

QuieW

Quality-based View of the Web

P.Avesani, D.Sona
ITC-irst

Web Directory

http://www.dmoz.org

The screenshot shows a browser window titled "ODP - Open Directory Project" with the address bar containing "http://www.dmoz.org/". The browser's menu bar includes "Razmataz", "Post", "IRST", "Stats", "Apple", "Events", "News", "Bookmarklet", "Meteo", and "CfP". The page header features the "dmoz" logo and the text "open directory project" in a green bar, with "In partnership with AOL search" on the right. Below the header is a navigation menu with links for "about dmoz", "suggest URL", "help", "link", and "editor login". A search bar with a "Search" button and a link to "advanced" search is present. The main content area is a grid of category links, each with a sub-menu:

- Arts**: [Movies](#), [Television](#), [Music...](#)
- Business**: [Jobs](#), [Real Estate](#), [Investing...](#)
- Computers**: [Internet](#), [Software](#), [Hardware...](#)
- Games**: [Video Games](#), [RPGs](#), [Gambling...](#)
- Health**: [Fitness](#), [Medicine](#), [Alternative...](#)
- Home**: [Family](#), [Consumers](#), [Cooking...](#)
- Kids and Teens**: [Arts](#), [School Time](#), [Teen Life...](#)
- News**: [Media](#), [Newspapers](#), [Weather...](#)
- Recreation**: [Travel](#), [Food](#), [Outdoors](#), [Humor...](#)
- Reference**: [Maps](#), [Education](#), [Libraries...](#)
- Regional**: [US](#), [Canada](#), [UK](#), [Europe...](#)
- Science**: [Biology](#), [Psychology](#), [Physics...](#)
- Shopping**: [Autos](#), [Clothing](#), [Gifts...](#)
- Society**: [People](#), [Religion](#), [Issues...](#)
- Sports**: [Baseball](#), [Soccer](#), [Basketball...](#)
- World**: [Deutsch](#), [Español](#), [Français](#), [Italiano](#), [Japanese](#), [Nederlands](#), [Polska](#), [Dansk](#), [Svenska...](#)

At the bottom, there is a "Become an Editor" button with the text "Help build the largest human-edited directory of the web" and a small green lizard illustration. The footer contains the copyright notice "Copyright © 1998-2006 Netscape" and the statistics "5,281,960 sites - 72,793 editors - over 590,000 categories".

SegnaWeb

http://www.segnaweb.it

The screenshot shows a web browser window with the address bar displaying "http://www.segnaweb.it/". The browser's menu bar includes options like "Razmataz", "Post", "IRST", "Stats", "Apple", "Events", "News", "Bookmarklet", "Meteo", "CfP", "Book", "Mappe", "QuieW", and "LanGam". The page content is organized into several sections:

- Left sidebar (green background):**
 - SegnaWeb logo (a circular graphic with letters W, G, N, A, B, S).
 - Text: "What is SegnaWeb", "Cos'è SegnaWeb".
 - Section: "LE RISORSE di SegnaWeb".
 - Text: "Criteri di selezione delle risorse", "Scambio dei link".
 - Section: "Il mio SegnaWeb (?)".
 - Text: "Login", "Logout".
 - Section: "Registrazione personale".
 - List: "Profilo utente", "Aspetto delle pagine".
- Main content area:**
 - Section: "Motori di ricerca (?)".
 - Search box with "cerca in SegnaWeb" and a "Cerca" button.
 - Section: "SegnaWeb: risorse Internet selezionate dai bibliotecari italiani".
 - Collaboration box: "Una collaborazione CILEA" with logos for AIB and CILEA. Text: "Un servizio per tutti, libero da interessi commerciali, teso a garantire informazioni corrette e attendibili." and "La redazione Per proposte, suggerimenti e consigli: segnaweb@aib.it".
 - Section: "Cerca per argomenti (?)".
 - Section: "Chiedi @lla tua biblioteca (?)".
 - Section: "Biblioteche e Cataloghi (?)".
 - Section: "Altri come SegnaWeb (?)".
 - Section: "Affari e economia" with sub-topics: "Commercio · Consumi e utenze · Economia, imprese e management · Finanza, moneta e banche · Fisco · Lavoro e concorsi · Pensioni e sicurezza sociale · Pesì e misure · Statistiche · Trasporti e traffico".
 - Section: "Arte e cultura" with sub-topics: "Architettura · Arti visive e plastiche · Cinema · Collezionismo · Fotografia · Fumetti · Letteratura italiana · Letterature · Libri e editoria · Moda · Musei, gallerie e mostre · Musica · Spettacoli e teatro · Televisione e radio · Urbanistica e paesaggio".
 - Section: "Attualità e media" with sub-topics: "Ecologia, meteo e ambiente · Emergenze · Eventi e appuntamenti nazionali · Giornali e agenzie di stampa · Lotto, lotterie, premi · Ora esatta · Sport · Televisione e radio".
 - Section: "Divertimento e tempo libero" with sub-topics: "Animali domestici · Auto e moto · Casa, giardino e bricolage · Cucina · Feste, sagre e folklore · Giochi e hobbies · Lotto, lotterie, premi · Sport · Televisione e radio · Turismo e viaggi".
 - Section: "Informatica e internet" with sub-topics: "E-mail gratuita · Grafica e immagini · Informatica · Internet · Motori di ricerca e directory · Software gratuito".
 - Section: "Istruzione e formazione".

