

Impronte che non si cancellano

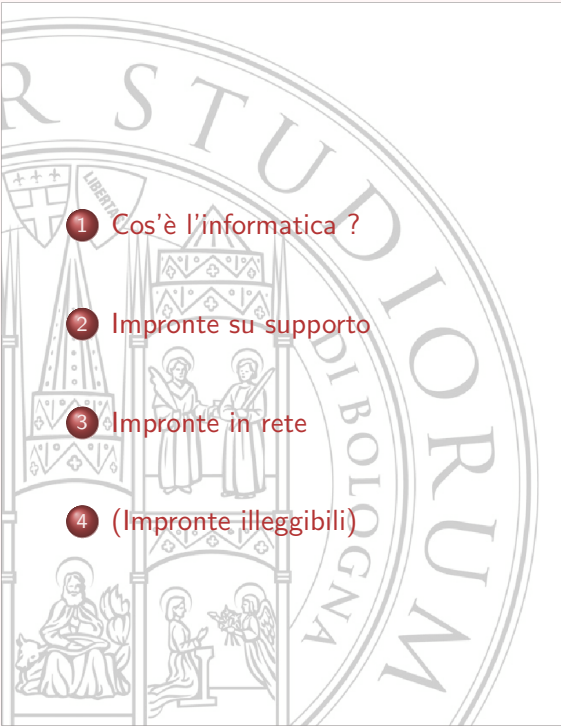
Internet, strumenti di memorizzazione e riservatezza

Simone Martini

Dipartimento di Scienze dell'Informazione
Alma mater studiorum • Università di Bologna

Riflessioni su Scienza e Società – Bologna, 13 aprile 2010

Outline

- 
- The background of the slide features a large, faint watermark of the University of Bologna seal. The seal is circular and contains the text 'R STUDI' at the top, 'IORUM' on the right, and 'BOLOGNA' at the bottom. In the center, there is a shield with a cross and three stars, and below it, a figure holding a book. The seal is partially obscured by the text of the outline.
- 1 Cos'è l'informatica ?
 - 2 Impronte su supporto
 - 3 Impronte in rete
 - 4 (Impronte illeggibili)

Outline

- 1 Cos'è l'informatica ?
- 2 Impronte su supporto
- 3 Impronte in rete
- 4 (Impronte illeggibili)



L'informatica è:

- un insieme di applicazioni?
- una tecnologia?
- una scienza?



L'informatica è:

- un insieme di **applicazioni**?
- una tecnologia?
- una scienza?



L'informatica è:

- un insieme di **applicazioni**?
- una **tecnologia**?
- una **scienza**?



L'informatica è:

- un insieme di **applicazioni**?
- una **tecnologia**?
- una **scienza**?

Una definizione comune

*L'informatica è la scienza che si occupa della rappresentazione, dell'elaborazione e della conservazione dell'**informazione**, e in generale di tutti gli aspetti del trattamento dell'informazione mediante procedure automatizzabili.*



L'informatica è:

Cos'è l'informatica?

- un insieme di **applicazioni**?
- una **tecnologia**?
- una **scienza**?

L'informatica è:

- una **scienza**
- che fonda una **tecnologia**
- che costruisce **applicazioni**



L'informatica è:

Cos'è l'informatica?

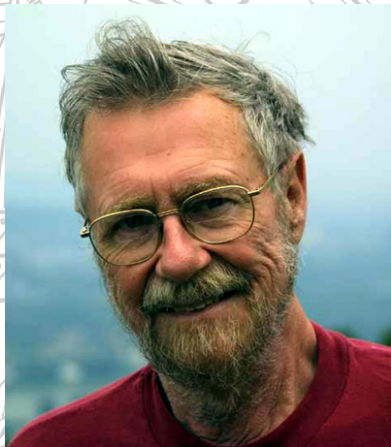
- un insieme di applicazioni?
- una tecnologia?
- una scienza?

L'informatica è:

- una scienza
- che fonda una tecnologia
- che costruisce applicazioni

Informatica e calcolatori

Edsger W. Dijkstra (1930 – 2002), Premio Turing 1972



Computer Science is no more about computers than astronomy is about telescopes

Informatica...

- Cos'è questa **informazione** che i calcolatori manipolano? Può essere misurata?
- Con quali strumenti **linguistici** avviene questo trattamento?
- Con quali tecniche (linguistiche e **algoritmiche**) possiamo risolvere problemi di gestione dell'informazione?
- Esistono dei **limiti intrinseci** (*"insiti nella natura"*) al trattamento dell'informazione mediante calcolatore?
- Quali **conseguenze sociali, etiche od economiche** ha un particolare uso dell'informatica?

