

DATASHEET

KENDO



Making Innovation

Sommario

1	INTRODUZIONE	3
2	OVERVIEW	4
2.1	Target di riferimento	4
2.2	Scenari d'uso e applicazioni	5
3	L'ARCHITETTURA DI KENDO®	6
3.1	Analyzing	7
3.1.1	Crawling.....	8
3.1.2	Grabbing & Storing	8
3.1.3	Reporting	9
3.1.4	Tipologie di Report	10
3.2	Editing	13
3.2.1	Automatic Generated Content.....	14
3.2.2	User Generated Content	14
3.3	Delivery & Publishing	15
3.3.1	Workflow designer.....	16
3.3.1.1	Data source	17
3.3.1.2	Content cleaner	17
3.3.1.3	Url mapper.....	18
3.3.1.4	Replacer	18
3.3.1.5	Image converter.....	18
3.3.1.6	Exporter	19
4	REQUISITI TECNICI	19
4.1	Tecnologie utilizzate	19

1 Introduzione

Forniamo, con questo documento, una panoramica completa sulle funzionalità della soluzione KENDO®.

Il documento descrive le fasi che identificano un tipico flusso di lavoro e come queste consentono di analizzare, trasformare e correggere in poco tempo un sito fino ad ottenere un prodotto accessibile secondo gli standard internazionali e le leggi vigenti in Italia in materia.

L'attuale versione è efficace per i siti di contenuto, statici o dinamici, con l'esclusione di quelli ad accesso riservato tramite autenticazione o con logiche applicative, sia client che server, che alterano il contenuto riferito alla URL (unified resource location)

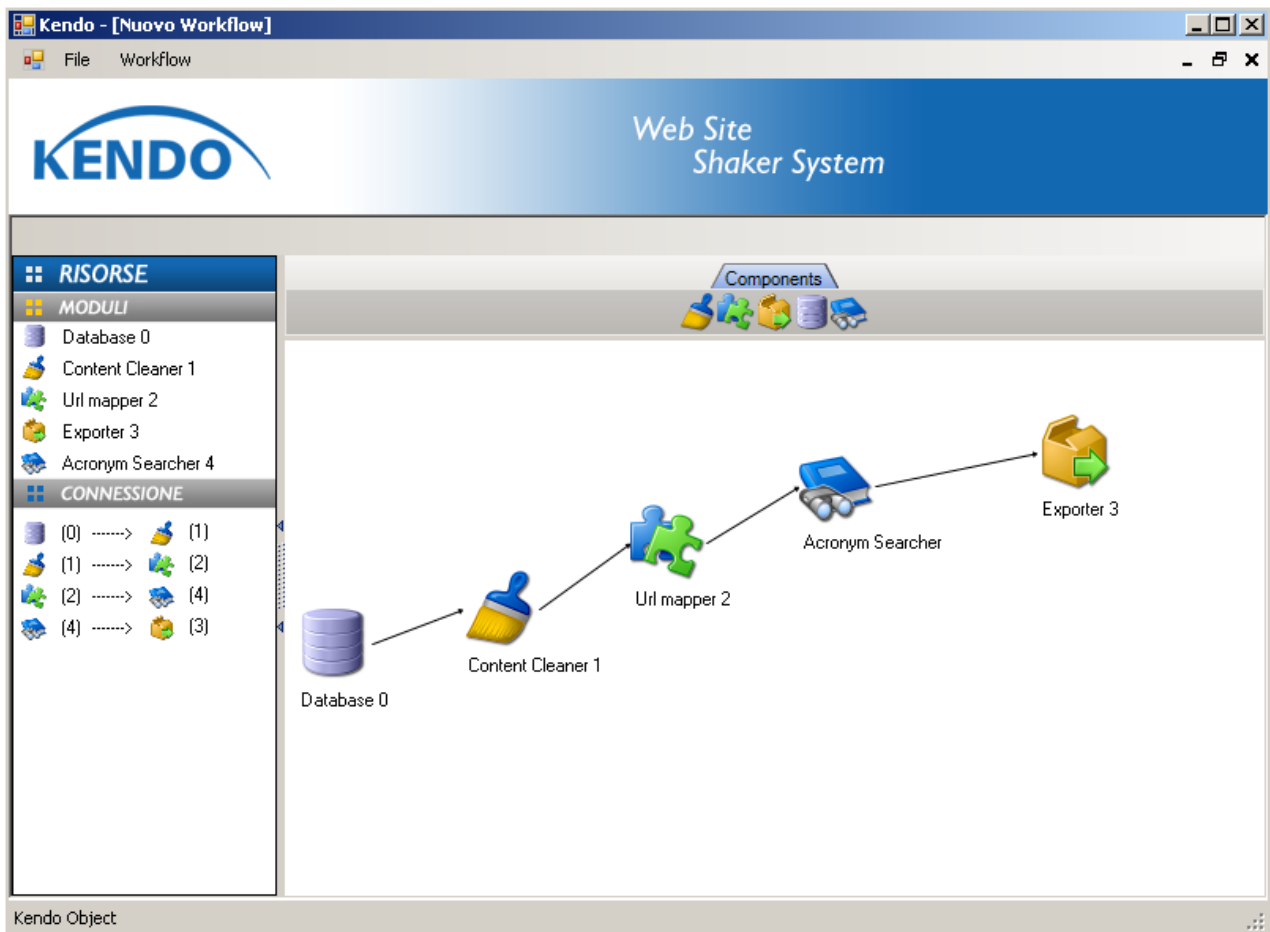


Immagine 1 - Kendo Workflow Designer

2 Overview

Nato dall'esperienza di Eustema S.p.A. nella realizzazione di decine di grandi portali per la Pubblica Amministrazione, KENDO® è un'avanzata piattaforma software dotata di un *workflow designer* con diverse risorse e moduli basati sulle più innovative tecnologie web e multimediali sviluppate da Eustema S.p.A.

KENDO® opportunamente configurato, è in grado di importare, analizzare, editare, modellare, trasformare ed esportare autonomamente un sito web in poche ore.