Renewal

- **Quality-based Sampling**
Manual selection of web sources
- **Ranking Seeds**
Dealing with Google Bombs
- **Shallow Ontologies**
Extensional definition of concepts
- **Context-driven Queries**
Focused search

Sampling

Correlec | correlated electronics

http://www.correlec.com/ Google

Razmataz Post IRST Stats Apple Events News Bookmarklet Meteo CFP Book Mapped QuieW LanGam Matteo

KDD 2005 - KDD Cup 2... Correlec | correlated ele...

Database Updated May 24 2006

correlec
correlated electronic

Innovative solutions for enhanced productivity.

start here

home | about us | what we offer | contact us

Correlated data how you want it, when you want it

More data than ever before. Correlated like never before.

We know you are busy. Here is a **unique on-line working tool** for professionals who prefer evaluating results to searching innumerable web sites.

- Created and updated by professionals who understand your language, your needs, your frustrations.
- A few clicks will allow you to group, correlate, compare all products of the same category supplied by the manufacturers listed on the right of your screen.
- A few more clicks and you have a ready-to-use report or pdf file to share with your colleagues containing all the data you have selected, in the format of your choice. **YOU are in control!**

Want to know more? A few minutes now will save you days of hard work later. Just click "start here" above

ORIGINAL MANUFACTURERS

- Hitachi Ltd.
- Fujl Electric
- Infineon Technologies AG
- Mitsubishi Electric
- Nihon Inter Electronics Corporation (NIEC)
- Sanyo Electric Co. Ltd
- Toshiba Corporation

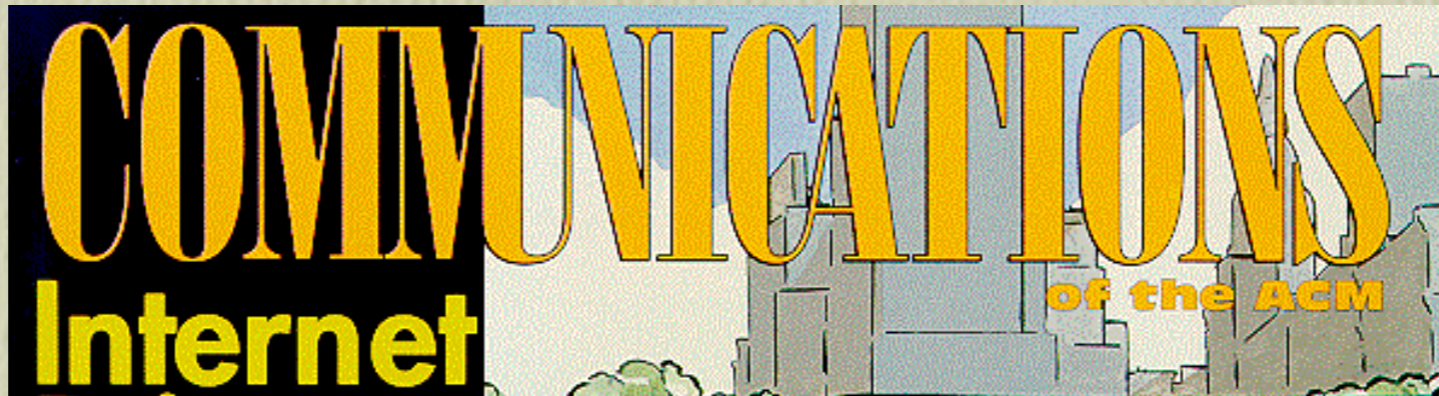
We transform dispersed data into *correlated information*, ready to use - © Copyright 2006 Correlec

http://www.correlec.com

Renewal

- **Quality-based Sampling**
Manual selection of web sources
- **Ranking Seeds**
Dealing with Google Bombs
- **Shallow Ontologies**
Extensional definition of concepts
- **Context-driven Queries**
Focused search