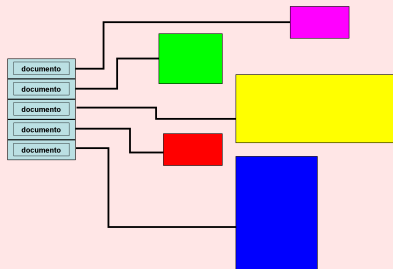
Outline

- 1 Cos'è l'informatica ?
- 2 Impronte su supporto**
- 3 Impronte in rete
- 4 (Impronte illeggibili)

Svuota il cestino...

Cancella solo l'*indice* delle informazioni

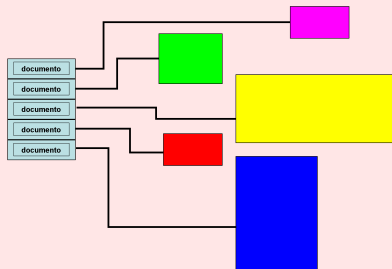
prima



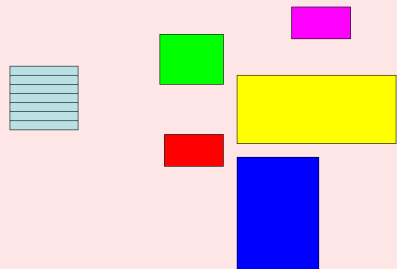
Svuota il cestino...

Cancella solo l'*indice* delle informazioni

prima



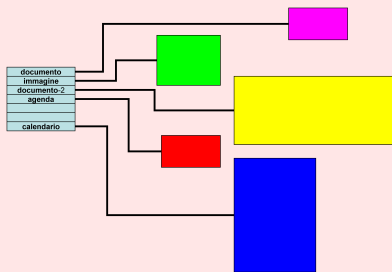
dopo



Riscrivi sopra...

- Una serie di zeri ?
- Una serie di numeri casuali ?

prima

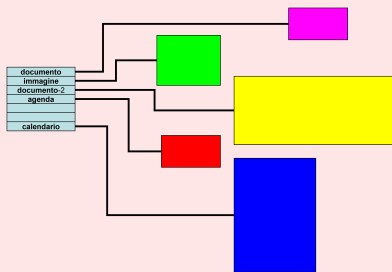


È sicuro ?

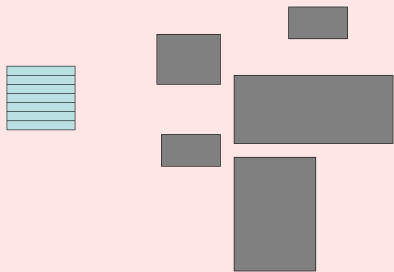
Riscrivi sopra...

- Una serie di zeri ?
- Una serie di numeri casuali ?

prima



dopo



È sicuro ?

A più basso livello

Usare il segnale analogico prima della digitalizzazione

Digital signal

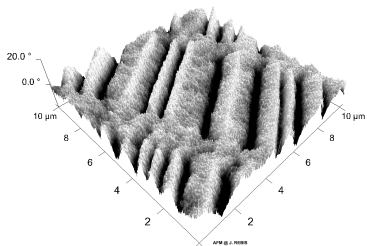


Analog signal

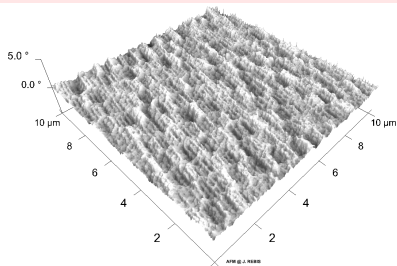


Le differenze tra il segnale analogico rilevato e quello teorico possono rivelare il contenuto cancellato

