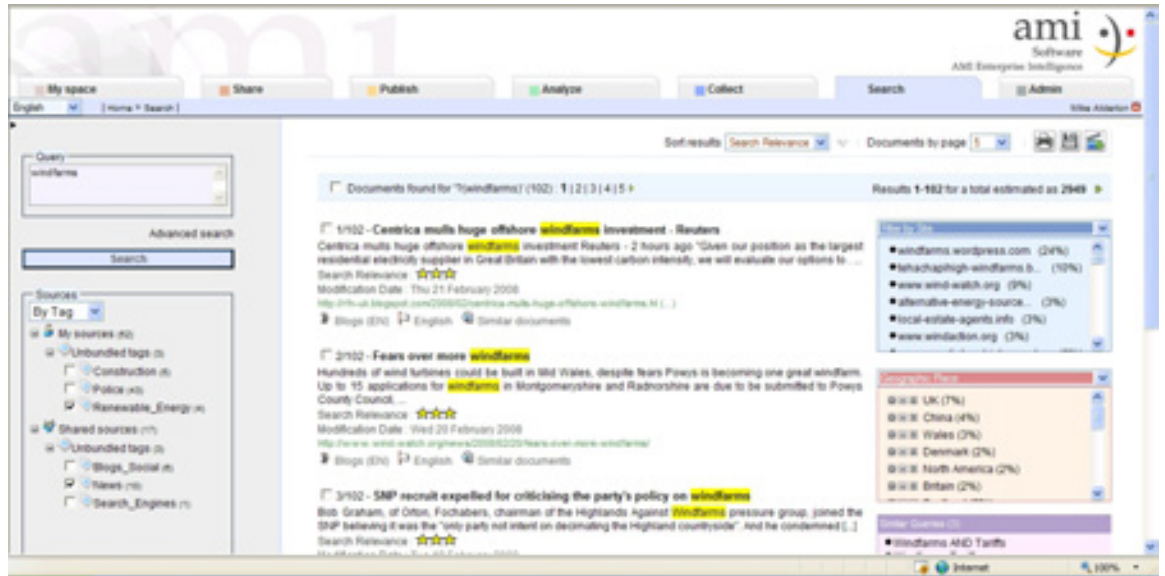
**4. Grains Council and the AGO data gap** 92 %

Document source: [carboncredits.blogspot.com](http://carboncredits.blogspot.com) (English Blogs)

## 7.3 The "Search" application

### User interface

The search application allows users to carry out search requests across both internal and external sources. Sources can be selected either individually by name or by group tag.



### 7.3.1 Conducting a Search

Both Boolean and everyday language queries are supported.

The information sources which are queried can vary (internal, external, web, content management systems...) but are seen in a consistent and unified way via the application's interface. The list of sources on the left hand of the user interface shows:

- ✓ Shared sources
- ✓ Private sources

The queries performed on these sources and the results returned benefit from **AMI's** advanced functional capabilities.

A fully comprehensive range of pre-search, post-filtering and post-ranking functions are automatically available within AMI Search

### 7.3.2 Search results and manual population of the Publishing space

The search results are composed of a list of documents with the following information:

- ✓ the document title
- ✓ the source where the document was found
- ✓ an abstract of the document
- ✓ a relevance level
- ✓ the words and expressions, from the query or from a thesaurus, on which the document was found and selected.

Many other functions can be automatically called by the user such as similarity search, automatic classification, clustering or identification of related themes.

## Memorising a Search

A Search request can be saved and converted into a scenario request within AMI Collect.

## Detection of themes, topics and named entities

AMI E.I. analyses all of the results of a request and identifies the most important key words and expressions. These expressions make it possible to extend the original request of the user. They are characterised by the type of information which they represent including Names, Locations, Organisations and General Concepts.

## Advanced Search

An “Advanced Search” screen allows search terms and parameters to be more fully defined.



## 7.4 “Contribute” Application

The “contribute” application makes it possible to quickly record “unpublished” information by creating a memo note or new document from the desktop for automatic submission to the knowledgebase and inclusion in any subsequent process of distribution and analysis.



Unpublished information could be information learned from a meeting, seminar or conversation. Using the “I Contribute” function the value of such information can be readily captured and fully capitalised upon.

The “Contribute” application allows new information to be quickly added in three ways

- ✓ By the user creating new content themselves;
- ✓ By including an existing file.
- ✓ By giving a URL that points to a specific document

The contents of this information can be directly published or await validation as with any other collected information in the publish workflow.

## 8 Developing Information: Publish, Analyse

### 8.1 Main principles of Publish and Analysis

Publishing and Navigation are provided by **AMI E.I.** and permit a collaborative Publishing space which is organised as a classification plan.

The publishing space can be populated in three different ways:

- ✓ automatically via the "Collect" application
- ✓ manually via the "Search" application
- ✓ by adding documents manually via the "Contribute" function

The file classification plan can be managed via the "Publish" application and allows the user to classify and share documents contained in the Publishing space.

Publishing consists of assigning a status to each document. This action is performed from within the Publish application.

The principle of content analysis is dealt with by "Analyze" application, details of which are provided in section 7.2.7 onwards.

- ✓ Determination of a the extent of an "information perimeter" via the choice of a section in the classification plan, by filtering by expressions over a period of time and/or a geographical distribution (when integrated with a third party application such as ESRI Geographical Information Systems)
- ✓ Use of statistical or text-mining tools.

### 8.2 The "Publish" application

The Publish application includes functions to create a new document from url or file, provides workflow for reviewing and approving documents collected both manually and automatically, creating a classification plan, assigning document rights, and managing document movements.

#### 8.2.1 Function: "New Document"

These functions make it possible for users to contribute to the process of information acquisition and distribution by submitting reports, minutes of meetings, notes, or other documents found on line, etc

Existing documents can be added via the "New document from URL or File" function whilst newly documents can be created and added via the "New document" function.