Ranking



The bubble of web visibility

Marco Gori* and Ian Witten*

* *Dipartimento di Ingegneria dell'Informazione, Università di Siena*
◊ *Department of Computer Science, University of Waikato, New Zealand*

The web seems like a Bergsonian library with a huge amount of information. Access to this treasure is mediated by dragons who compete amongst themselves for dominance. The battleground is so hostile that few will survive indeed, just one will likely achieve overall dominance in all but specialist corners of the library. The allegory is attractive except that the treasure is not private property but a public good, and the dragons must serve their own business interests rather than the public cause to compete successfully.

A vital source of information in everyday life, all the content of the web is readily available in principle. But the dominant mode of access is through search engines, and—in practice—the view they offer is restricted. Typical queries match hundreds of thousands of documents, which search engines sort according to ranking heuristics. For pragmatic reasons users view only the highest ranked pages, which gives rise to the notion of web page “visibility” as seen through the lens provided by search engines.

The heuristics adopted by popular search engines favor the construction of artificial communities that are expressly



Figure 1: How to promote a web page by playing on the number of input links. A keyword-page is created which links to a set of artificial pages containing the selected keywords which, in turn, link to the target page.

Renewal

- **Quality-based Sampling**
Manual selection of web sources
- **Ranking Seeds**
Dealing with Google Bombs
- **Shallow Ontologies**
Extensional definition of concepts
- **Context-driven Queries**
Focused search

Renewal

- **Quality-based Sampling**
Manual selection of web sources
- **Ranking Seeds**
Dealing with Google Bombs
- **Shallow Ontologies**
Extensional definition of concepts
- **Context-driven Queries**
Focused search

Query

http://www.acm.org/sigs/sigkdd/kdd2005

The screenshot shows a web browser window with the URL <http://www.acm.org/sigs/sigkdd/kdd2005/kddcup.html>. The page title is "KDD 2005 - KDD Cup 2005: Aug 21-24, Chicago, IL. USA". The main content area features a large "KDD-05" logo and the text: "The Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Hyatt Regency - Chicago, IL. USA. August 21-24". Below this is a navigation menu with links: home, program, awards, kdd-cup, submissions, hotel/travel, registration, and organizers. The main content is titled "KDD Cup 2005" and includes a "News" section with four entries: "August 30, 2005: [KDD-Cup presentation slides from the KDD conference](#)", "August 30, 2005: [Winning Teams](#)", "August 30, 2005: [Labeled Query Data](#)", and "August 10, 2005: [Solution Evaluation Result](#)". There is also an "Introduction" section and a "Contest Rules" section with sub-sections for "Agreement" and "Eligibility". The right sidebar contains "Organizational Sponsors" with logos for ACM, SIGKDD, SIGMOD, and AAAI, and links for "Student Travel Awards" and "Student Room Sharing".

KDD 2005 - KDD Cup 2005: Aug 21-24, Chicago, IL. USA

<http://www.acm.org/sigs/sigkdd/kdd2005/kddcup.html> kdd 2005 competition

Razmataz Post IRST Stats Apple Events News Bookmarklet Meteo CfP Book Mapped QuieW LanGam Matteo

KDD 2005 - KDD Cup 2... Correlec | correlated ele...

KDD-05 The Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Hyatt Regency - Chicago, IL. USA. August 21-24

home program awards kdd-cup submissions hotel/travel registration organizers

KDD Cup 2005

News

August 30, 2005: [KDD-Cup presentation slides from the KDD conference](#)

August 30, 2005: [Winning Teams](#)

August 30, 2005: [Labeled Query Data](#)

August 10, 2005: [Solution Evaluation Result](#)

Introduction

The KDD-Cup 2005 Knowledge Discovery and Data Mining competition will be held in conjunction with the Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. The task is selected to be interesting to participants from both academia and industry. In particular, we encourage the participation of students. This year's competition is about classifying internet user search queries. We are looking forward to an interesting competition and encourage your participation.