Il microscopio a forza magnetica



Disco 3.2GB



Disco 30GB

Il metodo di Gutmann

| | |
|----|----------------------------|
| 1 | (Random) |
| 2 | (Random) |
| 3 | (Random) |
| 4 | (Random) |
| 5 | 01010101 01010101 01010101 |
| 6 | 10101010 10101010 10101010 |
| 7 | 10010010 01001001 00100100 |
| 8 | 01001001 00100100 10010010 |
| 9 | 00100100 10010010 01001001 |
| 10 | 00000000 00000000 00000000 |
| 11 | 00010001 00010001 00010001 |
| 12 | 00100010 00100010 00100010 |
| 13 | 00110011 00110011 00110011 |
| 14 | 01000100 01000100 01000100 |
| 15 | 01010101 01010101 01010101 |
| 16 | 01100110 01100110 01100110 |
| 17 | 01110111 01110111 01110111 |
| 18 | 10001000 10001000 10001000 |
| 19 | 10011001 10011001 10011001 |
| 20 | 10101010 10101010 10101010 |
| 21 | 10111011 10111011 10111011 |
| 22 | 11001100 11001100 11001100 |
| 23 | 11011101 11011101 11011101 |
| 24 | 11101110 11101110 11101110 |
| 25 | 11111111 11111111 11111111 |
| 26 | 10010010 01001001 00100100 |
| 27 | 01001001 00100100 10010010 |
| 28 | 00100100 10010010 01001001 |
| 29 | 01101101 10110110 11011011 |
| 30 | 10110110 11011011 01101101 |
| 31 | 11011011 01101101 10110110 |
| 32 | (Random) |
| 33 | (Random) |
| 34 | (Random) |
| 35 | (Random) |

- Per specifici tipi di dischi, alcuni *pattern* sono più efficaci di altri
- 35 passate di riscrittura
- “ For any modern [. . .] drive, a few passes of random scrubbing is the best you can do. ”

Non c'è alcuna letteratura sulla effettiva ricostruzione di dati dopo uno *scrubbing*.

La soluzione dell'MI6

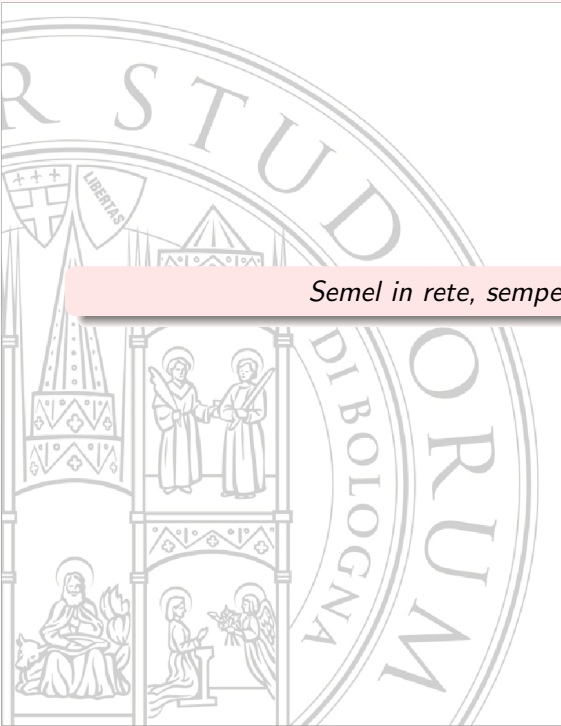
- per il Secret Intelligence Service di Sua Maestà Britannica...
- nessuna informazione è cancellata fino alla completa distruzione del suo supporto fisico

La soluzione dell'MI6

- per il Secret Intelligence Service di Sua Maestà Britannica...
- nessuna informazione è cancellata fino alla completa distruzione del suo supporto fisico

Outline

- 1 Cos'è l'informatica ?
- 2 Impronte su supporto
- 3 Impronte in rete**
- 4 (Impronte illeggibili)



Semel in rete, semper in rete

Un compagno di strada

Funes el Memorioso

*Jorge Luis Borges, 1942
in Ficciones, trad. F. Lucentini, Einaudi*



Internet

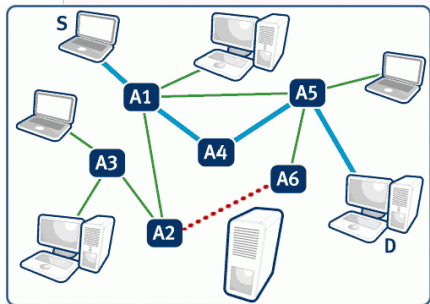
Internet è il sistema informativo globale che:

(i) è logicamente connesso mediante un unico spazio globale di indirizzi [...]

(ii) permette di supportare le comunicazione [...]

(iii) fornisce, utilizza o rende accessibili, in modo pubblico o privato, servizi ad alto livello sfruttando i livelli di comunicazione e le infrastrutture che sono stati descritti ai punti precedenti.