Add a new document from url or specified location

**EDIT DOCUMENTS**

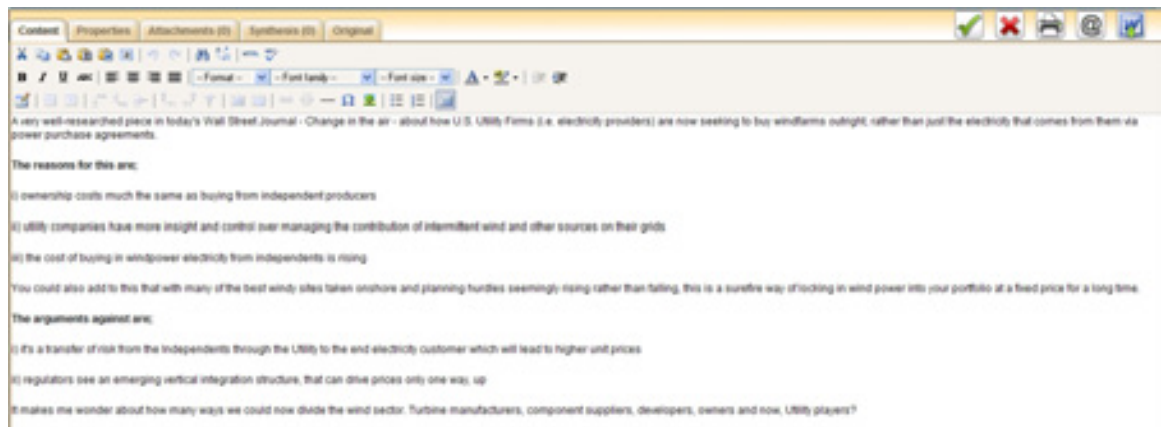
Create a new document from URL or File

Section:  M&A Knowledgebase

Url:

File:  Browse...

Editing a new document



### 8.2.2 Function: "Document Validation" (Publication)

The Validate document function allows the user to publish documents, i.e. assign a status to the document as follows:

- ⇒ Collected
- ⇒ In Progress
- ⇒ Waiting for Validation
- ⇒ Validated
- ⇒ Published
- ⇒ Refused

Documents with a publish status can only be seen from the "view" mode if the user has the necessary access rights to the Category in which the document appears.

- ✓ View the original document at source.
- ✓ *Review and/or edit the Image text of the document.* This Image text is extracted automatically from the original which may be a document file or an audio or video sequence. An automatic process of cleaning is applied to the process making it possible to exclude irrelevant information or “noise” (advertising, banners etc). It can be published and formatted to improve the presentation and/or to insert images or add a “comment in the text” of the document. Content storage is done in xhtml and the structure of the page is preserved.
- ✓ *Add a comment to the document:* An added comment is a property of the document: it does not form part of the text of the document itself. It can coexist with the “comments in the text” described above.
- ✓ Review and/or publish the document properties such as the Date, Relevance level, Title etc.
- ✓ Add attachments to the document (in the same way as with e-mail)
- ✓ Automatically create newsletters via export in Office Word format

A contextual menu (shown as a blue triangle) makes it possible to carry out transverse actions on collected documents. It enables actions such as: “validate documents by heading”, “publish subjects by heading”, “review documents by heading” etc.

The Publish application also allows requests to be applied to all documents contained within the classification plan if and where required.

### **Help with reading the results**

Comprehensive functions designed to help with the reading and interpretation of outputs from AMI Analyze readily highlight sections from a text that meet or answer certain criteria. Automatic selection of the most significant expressions makes it possible to quickly identify the overall meaning of an article or document that may be many pages in length.

### **Search engine**

The search function within AMI Publish makes it possible to search across all documents contained within the classification plan. The search can be simple or be advanced, the advanced mode allowing selection by all properties associated with the documents.

## Summary View

The screenshot shows the AMI Publish Summary View. The interface includes a top navigation bar with 'My space', 'Share', 'Publish', 'Analyze', 'Collect', 'Search', and 'Admin'. A breadcrumb trail reads: 'Home > Publish > Validate Documents > Renewable Energy > UK and Ireland > Newswatch Blogs'. On the left, a 'Classification' tree is visible with categories like 'Classification (1887)', 'All UK (76)', 'Reviews', 'Construction', 'Construction Skills (1538)', 'Contractor', 'Energy & Utility', 'Investment Analysis', 'Legal', and 'Postal Services (48)'. Below this is a 'Query' field and a 'Search' button. The main content area is titled 'Full Newsaries of' and shows a list of documents. The first document is '1 - Emergency meeting about windfarms', published on 21 February 2008 by Mike Alerton. The second document is '2 - SNP support windfarms?', also published on 21 February 2008 by Mike Alerton. The interface also shows 'Type: Articles', 'Status: Created', and 'Sort by: Date of publication'.

## View by title

The screenshot shows the AMI Publish View by title interface. The top navigation bar and breadcrumb trail are the same as in the Summary View. The left sidebar is identical. The main content area is titled 'Newsaries of' and displays a table of articles. The table has columns for 'Reference', 'Section', 'Title', 'Created on', and 'Published on'. The data rows are as follows:

Reference	Section	Title	Created on	Published on
7 7 7	Newswatch: Mainstream	Fears over more windfarms	21/02/2008	ndash
7 7 7	Newswatch: Mainstream	SNP recruit expelled for criticising the party's policy [...]	21/02/2008	ndash
7 7 7	Newswatch: Mainstream	Wish out windfarms, the Utilities may be coming to bay [...]	21/02/2008	ndash
7 7 7	Newswatch: Mainstream	Windfarms cannot meet UK Government's renewable energy [...]	21/02/2008	ndash
7 7 7	Newswatch: Mainstream	Centrica mulls huge offshore windfarms investment	21/02/2008	ndash
7 7 7	Newswatch: Mainstream	Are windfarms going to be banned in favour of nuclear [...]	21/02/2008	ndash
7 7 7	Newswatch: Mainstream	Windfarm to be upgraded	21/02/2008	ndash

Below the table, there is a section titled 'FEARS OVER MORE WINDFARMS' with a 'In summary' link. The 'Origin of the document (URL):' is 'http://www.wind-watch.org/news/2008/02/20/fears-over-more-windfarms/'. The 'Section:' is 'Newswatch: Mainstream' and the 'Author:' is 'Mike Alerton'.

### 8.2.3 Synthesis

AMI Publish allows the creation and management of document "syntheses" which themselves reference other documents in the classification plan. A synthesis can combine sets of documents and also include elements from other applications such as the results of analysis and search.

A synthesis can, like any other document, be exported in Word format to create a customised newsletters.

#### **8.2.4 Function “Search Documents”**

This function allows the user to search the Publishing space using Boolean or natural language queries. Document properties can be sorted on or selected.