Grazie alla sua architettura distribuita e basata su una moltitudine di moduli specializzati, KENDO® è in grado di fornire un mix di tecnologie adatte a risolvere i più svariati problemi legati alla produzione di contenuti testuali e multimediali sul web.

KENDO® infatti, è in grado di **analizzare (Analyzing)** profondamente la struttura ed ogni tipo di risorsa presente in un sito web fornendo dettagliati report di tipo statistico e quantitativo; **modellare ed editare (Editing)** tali risorse attraverso regole automatiche suggerite dal sistema o attraverso il modulo UGC (User Generated Content) che permette all'operatore di modificare e "arricchire" direttamente i contenuti in modo intuitivo e veloce agendo solo ed esclusivamente nei punti in cui si verificano particolari anomalie rispetto gli standard W3C – WAI, e i 22 requisiti della vigente normativa italiana sulle "Disposizioni per favorire l'accesso dei soggetti disabili agli strumenti informatici" (Legge 9 gennaio 2004, n. 4).

Le fasi di analisi e modellazione sono propedeutiche all'ultima e più importante fase di **esportazione e ripubblicazione (Delivery & Publishing)** dei contenuti, dove affidandosi al componente *workflow designer* è possibile configurare diversi *output device* e quindi scegliere se ripubblicare (semplicemente o applicando un nuovo template grafico appositamente predisposto), l'intero sito in formato XHTML 1.0 Strict e accessibile, esportare i contenuti verso un CMS, proprietario o open source, oppure esportare in formato XML.

In sintesi

KENDO® è la soluzione completa, rapida ed economica per rendere subito compatibile un sito web con le norme di accessibilità italiane e internazionali.

2.1 Target di riferimento

I principali ambiti di intervento di KENDO® sono:

- Enti ed Istituti della Pubblica Amministrazione;
- Aziende pubbliche e private che necessitano di distribuire informazioni sul web senza "barriere tecnologiche";
- Aziende ICT fornitrici di soluzioni web per la PA.

- Aziende che si occupano di censimento statistico e monitoraggio di siti web

Per questi principali *stakeholders* KENDO® si pone come strumento per il monitoraggio della qualità e il *problem solving* in materia di accessibilità di portali, siti e applicazioni web.

Fino ad oggi il problema della manutenibilità dei portali accessibili, ha spinto molte aziende e associazioni ad introdurre nel mercato IT, soluzioni prettamente specifiche per la validazione automatica del codice e di individuazione dei possibili punti di intervento.

KENDO® è la **prima soluzione che va oltre la semplice validazione**, infatti è in grado di frammentare un intero portale web, analizzarlo e ricostruirlo attraverso una moltitudine di processi automatizzati sempre e comunque “orchestrati” dall’operatore che, grazie ad un’interfaccia intuitiva, possiede un elevato *feed-back* ad ogni contesto e la possibilità di decidere il livello di approfondimento dei *tasks* da eseguire compatibilmente con il proprio *know-how* o i requisiti espressi dal cliente.

2.2 Scenari d’uso e applicazioni

Eustema S.p.A. propone la soluzione KENDO® e l’alta professionalità dei suoi esperti in tutti quei contesti in cui è presente una redazione web che lavora su grosse mole di informazioni e non dispone di risorse interne in grado di migrare in forma accessibile i contenuti.

KENDO® può essere erogato come:

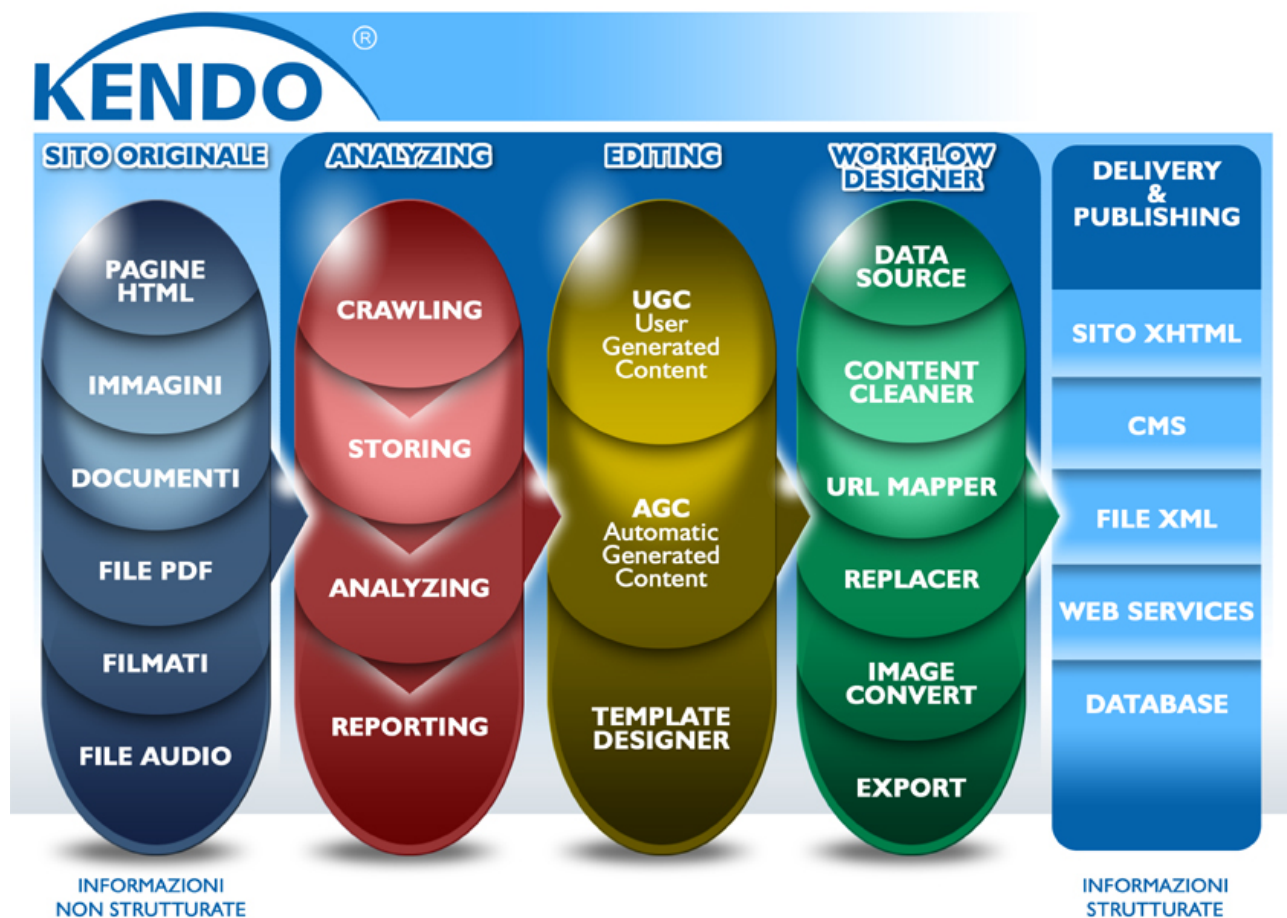
- **servizio “one shot”**, porting alla pari del sito in regime di accessibilità;
- **servizio continuativo** con interventi mirati a controlli periodici;
- **progetti ad-hoc** e soluzioni integrate;
- **Servizio ASP on demand** per la validazione e controllo a campione.