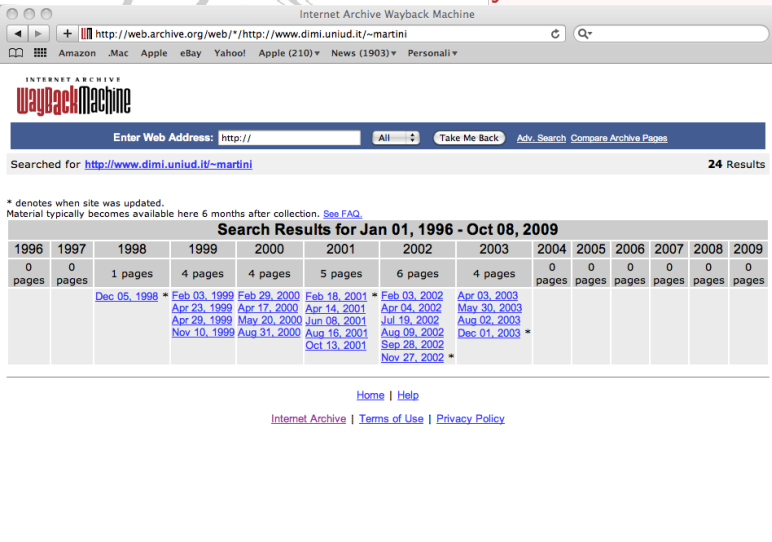
Federal Networking Council, 1995



Il Web

*Il World Wide Web (WWW, o Web)
è un gigantesco ipertesto multimediale distribuito,
basato su Internet
e dotato di un'interfaccia di facile uso.*

The Internet Archive's Way Back Machine



Internet Archive Wayback Machine

http://web.archive.org/web/*/http://www.dimi.uniud.it/~martini

Enter Web Address: All [Adv. Search](#) [Compare Archive Pages](#)

Searched for <http://www.dimi.uniud.it/~martini> **24 Results**

* denotes when site was updated.
Material typically becomes available here 6 months after collection. [See FAQ](#).

Search Results for Jan 01, 1996 - Oct 08, 2009

| 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
|---------|---------|--------------------------------|--|--|--|--|--|---------|---------|---------|---------|---------|---------|
| 0 pages | 0 pages | 1 pages | 4 pages | 4 pages | 5 pages | 6 pages | 4 pages | 0 pages | 0 pages | 0 pages | 0 pages | 0 pages | 0 pages |
| | | Dec 05, 1998 * | Feb 03, 1999 Apr 23, 1999 Apr 29, 1999 Nov 10, 1999 | Feb 29, 2000 Apr 17, 2000 May 20, 2000 Aug 31, 2000 | Feb 18, 2001 * Apr 14, 2001 Jun 08, 2001 Aug 16, 2001 Oct 13, 2001 | Feb 03, 2002 Apr 04, 2002 Jul 19, 2002 Aug 09, 2002 Sep 28, 2002 Nov 27, 2002 * | Apr 03, 2003 May 30, 2003 Aug 02, 2003 Dec 01, 2003 * | | | | | | |

[Home](#) | [Help](#)

[Internet Archive](#) | [Terms of Use](#) | [Privacy Policy](#)

La pagina di Simone a Udine: 1998

Simone Martini's Home


http://web.archive.org/web/19990203181855/www.diml.uniud.it/~martini/

Amazon .Mac Apple eBay Yahoo! Apple (210) News (1903) Personali




Simone Martini

Professor of Computer Science

[Dipartimento di Matematica e Informatica](#)
[Facoltà di Scienze MM.FF.NN.](#)
[Università di Udine](#)
Italy.



Guidoriccio da Fogliano
by another S.M.

Research Activities

- Type systems for programming languages
- Logic in computer science
- Lambda-calculus


[EACSL](#): The European Association for Computer Science Logic

[Fun!](#) The Italian Functional Programming Group

[A.I.L.A.](#) Associazione Italiana Logica e Applicazioni

A Ersilia, per stabilire i rapporti che reggono la vita della città gli abitanti tendono dei fili tra gli spigoli delle case, bianchi o neri o grigi o bianco-e-neri a seconda se segnano relazioni di parentela, scambio, autorità, rappresentanza. Quando i fili sono tanti che non ci si può più passare in mezzo, gli abitanti vanno via: le case vengono smontate; restano solo i fili ed i sostegni dei fili.

Italo Calvino; da: Le città invisibili



In Ersilia, to establish the links underlying the city life, people stretch strings across the corners of the houses, white or black or gray or black-and-white according to whether they stand for a relationship of blood, trade, authority, agency. When the strings are so many that one cannot pass through them any longer, people leave. The houses are dismantled; only the strings and their supports remain.