##### **Advanced Search**

Advanced Search allows a finer selection of the documents: seek multi-heading, language running or Boolean, control of the generation of assumptions, choices of dates are the principal criteria.

#### **8.2.5 Function: Classification Plan**

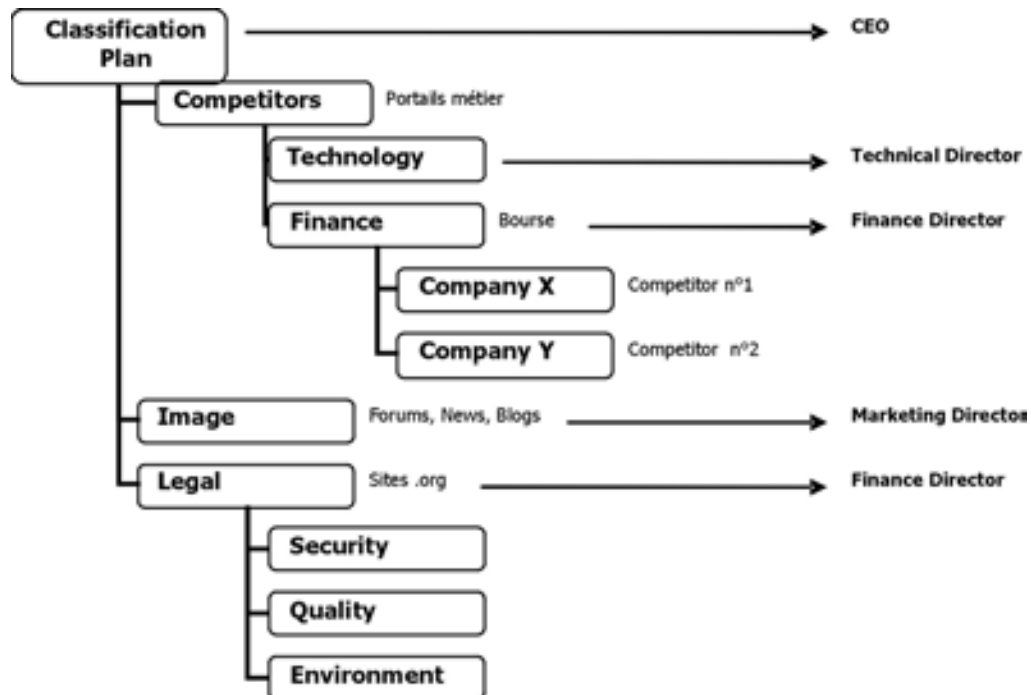
The Classification plan is a tree structure of Headings. The documents published or on standby for publication are attached to one or more headings of the classification plan.

Each Heading has:

- ✓ A position in the directory specified by its parent header
- ✓ A list of reference Subjects, extracts of the collect plan
- ✓ A list of Users rights (read/right access) within in the plan

Like the Collect plan, the Classification plan is global within the application.

Example of a classification plan.



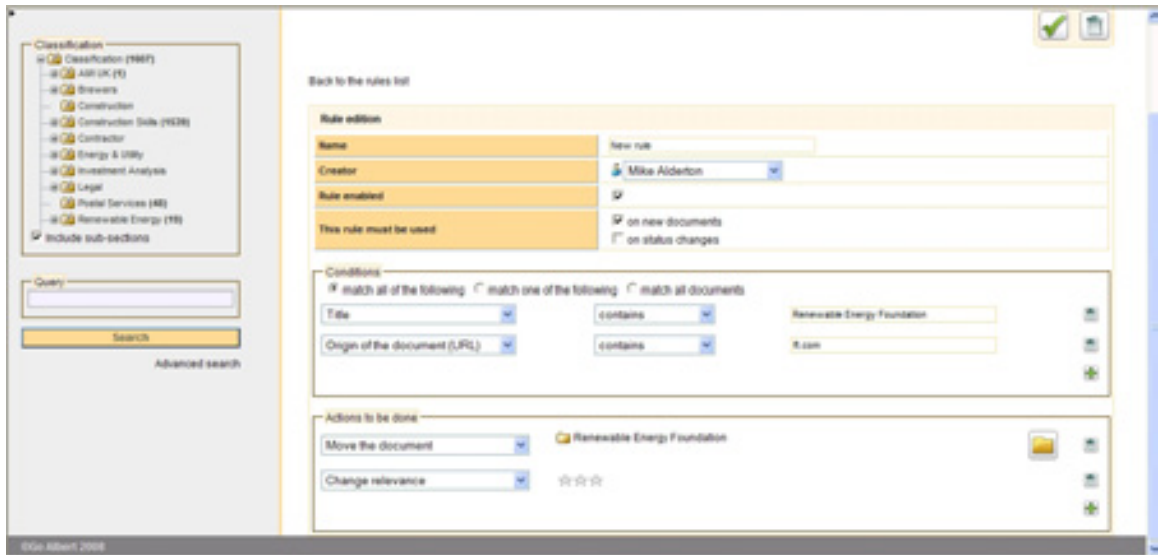
### 8.2.6 Function: Document Rules

This function is an available option within the “Publish” application. It is driven by a rules engine designed to optimise the management of collected documents. Documents which meet particular “conditions” can be subjected to one or a number of specific “actions”.

“conditions” relate to the status of the document such as its properties (level of proof, relevance), its lifecycle (creation and/or expiry date) and the metadata associated to it.

“actions” allow intervention on the above mentioned conditions or the classification of a document making it possible to automatically indicate the level of relevance or proof, the importance, the date of publication, whether to move a document to a certain location or delete it, etc.

Below: An example of a rule executed based on document title and source with subsequent actions defined for automatic relevance rating and classification if the first conditions are met.