3 L'architettura di KENDO®

KENDO® scompone il processo di elaborazione di un sito internet in tre fasi principali:

- Analyzing
- Editing
- Delivery & Publishing

Per ognuna di queste fasi, la soluzione mette a disposizione degli operatori una suite di componenti visuali che agevolano e automatizzano il processo di correzione e arricchimento dei contenuti del sito, nonché la sua esportazione e ripubblicazione.



Di seguito si dettagliano le singole fasi e le relative funzionalità.

3.1 Analyzing

La fase di Analisi di KENDO[®] è una procedura completamente automatizzata che richiede come input semplicemente la URL di un sito o di un'applicazione web e la possibilità di scegliere se salvare direttamente il sito in locale (Grabbing & Storing) o meno.

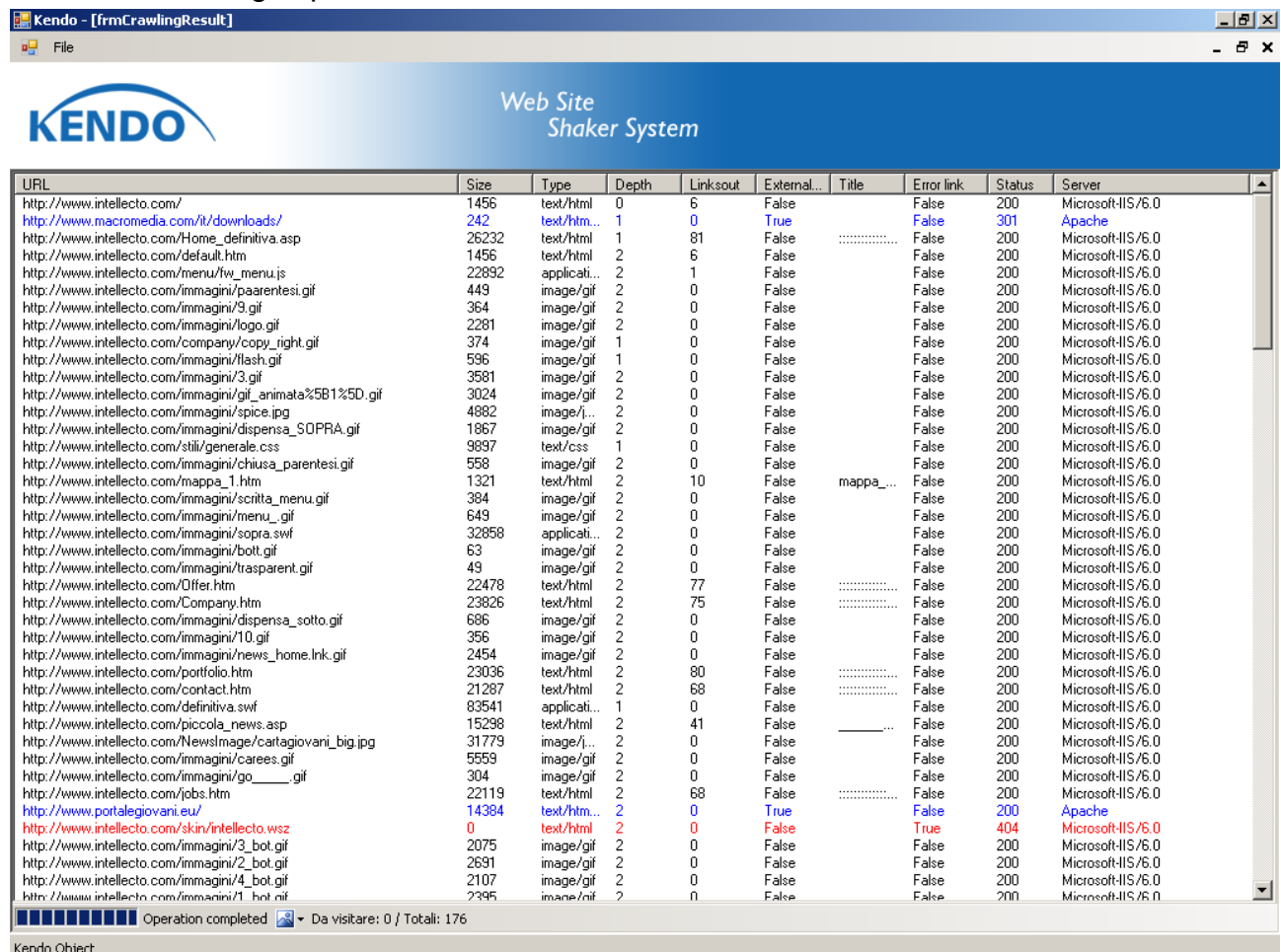
KENDO[®] effettua in tutte le pagine e le risorse del sito le seguenti tipologie di analisi:

- di tipo **statistico**:
- di **composizione logica** per la separazione tra dati e presentazione;
- di struttura **gerarchica e relazionale**;
- di **sintassi del codice** usato.