Italo Calvino; da: Le città invisibili

[Simone.Martini\(martini@diml.uniud.it\)](mailto:Simone.Martini(martini@diml.uniud.it))

La pagina di Simone a Udine: 2001

Simone Martinì's Home


http://web.archive.org/web/20010720093930/www.diml.uniud.it/~martini/

Amazon .Mac Apple eBay Yahoo! Apple (210) News (1903) Personali

Simone Martinì




Professor of Computer Science

[Dipartimento di Matematica e Informatica](#)
[Facoltà di Scienze MM.FF.NN.](#)
[Università di Udine](#)
Italy.



Guidoriccio da Fogliano
by another S.M.

[Laboratorio di Algoritmi e Strutture Dati](#)




Research Activities

- Type systems for programming languages
- Logic in computer science
- Lambda-calculus

[EACSL](#): The European Association for Computer Science Logic

[A.I.L.A.](#) Associazione Italiana Logica e Applicazioni

A Ersilia, per stabilire i rapporti che reggono la vita della città gli abitanti tendono dei fili tra gli spigoli delle case, bianchi o neri o grigi o bianco-e-neri a seconda se segnano relazioni di parentela, scambio, autorità, rappresentanza. Quando i fili sono tanti che non ci si può più passare in mezzo, gli abitanti vanno via: le case vengono smontate; restano solo i fili ed i sostegni dei fili.



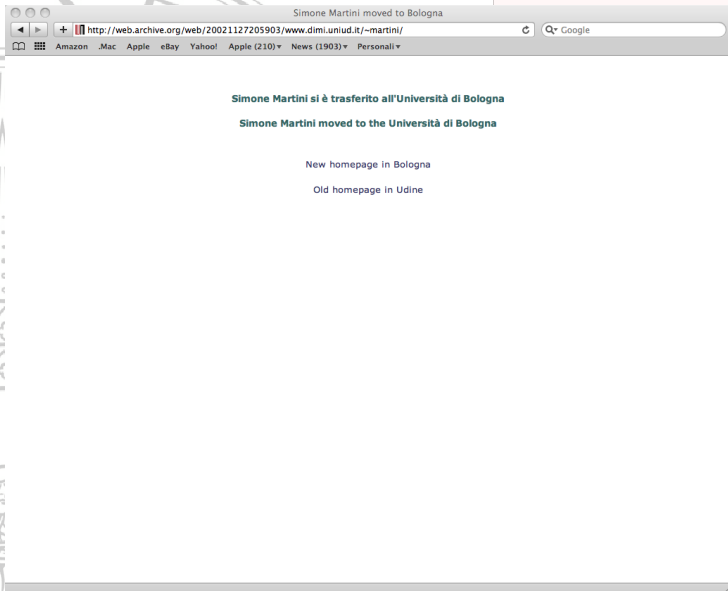
Italo Calvino; da: Le città invisibili

In Ersilia, to establish the links underlying the city life, people stretch strings across the corners of the houses, white or black or gray or black-and-white according to whether they stand for a relationship of blood, trade, authority, agency. When the strings are so many that one cannot pass through them any longer, people leave. The houses are dismantled; only the strings and their supports remain.

Italo Calvino; da: Le città invisibili

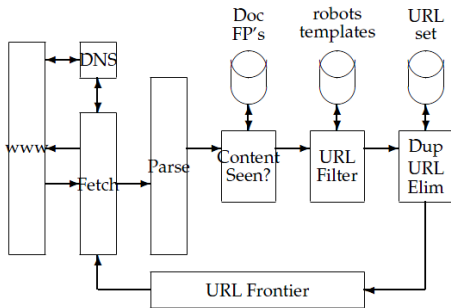
[Simone Martinì \(martini@diml.uniud.it\)](mailto:martini@diml.uniud.it)

La pagina di Simone a Udine: 2002



Robots

Ogni pagina è presto analizzata e digerita dai robot dei motori di ricerca, degli archivi, degli antiterrorismi, ecc.



The basic crawler architecture.

Mirrors

- Molte info sono replicate su mirror
- Foto e testi sono ripresi da altri

Chi mette informazione in rete non ha alcun controllo sul suo uso, duplicazione, memorizzazione

Oscurare un sito?

- “Oscurare” non è linguaggio tecnico
- Perché non si può davvero oscurare nulla, a meno di non distaccarlo dalla rete

Oscurare un sito?