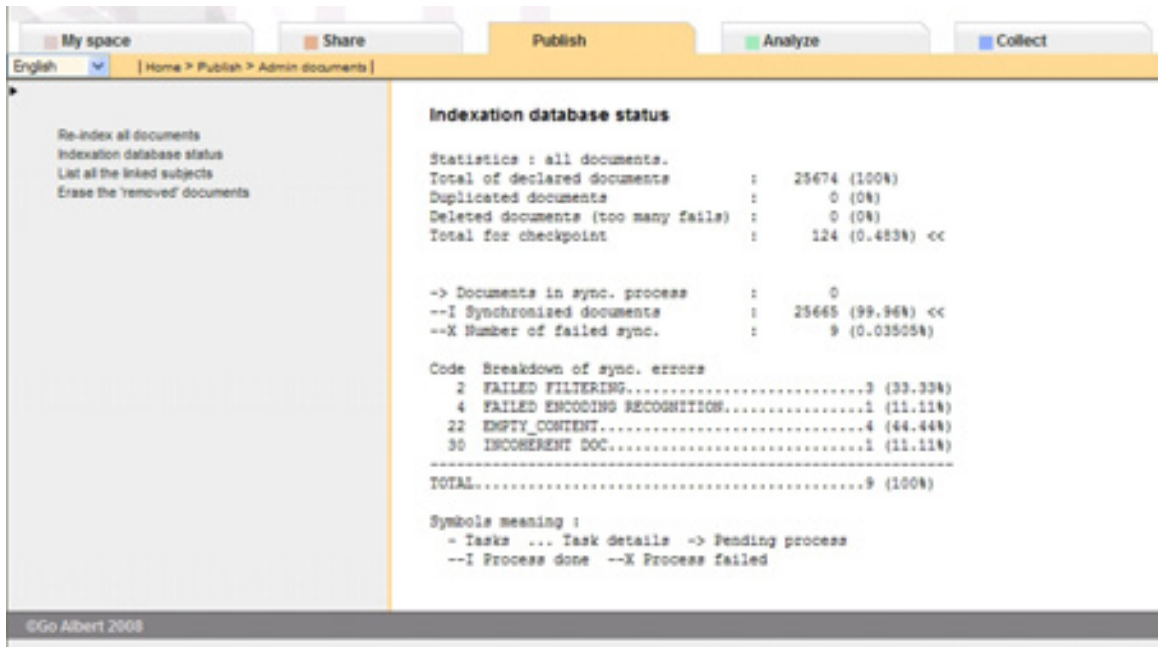


### 8.2.7 Function: Document Administration





























This function integrates three processes: “Re-index all the documents”, “Indexing status” and “List of all the objects referenced”.

The first makes it possible to erase the knowledgebase index of shared documents and re-index all the documents present in the knowledgebase.

The second presents a detailed report as to the correct operation of the indexing process. The process takes account of and utilises every available statistic.



"List all the linked subjects" presents all the subjects covered in the knowledgebase indicating for each one whether the subject is active and which section of the classification plan it is related to.

List all the linked subjects			
Enabled	Subject	Section	
X	AMI EI v3.0 : 350 Grade Steel	350-grade Steel	   
X	AMI EI v3.0 : Bank Profits and Opinion	Bank Profits	   
X	Mike Alderton : Barclays and A&N Amro	Merger	   
X	Prof. Arun Holden : Computational Biology	Computational Biology	   
X	Mike Alderton : Geophysical Services	Geophysical Published Source	   
X	Carbon Exchange : Carbon Exchange News	Carbon Exchange	   
X	Bus Operators : Bus Operator Research	Bus Operators	   

### 8.3 The "Analyze" application

The Analyze application allows the user to extract statistical information and semantics from of group of documents, called the "perimeter", as defined by the user.

The Analyze application allows the following:

- ✓ Define the Analysis criteria across a particular category from within the File plan
- ✓ Submit a search request analysing only the documents which match it
- ✓ The status of documents concerned
- ✓ All document properties (relevance, level of confidence, etc.).
- ✓ Analysis of a batch of documents over a period of time between pre-selected dates
- ✓ A restriction in the documents Analysed limited to a specific geographical selection in cases where AMI is integrated with a Geographical Information System.

All results from any process of analysis undertaken can be recorded and exported in xml.

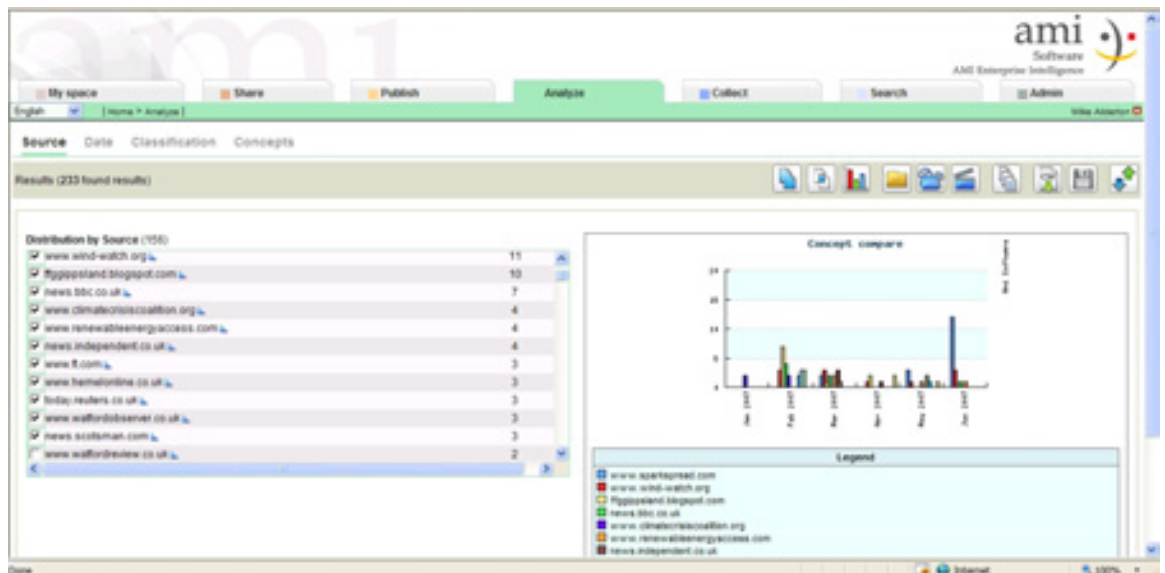
#### 8.3.1 Measuring contribution by source

The "Source" menu item shows the source and number of documents per source.