Di seguito si dettagliano le funzionalità del modulo Analyzer.

3.1.1 Crawling

L'Analyzer è dotato di un potente *Crawler* che analizza i contenuti in un modo metodico e automatizzato. Navigando all'interno del sito riferenzia tutte le risorse e struttura un dettagliato Data Base delle URI che consentirà poi in fase di Grabbing di prelevare tutte le risorse del sito.



URL	Size	Type	Depth	Linkout	External...	Title	Error link	Status	Server
http://www.intellecto.com/	1456	text/html	0	6	False		False	200	Microsoft-IIS/6.0
http://www.macromedia.com/it/downloads/	242	text/html	1	0	True		False	301	Apache
http://www.intellecto.com/Home_definitiva.asp	26232	text/html	1	81	False	False	200	Microsoft-IIS/6.0
http://www.intellecto.com/default.htm	1456	text/html	2	6	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/menu/fw_menu.js	22892	applicati...	2	1	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/parentesi.gif	449	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/3.gif	364	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/logo.gif	2281	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/company/copy_right.gif	374	image/gif	1	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/flash.gif	596	image/gif	1	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/3.gif	3581	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/gif_animata%5B1%5D.gif	3024	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/spice.jpg	4882	image/j...	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/dispensa_SOPRA.gif	1867	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/stili/generale.css	9897	text/css	1	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/chiusa_parentesi.gif	558	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/mappa_1.htm	1321	text/html	2	10	False	mappa_...	False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/scritta_menu.gif	384	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/menu_swf	649	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/sopra.swf	32858	applicati...	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/bot.gif	63	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/trasparent.gif	49	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/Offet.htm	22478	text/html	2	77	False	False	200	Microsoft-IIS/6.0
http://www.intellecto.com/Company.htm	23826	text/html	2	75	False	False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/dispensa_sotto.gif	686	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/10.gif	356	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/news_home.lnk.gif	2454	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/portofolio.htm	23036	text/html	2	80	False	False	200	Microsoft-IIS/6.0
http://www.intellecto.com/contact.htm	21287	text/html	2	68	False	False	200	Microsoft-IIS/6.0
http://www.intellecto.com/definitiva.swf	83541	applicati...	1	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/piccola_news.asp	15298	text/html	2	41	False	False	200	Microsoft-IIS/6.0
http://www.intellecto.com/NewsImage/cartagiovani_big.jpg	31779	image/j...	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/carees.gif	5559	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/go_...gif	304	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/jobs.htm	22119	text/html	2	68	False	False	200	Microsoft-IIS/6.0
http://www.portalegiovani.eu/	14384	text/html	2	0	True		False	200	Apache
http://www.intellecto.com/skin/intellecto.wsz	0	text/html	2	0	False		True	404	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/3_bot.gif	2075	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/2_bot.gif	2691	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/4_bot.gif	2107	image/gif	2	0	False		False	200	Microsoft-IIS/6.0
http://www.intellecto.com/immagini/1_bot.gif	2395	image/gif	2	0	False		False	200	Microsoft-IIS/6.0

Operation completed Da visitare: 0 / Totali: 176

Immagine 2 - Visualizzazione attività del Crawler

3.1.2 Grabbing & Storing

In questa fase avviene l'estrazione e la memorizzazione dell'intero sito analizzato da KENDO®.

L'applicazione memorizza i dati nelle seguenti modalità:

- **Data base relazionale** in tabelle separate e specializzate per:

- URI delle risorse interne al sito, con informazione sullo stato, dimensione, titolo eventuale codice di mark-up ecc;
 - Immagini (tag IMG) con attributi specifici e riferimenti di contesto;
 - Documenti e altri contenuti multimediali
 - Link (tag A) con attributi specifici e riferimenti di contesto;
 - Tabella “site map” nella quale sono memorizzati tutte le relazioni tra le risorse del sito.
- **File system**
- Il sito viene (se richiesto) completamente memorizzato in locale sulla macchina, ricostruendo la struttura dei file in maniera pedissequa all’analisi effettuata in fase di Crawling.

3.1.3 Reporting

La fase di reporting è il momento in cui l’operatore viene a conoscenza della situazione generale del sito. I moduli di reporting di KENDO[®] consentono di visualizzare una vasta mole di informazioni partendo da un livello generale in cui è possibile avere una visione statistica dei risultati fino ad arrivare ad un livello particolare in cui si scende nel dettaglio delle peculiari anomalie riscontrate al fine di poter agire direttamente attraverso la successiva fase di Editing alla risoluzione.

Questo metodo definito da Eustema S.p.A. “*discovery and fixing interface system*”, consente all’operatore di verificare, pianificare e agire sulle anomalie riscontrate con la possibilità di ripetere l’analisi fino ad ottenere la correzione totale

I report, organizzati per argomenti, sono concepiti per livelli di dettaglio e si dividono in:

■ **Report generale statistico**

Contiene informazioni espresse in valori statistici. Presenta un pulsante iconico che consente di accedere al report quantitativo e un’icona che simboleggia se sulla tipologia del report è richiesto un particolare intervento utente (accesso al modulo Editor) o meno.

- **Report quantitativi**

Contiene informazioni dettagliate e fornisce indicazioni precise su ogni aspetto analizzato o anomalia riscontrata. Le singole anomalie vengono espresse in valore quantitativo per numero di occorrenze e contesti in cui si presentano. Da questa vista è possibile poi accedere alla vista di dettaglio o di editing (se consentito).

- **Dettaglio/Editing**

Nella vista di dettaglio è possibile prendere visione in maniera puntuale di tutti i casi (o contesti) in cui si verifica una data anomalia. In questa fase è possibile agire come specificato nel paragrafo 3.2 in due modalità AGC e UGC.