- “Oscurare” non è linguaggio tecnico
- Perché non si può davvero oscurare nulla, a meno di non distaccarlo dalla rete

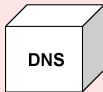
`http://www.abc.com ?`



Oscurare un sito?

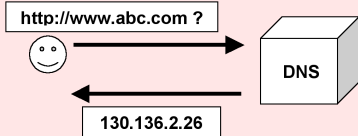
- “Oscurare” non è linguaggio tecnico
- Perché non si può davvero oscurare nulla, a meno di non distaccarlo dalla rete

http://www.abc.com ?



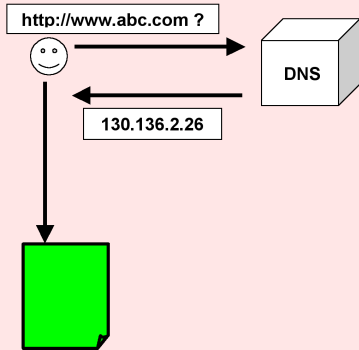
Oscurare un sito?

- “Oscurare” non è linguaggio tecnico
- Perché non si può davvero oscurare nulla, a meno di non distaccarlo dalla rete



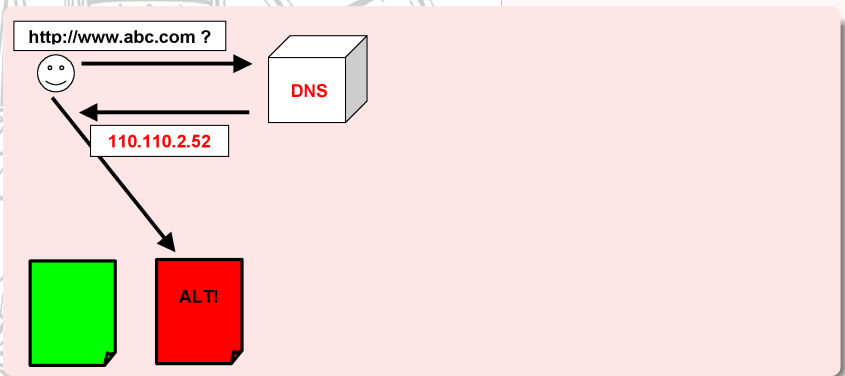
Oscurare un sito?

- “Oscurare” non è linguaggio tecnico
- Perché non si può davvero oscurare nulla, a meno di non distaccarlo dalla rete



Oscurare un sito?

- “Oscurare” non è linguaggio tecnico
- Perché non si può davvero oscurare nulla, a meno di non distaccarlo dalla rete



- Pubblicare sul web può essere molto più longevo che pubblicare su carta.
- Potenti strumenti di ricerca e aggregazione dei dati

HG: Today's newspapers will be lining tomorrow's waste paper bins.

JR: You really don't get it. This story gets filed. Every time anyone writes anything about me – they'll dig up these photos.
Newspapers last forever.

[Notting Hill, Screenplay by Richard Curtis, 1999]