Contest Rules

Agreement

By sending the registration email, you indicate your full and unconditional agreement and acceptance of these contest rules.

Eligibility

The contest is open to any party planning to attend KDD 2005. A person can participate in only one group. Multiple submissions per group are allowed, since we will not provide feedback at the time of submission. Only the last submission before the deadline will be evaluated and all other submissions will be discarded.

Organizational Sponsors:

acm SIGKDD acm SIGMOD AAAI

[Student Travel Awards](#)

[Student Room Sharing](#)

Assumptions

- **Paper versus Web Sources**
 - Present: classification of knowledge on paper
 - Future: classification of knowledge on internet
- **Representation versus classification**
 - Extensional encoding of categories

Assumptions

- **Paper versus Web Sources**
 - Present: classification of knowledge on paper
 - Future: classification of knowledge on internet
- **Representation versus classification**
 - Extensional encoding of categories

Outline



Motivation

- Web directories revisited



Technology

- Web directories back-stage



Research

- Web directories intelligent supports

Technology Framework

Web



Web Directory

Front-End

User

Technology Platform

QUIEW - Editor Area

http://cork.dit.unitn.it/quiew/editor/explore/1149693624419011

Razmataz Post IRST Stats Apple Events News Bookmarklet Meteo CfP Book Mappe QuieW LanGam Matteo

QUIEW

My Workspace
User **Paolo Avesani** [Clipboard](#)


Library Browse by: [Directory](#) [Tags](#) [Resources](#)

[MeTe](#) > [Didattica](#) > Esperienze

Category: Esperienze [Edit](#) | [Add a Subcategory](#)

Subcategories [+]


QUIEW Recommends



Esperienze in prima persona
Esperienze in prima persona
Metodo di ricerca efficace >>
Inizio pagina > Esperienze in
prima persona | Sull' ...

[Certify me](#) | [Reject](#) | [Next](#)

QUIEW Recommends



Esperienze.net
Esperienze.net La Missione di
Esperienze.net è garantire la
qualità dell'apprendimento a
distanza Esperienze.net ...

[Certify me](#) | [Reject](#) | [Next](#)

Resources Page 1 of 1 Order by: [Newer first](#)

[Indire, sito ufficiale](#)
Indire, sito ufficiale home chi siamo contatti help miur usr irre invalsi 07/06/2006 [Punto edu Europa]
Ata] [...
<http://www.indire.it/> [more...](#) [Remove certification](#) | [Clip it](#)

Page 1 of 1

[About](#) [Privacy Policy](#) [Credits](#) [Help](#) [Contact Us](#)

Technology Prototype

Web Directory

Taxonomy of categories

Classification

Qualifying web sources

Tagging

Anytime tags association

Ranking

Topic-driven relevance criteria

Bookmarking

Manual on-the-fly selection

Recommendation

Hints on source classification

Searching

Full text focussed search

Feeds Aggregation

Organization of RSS feeds

Crawling

Monitoring of web sources

Reviewing

Annotations of web resources

Technology Prototype

Web Directory

Taxonomy of categories

Classification

Qualifying web sources

Tagging

Anytime tags association

Ranking

Topic-driven relevance criteria

Bookmarking

Manual on-the-fly selection

Recommendation

Hyperlink source classification

Searching

Full text focussed search

Top-Down Strategy

versus

Bottom-up Strategy

Aggregation

Aggregation of RSS feeds

Crawling

Monitoring of web sources

Reviewing

Annotations of web resources

Technology Prototype

Web Directory

Taxonomy of categories

Classification

Qualifying web sources

Tagging

Anytime tags associated

Static Perspective
versus

Open relevance criteria

Bookmarking

Manual on-the-fly selection

Dynamic Perspective

Recommendation

Hints on source classification

Searching

Full text focussed search

Feeds Aggregation

Organization of RSS feeds

Crawling

Monitoring of web sources

Reviewing

Annotations of web resources

Outline

- **Motivation**
 - Web directories revisited
- **Technology**
 - Web directories back-stage
- **Research**
 - Web directories intelligent supports

Research

- **Taxonomy-driven Classification**
 - Learning categorization of web sources
- **Category-driven Ranking**
 - Learning custom relevance criteria
- **Sharing Lexicon**
 - Distributed coordination on tagging annotation