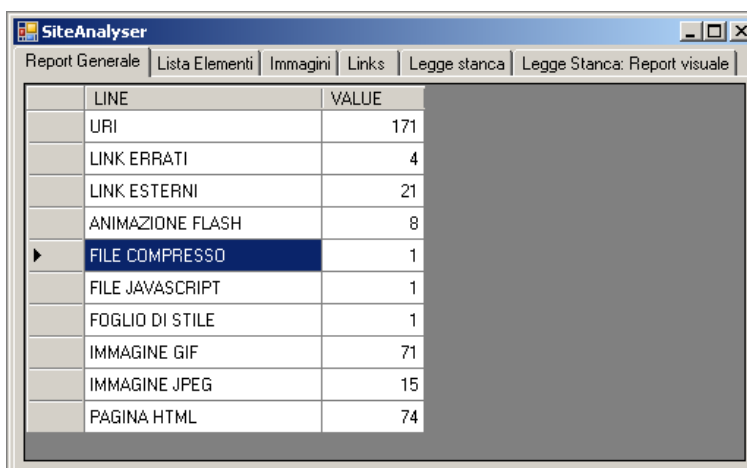
3.1.4 Tipologie di Report

KENDO® presenta le seguenti macro aree di reportistica che danno poi luogo a report quantitativi e di dettaglio o editing.

- **Censimento**

E' un report prevalentemente riassuntivo, configurabile ed espandibile, che fornisce all'operatore una panoramica generale del sito analizzato fornendo informazioni circa:

- Analisi statistica delle risorse individuate dal crawler;
- N. pagine e risorse analizzate organizzate per:
 - Mime type
 - estensione
 - dimensione
 - URI
 - Status code



LINE	VALUE
URI	171
LINK ERRATI	4
LINK ESTERNI	21
ANIMAZIONE FLASH	8
FILE COMPRESSO	1
FILE JAVASCRIPT	1
FOGLIO DI STILE	1
IMMAGINE GIF	71
IMMAGINE JPEG	15
PAGINA HTML	74

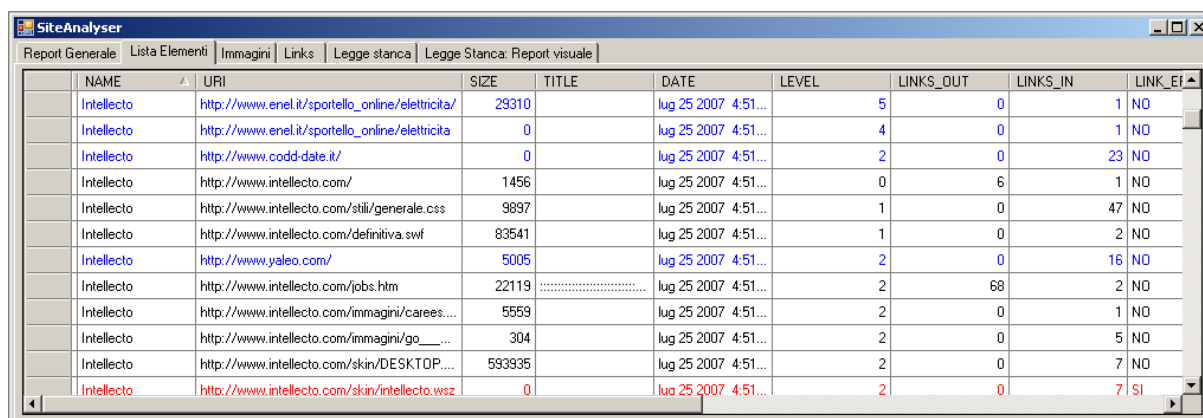
Immagine 3 - Report generale

■ **Efficienza e Qualità**

La finalità di questo report è di individuare i fattori critici che possono minacciare l'efficienza di un sito web e quindi la qualità della sua navigabilità e disponibilità sul web.

In particolare fornisce informazioni su:

- Link a riferimenti interni o esterni non raggiunti dal Crawler;
- Peso delle pagine e delle risorse scaricate dal Grabber evidenziando i file che superano la soglia massima definita dalle euristiche web;
- Tempo di elaborazione medio per URI.



NAME	URI	SIZE	TITLE	DATE	LEVEL	LINKS_OUT	LINKS_IN	LINK_EF
Intellecto	http://www.enel.it/sportello_online/elettricit...	29310		lug 25 2007 4:51...	5	0	1	NO
Intellecto	http://www.enel.it/sportello_online/elettricit...	0		lug 25 2007 4:51...	4	0	1	NO
Intellecto	http://www.codd-date.it/	0		lug 25 2007 4:51...	2	0	23	NO
Intellecto	http://www.intellecto.com/	1456		lug 25 2007 4:51...	0	6	1	NO
Intellecto	http://www.intellecto.com/stili/generale.css	9897		lug 25 2007 4:51...	1	0	47	NO
Intellecto	http://www.intellecto.com/definitiva.swf	83541		lug 25 2007 4:51...	1	0	2	NO
Intellecto	http://www.yaleo.com/	5005		lug 25 2007 4:51...	2	0	16	NO
Intellecto	http://www.intellecto.com/jobs.htm	22119	lug 25 2007 4:51...	2	68	2	NO
Intellecto	http://www.intellecto.com/immagini/carees....	5559		lug 25 2007 4:51...	2	0	1	NO
Intellecto	http://www.intellecto.com/immagini/go_....	304		lug 25 2007 4:51...	2	0	5	NO
Intellecto	http://www.intellecto.com/skin/DESKTOP....	593935		lug 25 2007 4:51...	2	0	7	NO
Intellecto	http://www.intellecto.com/skin/intellecto.wsz	0		lug 25 2007 4:51...	2	0	7	SI

Immagine 4 - Report dettaglio risorse analizzate

■ **Accessibilità (rif. Decreto Ministeriale 8 luglio 2005 - Legge 9 gennaio 2004)**

Il report di accessibilità presenta le seguenti informazioni a seconda dei tre livelli di visualizzazione:

○ **Livello1: report generale**

Fornisce un dato statistico in percentuale sulla congruità del sito ai 22 requisiti e un dato numerico sempre statistico che restituisce quanti dei 22 requisiti sono soddisfatti per l'intero sito analizzato, discriminando la tipologia di messaggio (error, warning).