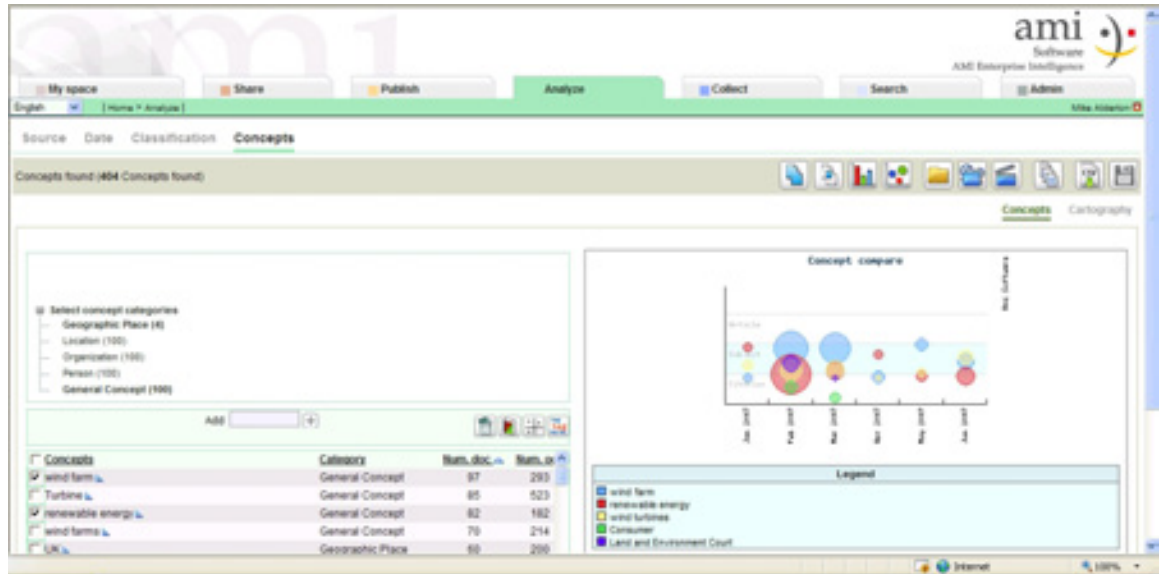
The source analysis function allows the user to carry out comparative analysis across wide ranges of sources. It makes it possible to identify which are the most useful sources, review the links between documents, carry out comparisons of document volume collected over time or comparisons by classification plan, scenario or subject. The objective is to support transversal actions between the “Analyze”, “Collect” and “Publish” applications.

Illustration of Transversal Analysis:



### 8.3.2 "Volume and Centering"

Example of the Time/Date analysis with Volume and Centering



Volume shows the number of documents collected against a scale of time. Centering is a measure of representation and indicates whether the document is an article about a subject, whether a particular subject is mentioned by name or whether a citation is included in the text.

The "view change" button makes it possible to carry out analysis based on document creation date, identify related documents and compare their distribution by classification plan, scenario or subject.

### 8.3.3 Transversal Analysis of the Classification Plan

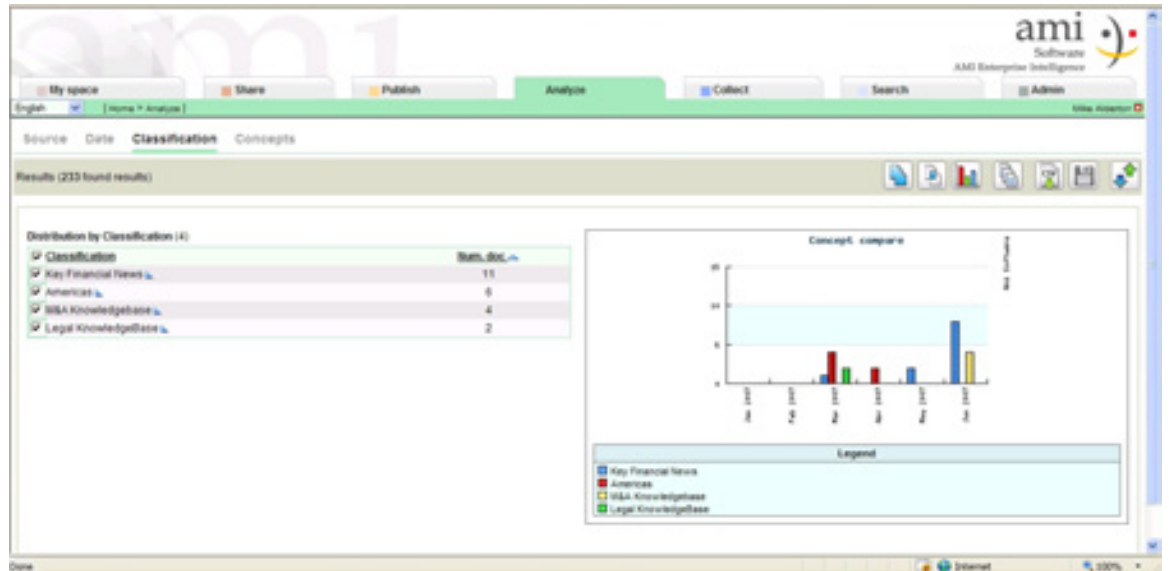
Transversal Analysis makes it possible to study the volume of documents processed relating to the respective Classification Plan headings.

The "view change" button into list mode makes it possible to carry out comparative analysis across the classification plan folders identifying document links and intersections and a comparison of volume against time

### 8.3.4 Analysis Tools: “Concepts” (detection of tendencies)

AMI automatically extracts entities, i.e. the principal correlated concepts (terms or expressions). Some are typified by specific subject such as People, Organisation, and Location, others are extracted and presented by criteria of importance under General Concepts.

Multiple entities can be selected and their relevance and relationship to each other graphically displayed.



Visualisation as a “Concept Cloud” is also a feature.

The Analyze application has a dedicated knowledgebase used in the detection of concepts. The knowledgebase features synonym management making it possible to define that one expression is synonymous with another (for example “Lawyer” is the same as “Attorney”). It is also possible to block a concept or to change its category.

As with other stated analysis functions it is possible to review the links between documents, compare concepts by volume, date, time and by level of centrality, or study their distribution by classification plan, scenario or subject.

### 8.3.5 Analysis Tools “Ontologies”

One or more ontologies can be imported into the AMI Enterprise Intelligence v4.0. knowledgebase. The contents of these ontologies can be highlighted through the AMI Analyze application.

For example, if the company implementing AMI EI has modeled their Organisation (branches, offices, factories and branches etc) in the form of RDF diagrams then documents collected can be broken down according to the Organisation’s structure. In the same way browsing within the classifications of various entities of the group provides access to related documents.