Geographically-referenced data

una foto

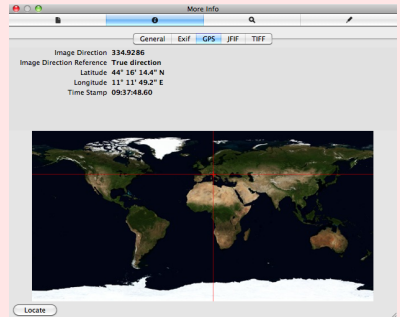


Geographically-referenced data

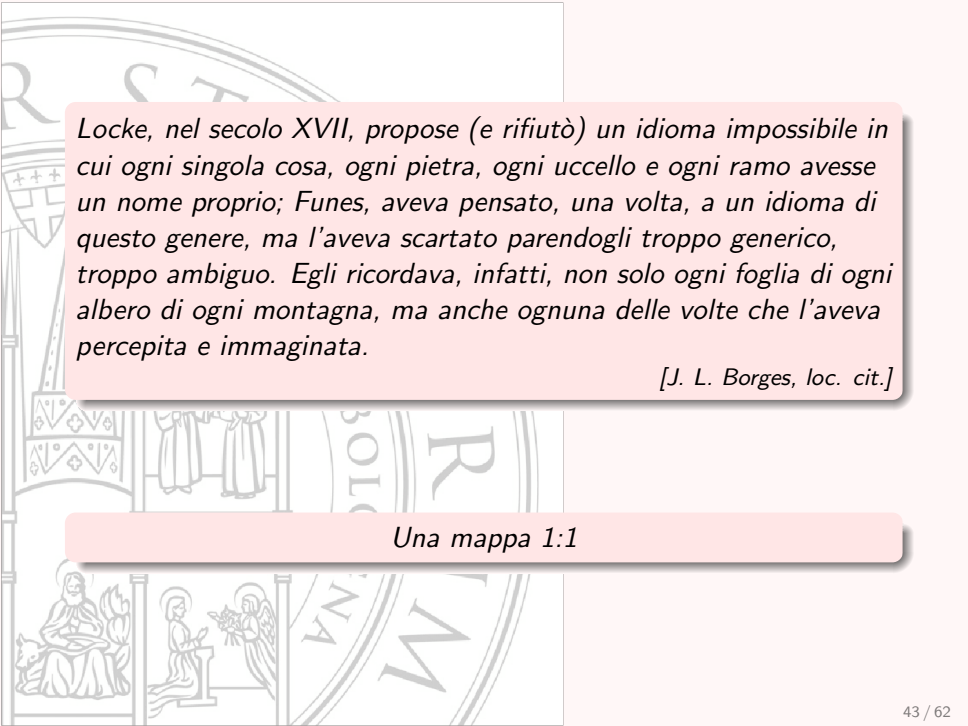
una foto



e alcuni dati collegati



Latitude: 44° 16' 14.4" N
Longitude: 11° 11' 49.2" E



Locke, nel secolo XVII, propose (e rifiutò) un idioma impossibile in cui ogni singola cosa, ogni pietra, ogni uccello e ogni ramo avesse un nome proprio; Funes, aveva pensato, una volta, a un idioma di questo genere, ma l'aveva scartato parendogli troppo generico, troppo ambiguo. Egli ricordava, infatti, non solo ogni foglia di ogni albero di ogni montagna, ma anche ognuna delle volte che l'aveva percepita e immaginata.

[J. L. Borges, loc. cit.]

Una mappa 1:1

Metadati

- I dati geografici sono un esempio di **metadati**
- Informazione testuale aggiuntiva
- Informazione “semantica”
- Permette ricerche, mappe concettuali, **ontologie**

Searchable representation

Il web è una rappresentazione della realtà a scala molto alta.

*È possibile farvi ricerche sofisticate
in modo efficiente e poco costoso.*

Searchable representation

Il web è una rappresentazione della realtà a scala molto alta.

*È possibile farvi **ricerche** sofisticate
in modo **efficiente** e **poco costoso**.*

Dati pubblici

- Dati raccolti da amministrazioni, aziende, ecc.
- Elenchi pubblici
- Dati forniti da noi in contesti diversi (newsgroups, blogs, social networks, etc.)

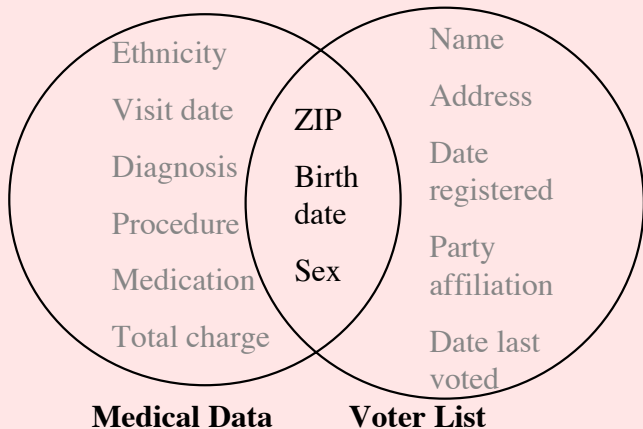


Figure 1 Linking to re-identify data

I cookies di terze parti

QWERTY! asdsadasdsa! asdsada!
FREE TRIAL!

Central Intelligence Agency

Home | Notices | Privacy | Security | Computers | Site Map | Index | Search | Feedback

About the CIA
About the CIA
CIA Vision, Mission, and Values
CIA Today
Virtual tour of the CIA
CIA Museum
George South, Clerk of Intelligence
H-2O Data for No Fear Act
CIA Frequently Asked Questions
Join the Update Service

USA Embassy
CIA Salary
Review FAFSA
Life at CIA
Current Vacating Ad
CIA's 5th Anniversary

Intelligence Program
Director of National Intelligence
Director of National Intelligence
Intelligence Community
National Intelligence Council

Central Intelligence Agency
Directorate of Intelligence
Directorate of Science & Technology
Center for the Study of Intelligence
NSA Electronic Reading Room
Office of General Counsel
Office of Military Affairs
Office of Public Affairs

If you submitted an on-line resume between December 7-9, 2005...