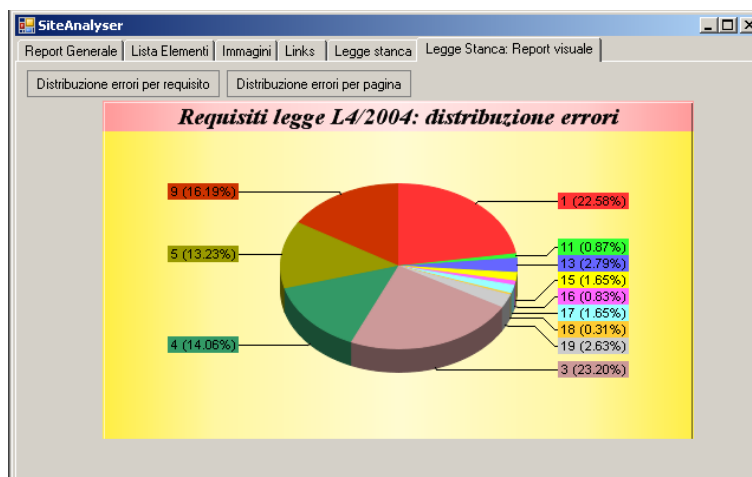
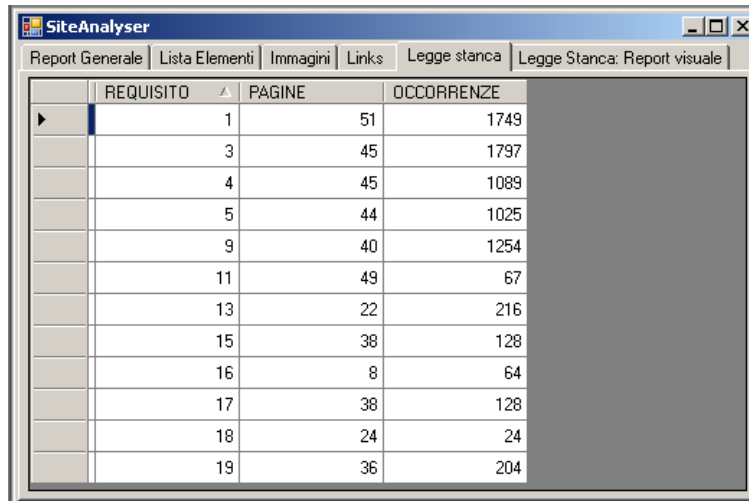


Immagine 5 - Report grafico statistiche accessibilità

○ **Livello 2: report quantitativo**

Visualizza in maniera più dettagliata per ciascun requisito il numero delle

occorrenze in cui si verificano particolari anomalie. Accanto ad ogni requisito è poi possibile aprire la vista di dettaglio/editing per intervenire attraverso la modalità AGC o UGC (vedi par. 3.2).



REQUISITO	PAGINE	OCCORRENZE
1	51	1749
3	45	1797
4	45	1089
5	44	1025
9	40	1254
11	49	67
13	22	216
15	38	128
16	8	64
17	38	128
18	24	24
19	36	204

Immagine 6 - Report quantitativo 22 requisiti Legge N. 4/2004

- **Livello 3: dettaglio/editing**

A questo livello è possibile **intervenire e correggere** le anomalie riscontrate. La vista di dettaglio in realtà si pre-configura in modalità editing fornendo dei potenti *tools* e *features* per agire su grosse mole di informazioni in maniera rapida e ricorsiva.

3.2 Editing

La fase di Editing permette all'operatore di aggiungere e arricchire i testi presenti nelle pagine del sito analizzato e di configurare e predisporre i template di presentazione dei formati editoriali.

L'Editor di Kendo dà la possibilità all'utente di intervenire sui contenuti attraverso due modalità:

- AGC - Automatic Generated Content;
- UGC - User Generated Content.

3.2.1 Automatic Generated Content

La fase di analisi come già specificato, restituisce in output dei report molto dettagliati con la possibilità di interagire direttamente con il modulo Editor e quindi di risolvere i vari problemi e anomalie riscontrate in maniera contestuale e filtrata.

Per le anomalie di tipo oggettivo riscontrate KENDO[®], è in grado di suggerire all'operatore alcune soluzioni da applicare in maniera puntuale o ricorsiva ai contenuti testuali e meta-dati attraverso il processo di generazione automatica dei contenuti.

3.2.2 User Generated Content

Nel caso in cui l'AGC non riesca a soddisfare i requisiti dell'operatore o comunque non riesca a risolvere problemi legati ad esempio alla contestualizzazione di determinati contenuti nell'ambito dell'alberatura generale del sito o nell'ambito di anomalie di tipo soggettivo come il colore e intrusività di alcuni oggetti di tipo APPLET, entra in gioco l'UGC ovvero lo **User Generated Content** che permette all'operatore di aggiungere, arricchire e manipolare ogni risorsa presente nel sito sia in maniera ricorsiva, operando con le selezioni multiple, che puntuale per ogni tipo di contenuto e contesto semantico.

Quindi l'Editor consente di agire sulle liste delle risorse (immagini, link, script ecc) del sito, intervenendo con una apposita finestra di dialogo per inserire e modificare testi o correggere eventuali anomalie.

In particolare, l'Editor permette di agire sulla lista delle pagine analizzate, correggendo i link non validi e sulle liste delle risorse individuate, immagini e link, inserendo testi alternativi, longdesc, dimensioni, fogli di style e altro ancora.

Nel caso di azioni compiute sulle risorse individuate nel sito, una apposita finestra di dialogo dell'Editor consente di filtrare gli elementi della lista per criteri di ricerca quali:

- mimetype della risorsa
- attributo mancante
- testo contenuto nei tag
- a altre ancora.