### 8.3.6 Analysis Tools “Cartography”

To allow the User or Analyst to spot early indicators of change AMI provides functionality for identifying “weak signals”.

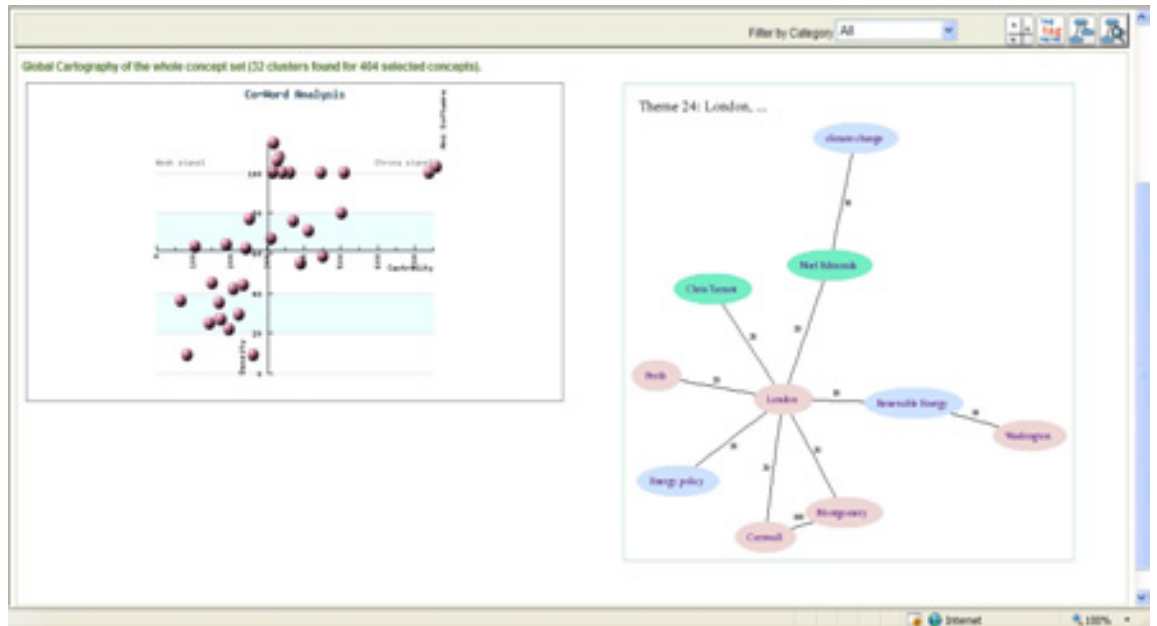
New events are detected by clusterisation of extracted concepts. Clusterisation is based on the co-occurrence of concepts in the same documents. Thus, concepts that recur in the same documents will be identified within the same cluster (an internal link), other concepts of lesser co-occurrence will be identified in different clusters and identified through an external link.

## Identification of Strong and Weak signals

AMI Analyze will identify and visualise Weak and Strong signals. Weak signals indicate something new and can be the “early warning” of a new event or the start of a new trend.

Similarly a Strong signal indicates a significant amount of related information on a particular topic within the knowledgebase

Example results shown in the Cartography menu:



All points on all graphs and charts are clickable allowing fast navigating directly to the related document within AMI and subsequent navigation to the original at source.



## 9 Sharing and Distributing Information

### 9.1 MyAMI Application



MyAMI provides a “global” view of all information contained within AMI E.I. to which the user has rights to view. In the simplest of cases it gives the end-user a universal view of all documents available to them whether they are published or not.

It makes it possible to organise, in one page, all key information presented in dedicated functional “blocks”. In addition to presenting information, functions accessible from within each “block” can be used to navigate within the system and perform various actions such as “Search” and “Contribute”

Those tasked with the process of organising information, e.g. collecting, validating, publishing or distributing it can use MyAMI to design highly accurate and informative decision management dashboards either for their own use or for use by their “customers”.

## Dashboard Menu

**Shown below:** Functions available from within MyAMI allowing the presentation and personalisation of information based very precisely on individual user requirements and areas of interest.

Information can be referenced from any of the AMI E.I. applications namely AMI Base Server (Sources and Indexing processes), AMI Publish (Documents, Published or Unpublished) AMI Analyze (Concept extraction and statistics), AMI Share (Shared documents available to the user), AMI Search (Search Engine, Internal and External sources).



An example of a personalised dashboard created using MyAMI with information referenced from AMI Publish and AMI Analyze highlighting key information for the decision maker.



Multiple dashboards reflecting varying user roles and requirements can be created managed and accessed easily via tabs.

Information presented in MyAMI can be exported in Microsoft Word<sup>1</sup> format allowing the creation of personalised newsletters etc.

<sup>1</sup>: Requires Microsoft Word version 2003 or higher or Open Office.

## 9.2 “Share” Application

The Share application, AMI Share, is an integrated portal that allows documents to be distributed and shared with users across an Organisation.

The primary use for AMI Share is the distribution and accessibility of the most relevant information to as broad an audience as is required. In this sense, it is a “public” application.

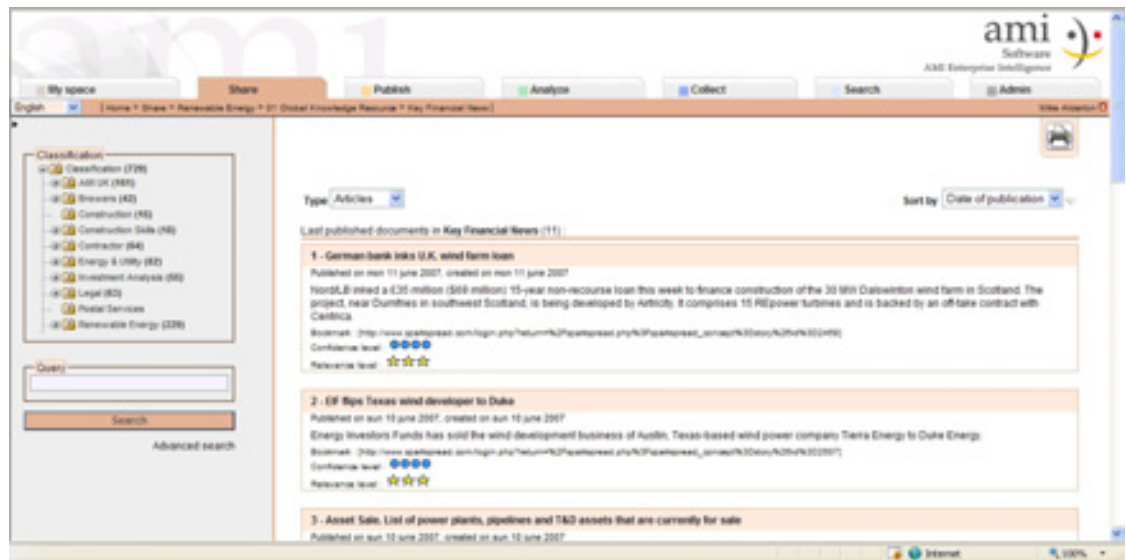
## 9.2.1 Viewing Documents

AMI Share makes available to the user the following:

- ✓ Categories the user has access to presented in a folder/tree structure.
- ✓ List of documents in the Publishing space or a selected part of the folder structure

For each document in the list, the user has an advanced range of information functions also available provided by AMI's core technology e.g. quotations, keywords, hyperlinks etc, and can reference the document either as text or as an original at source, (html, pdf, audio/video etc)

Example view screen in AMI Share




## Clustering

AMI Share features powerful clustering functionality automatically suggesting similar documents and documents by the same author.

## Search

AMI Share features both a simple and an advanced search facility.

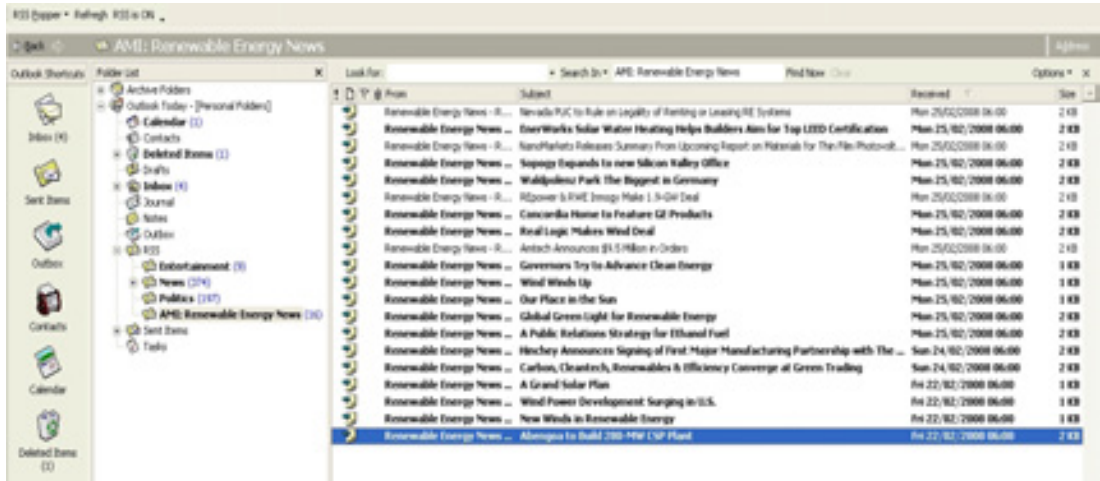
## 9.2.2 Generation of RSS feeds

New feed	
Name of the feed	Renewable Energy
User of the feed	Mike Alderton
Share policy	<input checked="" type="radio"/> Private feed (user described above only) <input type="radio"/> Shared feed for all registered users <input type="radio"/> Public feed (no authentication)
Type of the feed	<input type="radio"/> Access to "Share" module <input type="radio"/> Access to "Publish" module <input checked="" type="radio"/> Direct access to the websites the documents are from
Section to be syndicated	Classification 
Include sub-sections	<input type="checkbox"/>
Status of the documents	Published
Content of the feed	<input type="radio"/> Entire content of the document <input checked="" type="radio"/> Description of the document <input type="radio"/> No content (links only)
Shown date	<input checked="" type="radio"/> Creation date of the document <input type="radio"/> Publication date
Maximum number of documents in the feed	100

From AMI Share a user with assigned rights can create RSS feeds. A feed is associated with a category of the classification plan and can be customised according to the following parameters:

- ✓ Feed name
- ✓ Name of the person creating the feed
- ✓ Access rights: Whether personal or shared, password protected or public
- ✓ Feed type, i.e. value of the tag <link>:
  - ⇒ Pointing to the document in AMI Share
  - ⇒ Pointing to the original document, the source
- ✓ Document status (depending on type)
- ✓ Heading to amalgamate: can be a single folder, or entire folder classification
- ✓ Feed Content
  - ⇒ Entire content of the documents
  - ⇒ Document description
  - ⇒ Do not include content (links only)
- ✓ Date created, i.e. value of the tag <pubDate>:
  - ⇒ Document creation date
  - ⇒ Publication date
- ✓ Maximum number of documents to generate (from 10 to 100)

The application also provides access to the feed, the URL of which can also be saved in any RSS reader such as browser or e-mail client



## 10 Advanced functions

### 10.1 Advanced linguistic functions

The advanced linguistic functions of **AMI E.I.** enhance the efficiency of information retrieval in all the languages covered.

These functions rely on the patent g-MIL (Generator of Markers Independent of Language) technology which is at the heart of the **AMI** signature computation.

#### 10.1.1 Signature concept

The principle signature functions are:

- ✓ Compute, handle and store numeric information which represents the actual *meaning* of a text.
- ✓ Support a large variety of signatures ranging from a simple list of terms to a rich and complex set of hierarchically organised information related to the user's context.

Comparing the signatures of two documents allows the system to measure the "distance" between meanings conveyed between two texts. The more complex the signatures, the more accurate the distance measure. The distance between a simple and a complex signature can also be measured in which case the system grants a lower confidence indicator to the measure.

#### 10.1.2 Relevance Calculation

The *relevance level* is a score – between 0 and 100 – which reflects the value of a document with regards to the user's query.

The score is obtained by comparing the signatures of the query and the one of the document.

From the "Collect" application, the user can specify a *relevance threshold*.

#### 10.1.3 Identification of the Query language

Different algorithms are used to identify the user's language. They are based on the languages morphology as well as the user's query history. A recognition confidence score is assigned. Recognising the language in which the query is submitted is necessary to generate adapted search hypotheses.