Research Classification

QUIEW - Editor Area

http://cork.dit.unitn.it/quiew/editor/explore/1149693624419011

Razmataz Post IRST Stats Apple Events News Bookmarklet Meteo CfP Book Mappe QuieW LanGam Matteo

QUIEW

My Workspace
User **Paolo Avesani** [Clipboard](#)


Library Browse by: [Directory](#) [Tags](#) [Resources](#)

[MeTe](#) > [Didattica](#) > Esperienze

Category: Esperienze [Edit](#) | [Add a Subcategory](#)

Subcategories [+]


QUIEW Recommends



Esperienze in prima persona
Esperienze in prima persona
Metodo di ricerca efficace >>
Inizio pagina > Esperienze in
prima persona | Sull' ...

[Certify me](#) | [Reject](#) | [Next](#)

QUIEW Recommends



Esperienze.net
Esperienze.net La Missione di
Esperienze.net è garantire la
qualità dell'apprendimento a
distanza Esperienze.net ...

[Certify me](#) | [Reject](#) | [Next](#)

Resources Page 1 of 1 Order by: [Newer first](#)

[Indire, sito ufficiale](#)
Indire, sito ufficiale home chi siamo contatti help miur usr irre invalsi 07/06/2006 [Punto edu Europa]
Ata] [...
<http://www.indire.it/> [more...](#) [Remove certification](#) | [Clip it](#)

Page 1 of 1

[About](#) [Privacy Policy](#) [Credits](#) [Help](#) [Contact Us](#)

Thanks
Questions?

Acknowledgment: we are here thanks to the Geek Librarian :-)

Case Studies Scholar



Project MeTe

“Metodi e percorsi per ricerche tematiche su internet”

Partnership Liceo Tambosi

Grant Fondazione Caritro 2006-2007

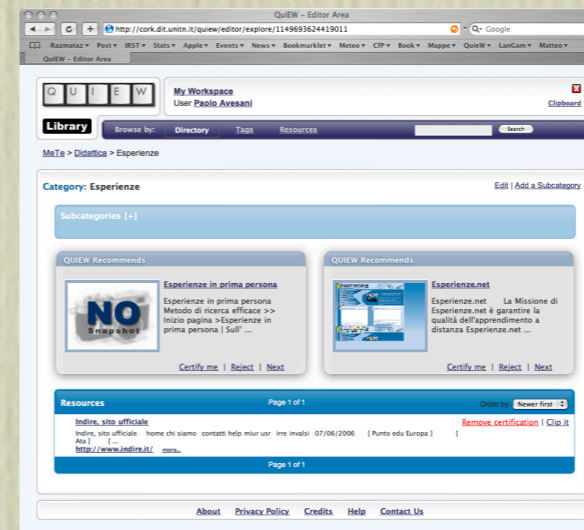
Assegnamento

Relazione

Bibliografia

Catalogazione

Metodo



Strumento


More

<http://quiew.itc.it>

QUIEW

GUI Preview

February 17th, 2006



We have designed the graphical user interface of QuieW prototype. It is conceived as a web application where to manage the discovery, the filtering, the classification, the organization, the ranking and the tagging of web resources.

Feel free to have a look at the full size snapshot clicking on the thumbnail above. Please leave a comment below on your first reaction (it has been proved that it is enough few milliseconds to recognize a user friendly interface).

Posted in [Case Studies](#), [Tools](#) | [Comments](#)

[« Previous Entries](#)

Main Contents

- [Project Overview](#)
- [Project Partners](#)
- [Project Team](#)
- [Documentation](#)
- [Contacts](#)

Post Categories

- [Case Studies](#)
 - [Library](#)
 - [School](#)
- [Documentation](#)
 - [Deliverables](#)
 - [Presentations](#)
 - [Publications](#)
 - [Reports](#)
- [Events](#)
 - [Conferences](#)
 - [Meetings](#)
 - [seminars](#)
- [Milestones](#)
- [Neighbours](#)
 - [Applications](#)
 - [Tools](#)
- [News](#)
- [Research](#)
 - [Quiew Filtering](#)
 - [Quiew Learning](#)
 - [Quiew Ranking](#)
- [Uncategorized](#)
- [Utilities](#)

[ITC](#) | [IRST](#) | [SRA](#) | [Login](#)

This site is not yet validated as XHTML 1.1 and CSS 2.0