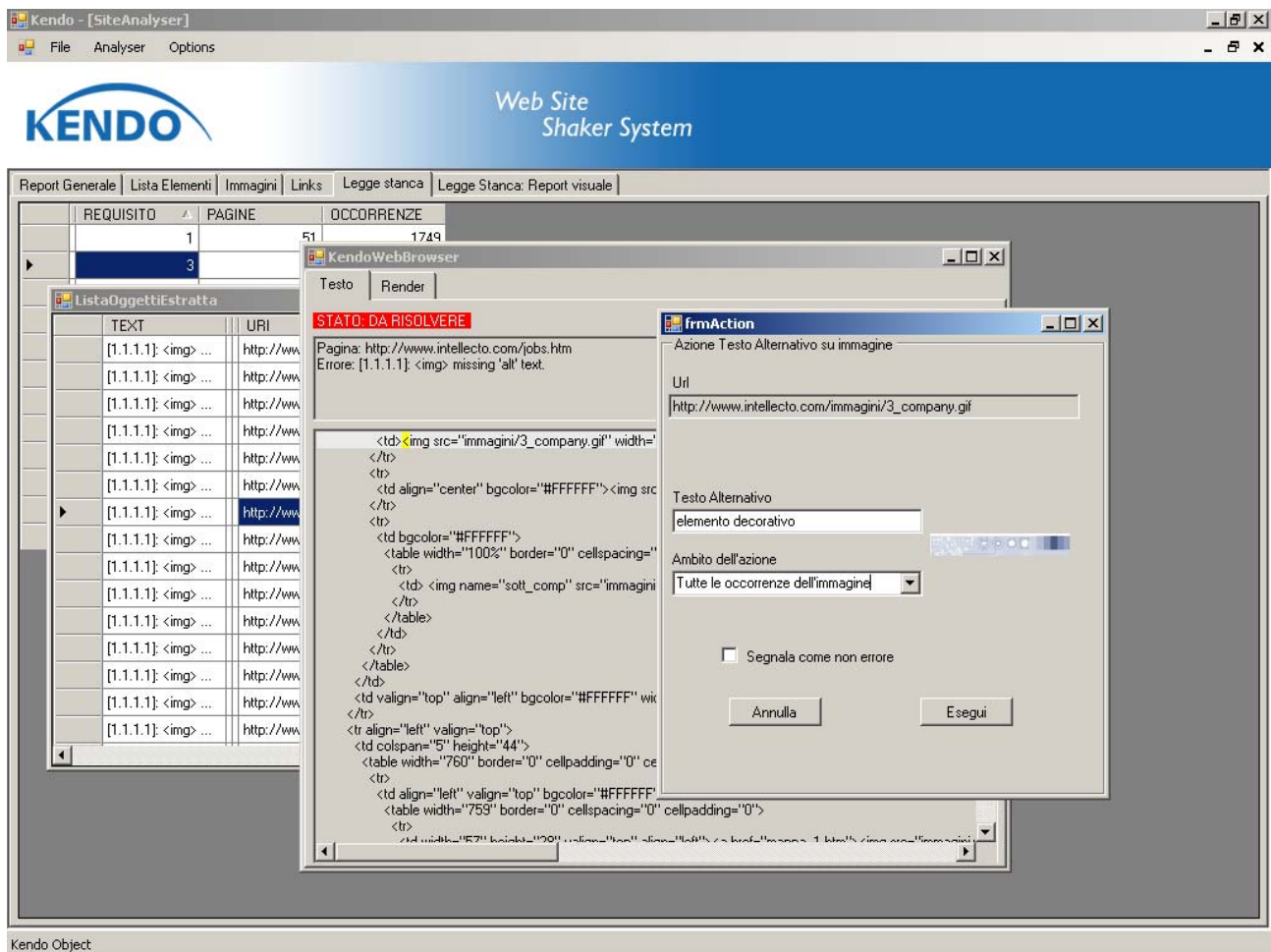


Immagine 7 - Esempio di intervento UGC - inserimento testi alternativi

Per quanto riguarda la predisposizione dei template è richiesto l'intervento di un web designer che adatta i modelli di presentazione disponibili nel sistema alla struttura delle pagine del sito originario. In futuro sarà messa a disposizione un'interfaccia WYSIWYG che agevola la costruzione dei template.

3.3 Delivery & Publishing

Effettuate le fasi di **Analyzing** ed **Editing** l'operatore dovrà configurare ed eseguire la fase di esportazione e pubblicazione del contenuto del sito esaminato.

Questa fase ha lo scopo di produrre il nuovo sito con una veste grafica vicina al concept editoriale originario, o eventualmente rinnovata nello stile di presentazione.

I contenuti e le altre informazioni estratte dal sito analizzato potranno essere esportate nei formati quali:

- Pagine statiche in formato html con relativi file allegati
- flusso dati per sistemi di content management, proprietari o open source
- file in formato XML
- flusso dati attraverso “Webservices”

Un'applicazione dove Kendo risulta particolarmente efficace è la pubblicazione di siti statici nel rispetto delle indicazioni date dalla vigente normativa sulle “Disposizioni per favorire l'accesso dei soggetti disabili agli strumenti informatici” (Legge 9 gennaio 2004, n. 4). Ciò avviene attraverso la progettazione di template di presentazione realizzati dai web designer sui requisiti di progetto.

La configurazione di questa fase avviene utilizzando un componente visuale di KENDO[®] denominato **Workflow designer**.

3.3.1 Workflow designer

Questo componente permette la definizione delle operazioni di pulizia, trasformazione ed esportazione da applicare al risultato della fase di analisi e di editing del sito.

Il workflow designer è una componente visuale, che tramite il metodo di *drag&drop* consente di disegnare un flusso di operazioni da effettuare su una base dati di Kendo.

Le operazioni sono rappresentate da moduli indicati sulla barra degli strumenti del componente. Ogni modulo può essere configurato dall'utente attraverso una apposita finestra di dialogo.

Di seguito sono illustrati i moduli principali del componente.