#### 10.1.4 Spelling approximation

The **AMI** kernel also caters for approximate spelling and spelling variations. Spelling and phonetic variations are added to the hypotheses generated by the system to collect information. This function allows the system to collect information from sources that may be very relevant but contain a high ratio of spelling mistakes, such as open forums.

#### 10.1.5 Synonym management

The **AMI** kernel handles multi-word synonyms and acronyms. It is also possible to assign a level of confidence to the relationship of synonymy. These synonyms are used to generate different hypotheses from the user's query.

#### 10.1.6 Stop word management

A list of *stop words* - i.e. words which are not indexed – can be implemented for each supported language.

In addition, regardless of the document language, information regarding the importance of terms is gathered during the indexing process to optimise the relevance of the results.

#### 10.1.7 Filtering out outdated documents

A specific feature of AMI I.E. is the evaluation of how “fresh” a piece of information is. Each document is evaluated so that only information which is new to the user is displayed.

#### 10.1.8 De-duplication

The **AMI** kernel evaluates the similarity between the meanings of different documents and can recognise that two documents bear the same information. Duplicate information is filtered out and is not shown to the user. The underlying algorithm is based on the comparison of the signature of documents and is therefore highly efficient.

A parameter defining the similarity threshold under which documents are filtered out can be set by the users.

#### 10.1.9 Extraction of Citations

Quotations most representative of the documents meaning are extracted and stored automatically to help the user in interpreting the results. This extraction is based on AMI's ability to distinguish within any given text what is essential and what can be regarded as “additional”. This functionality uses the signature calculation process.

### **Context sensitive abstract:**

The document sentences which contain the higher density of the words included in the scenarios – i.e. weak or strong condition – are shown to the user.

This feature is available in the Publish/view modes.

#### **10.1.10 Trend detection**

*Trends* are words or expressions which are relevant and frequently occur.

Trends can be tracked from the statistical analysis functions of AMI "Analyze". Trend tracking can be organised from the Administration screen.

#### **10.1.11 Automatic learning**

The knowledge base is managed and enriched from the dedicated administration application (see section 6.2.3). It can also be enriched through a learning process which can be activated, and which records and learns dedicated vocabulary from the documents collected.

This vocabulary is useful in terms of resolving ambiguities in the query terms.

### **10.2 Generic connector**

The *generic connector*, is an HTTP/XML protocol which acts as a logical connector.

It is embedded in the **AMI** kernel and can be called from the **AMI E.I.** web interface like a physical connector.

Examples and libraries are provided in PHP, Perl and Java.

## 11 Compatibility with other AMI applications

The **AMI Enterprise Intelligence** architecture is designed to ensure maximum compatibility with the other AMI solutions.

### 11.1 Source management compatibility

The AMI Enterprise Intelligence v4.0 application Base Server is fully compatible with preceding versions.

### 11.2 Knowledgebase compatibility

The AMI Enterprise Intelligence v4.0 application and knowledgebase is fully compatible with that of the preceding versions 3.x.

Existing knowledgebase content is preserved during any migration process across varying versions of AMI Enterprise Intelligence.

## 12 Compatibility with other applications

AMI currently supports technology integrations with the following vendors

- ESRI for GIS (geographical Information system)
- Factiva for news and market information
- Pikko for graphic concept presentations
- Systran for language translation
- Temis for text mining solutions

## 13 Documentation, contact and support

Documentation for AMI Enterprise Intelligence is available in either English or French in PDF or HTML form.

The documentation set helps the administrator install; configure and manage AMI Enterprise Intelligence. It describes the functional and technical parameters which the administrator can use to control the environment.

The documentation set is composed of:

- ✓ Reference guide
- ✓ Programmer guide
- ✓ User guide
- ✓ Technical pre-requisites

Contact information can be found on the web site <http://www.amisw.com>

## 14 Appendices

### 14.1 Technical Appendices

#### 14.1.1 Physical architecture

The physical architecture of **AMI Enterprise Intelligence** – i.e. the way files are organised on disk – strictly matches the functional architecture. Please refer to the document entitled *Technical pre-requisite of AMI Enterprise Intelligence*.

#### 14.1.2 Security

##### 14.1.2.1 Authentication

The user *authentication* mechanism is global to the application. It is centralised and relies on the LDAP mechanism.

Each user is granted rights for the whole application, and specific rights for each application.

Authentication ensures that specific access rights can be defined for each user.

##### 14.1.2.2 Web security

*Web security* is supported by the physical architecture and defines several document repository spaces:

- ✓ the public space: the space which can be accessed by anyone without authentication
- ✓ the identification space: the web space which can be accessed by authenticated users
- ✓ the internal space: the space which can only be accessed internally (Web Services)
- ✓ the “prohibited” space: the elements which should not be accessed.

Web security ensures that the physical integrity of documents is protected.

#### 14.1.3 Scripting language

AMI E.I. Base Server includes *albScript* - AMI's script language based on the ECMA-262 (JavaScript) language. It is open (XML) and allows exchange with Web Services, internal programmes or external libraries, as well as with ActiveX (COM) applications.

The *albScript* language provides all the functions of AMI which can be applied to any type of application. It provides functions to handle AMI objects and their properties.

## **14.2 Legal aspects on the use of information sources**

AMI Software takes this opportunity to highlight to prospective users that documents retrieved via both manual and automated search processes can bear intellectual property rights and copyrights. The use of information resources and, in particular, the documents available via them, is governed by specific laws in each country.

These rights usually restrict the duplication and distribution of documents without the copyright owners prior consent.

Compliance with relevant intellectual and copyright is the responsibility of the respective users of AMI Software solutions.

## 15 Index

albScript, 55  
authentication, 55  
automatic learning, 11  
Collect, 21  
connectors, 10  
Decision making, 8  
de-duplication, 10  
duplicate information, 50  
email alerts, 25  
generic connector, 51  
Generic connector, 10  
kernel, 9  
Knowledge base, 11, 24  
modules, 8  
noise, 8  
physical architecture, 55  
Publishing space, 29  
relevance level, 49  
Scenario, 22  
Scenario, activate, 23  
Search hypothesis, 24  
signature, 11  
Software solution, 8  
status, 29, 30  
stop words, 50  
Subscription, 22  
Topic, 22  
trends, 50  
web interface, 9  
web security, 55

.oOo.