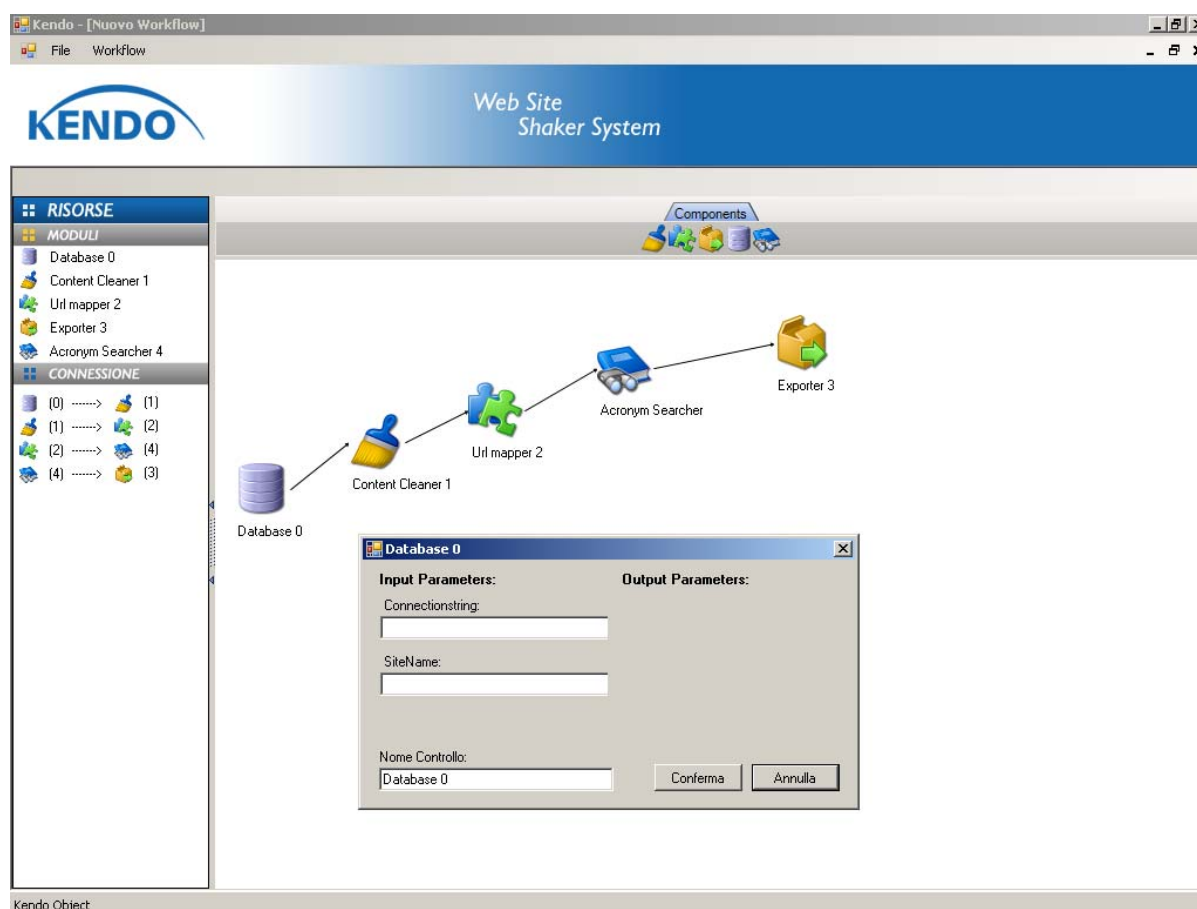


Immagine 8 - Workflow designer - form di configurazione modulo

3.3.1.1 Data source

Rappresenta la base di dati, restituita dalla fase di analisi ed editing, sulla quale agiranno i moduli configurati nel *workflow*.

Ogni attività sulla sorgente dati genera una nuova versione delle informazioni. Questo consente l'annullamento delle operazioni svolte dai moduli senza rischiare di perdere i dati originali.

La finestra di dialogo per la configurazione del modulo consente di impostare i dati per la connessione al database e di scegliere il sito analizzato tra quelli presenti.

3.3.1.2 Content cleaner

Il modulo consente di estrarre le informazioni principali dalle pagine elaborate in fase di analisi.

L'attività del modulo consiste nell'estrarre, in una struttura XML, il contenuto e i metadati che compongono le pagine del sito internet analizzato.

Il contenuto estratto, viene ripulito dalle parti di codice di presentazione e di blocchi html ripetitivi come ad esempio i menu, il footer ed altri elementi comuni a tutte le pagine del sito.

Il modulo consente, anche di:

- applicare delle sostituzioni al testo;
- aggiungere informazioni mancanti alle risorse che lo necessitano;

I dati ottenuti potranno in seguito essere utilizzati da altri moduli per essere inseriti in template al fine di produrre del nuovo codice XHTML accessibile.

La finestra di dialogo per la configurazione del modulo consente di:

- selezionare una cartella del sito sulla quale intervenire
- impostare valori di default per le informazioni mancanti, per ogni tipologia di risorsa analizzata
- impostare una o più sostituzioni testuali al contenuto estratto

3.3.1.3 Url mapper

Il modulo permette di applicare delle regole per trasformare i link censiti al fine di soddisfare la nuova struttura del sito, creata dopo le operazioni di estrazione del contenuto.

La finestra di dialogo per la configurazione del modulo consente di specificare la regola da utilizzare per modificare le URL presenti nei contenuti.

3.3.1.4 Replacer

Il modulo permette di ricercare e sostituire parole, frasi o porzioni di codice html presenti nelle pagine del sito con altre fornite dall'operatore.

3.3.1.5 Image converter

Immagini in formato non standard, quali tiff, bmp, psd, vengono convertiti nei formati web standard jpg o png. Inoltre il modulo effettua ridimensionamenti mirati a ottimizzare eventuali sovradimensionamenti delle immagini nel contesto in cui sono utilizzate.

3.3.1.6 Exporter

Il modulo Exporter consente di esportare il contenuto ottenuto tramite il modulo di "Content Cleaner" in diverse forme.

In particolare il modulo Exporter consente di:

- inserire i contenuti raccolti in un sistema di content management;
- inserire i contenuti in una base di dati attraverso il mapping delle informazioni con le tabelle del database di destinazione;
- esportare i contenuti generando pagine XHTML accessibile utilizzando degli appositi template;
- esportare le informazioni raccolte in un formato XML;
- restituire le informazioni attraverso l'esposizione di un "Webservice".

4 Requisiti tecnici

4.1 Tecnologie utilizzate

Java, .NET, Ajax, XML, XSLT, XHTML, data base Ms SQL Server, Javascript, CSS.