If you receive unsolicited e-mail appearing to be from the CIA...

What's New at the Central Intelligence Agency

cia.gov

QWERTY! asdsadasdsa! asdsada!
FREE TRIAL!

manual | discussion | edit | history | Log in / create account

MediaWiki

Manual:Parameters to index.php

This page is a **partial** list of the parameters to index.php, the main script of the MediaWiki software. Most of these arguments are usually given as GET parameters in the URL, but can also be passed as POST data. POST is actually required in some cases, such as the purge action for anonymous users.

Note: The information on this page is not complete

Contents [hide]

- Actions
- Preference overwriting
- Identifying a page or revision
- View and render
- History
- Flow
- Edit and submit

Navigation [Etc wiki]

- Main page
- Community portal
- Recent changes
- Help
- Manual
- FAQ
- Forum
- Support desk

Resources [elsewhere]

- IRC #mediawiki
- Mailing list
- Tag tracker
- Google SNIW
- Download from SVN

mediawiki.org

QWERTY! asdsadasdsa! asdsada!
FREE TRIAL!

lotofbanners.com



SOME RIGHTS RESERVED, GNU Free Documentation License

Incrociare...

- La “terza parte” vede le pagine visitate su siti diversi
- Le informazioni dei cookies sono anonime:
rivelano solo che un visitatore ha toccato una pagina
- Ma:
- Uno dei siti primari può avere informazioni sull'identità
- Se la passa alla “terza parte” ...
- Le info della terza parte sono identificate
- Legislazione in materia

Incrociare...

- La “terza parte” vede le pagine visitate su siti diversi
- Le informazioni dei cookies sono anonime:
rivelano solo che un visitatore ha toccato una pagina
- Ma:
- Uno dei siti primari può avere informazioni sull'identità
- Se la passa alla “terza parte” ...
- Le info della terza parte sono identificate
- Legislazione in materia

k-anonymity

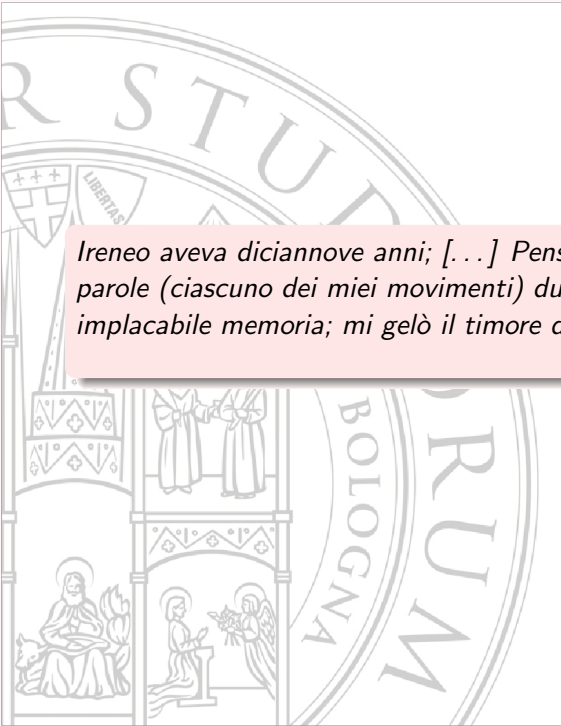
| | Race | Birth | Gender | ZIP | Problem |
|-----|-------------|--------------|---------------|------------|----------------|
| t1 | Black | 1965 | m | 0214* | short breath |
| t2 | Black | 1965 | m | 0214* | chest pain |
| t3 | Black | 1965 | f | 0213* | hypertension |
| t4 | Black | 1965 | f | 0213* | hypertension |
| t5 | Black | 1964 | f | 0213* | obesity |
| t6 | Black | 1964 | f | 0213* | chest pain |
| t7 | White | 1964 | m | 0213* | chest pain |
| t8 | White | 1964 | m | 0213* | obesity |
| t9 | White | 1964 | m | 0213* | short breath |
| t10 | White | 1967 | m | 0213* | chest pain |
| t11 | White | 1967 | m | 0213* | chest pain |

Figure 2 Example of k -anonymity, where $k=2$ and $QI=\{Race, Birth, Gender, ZIP\}$

L. Sweeney. k -anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10 (5), 2002; 557-570.

Ricordiamo tutto

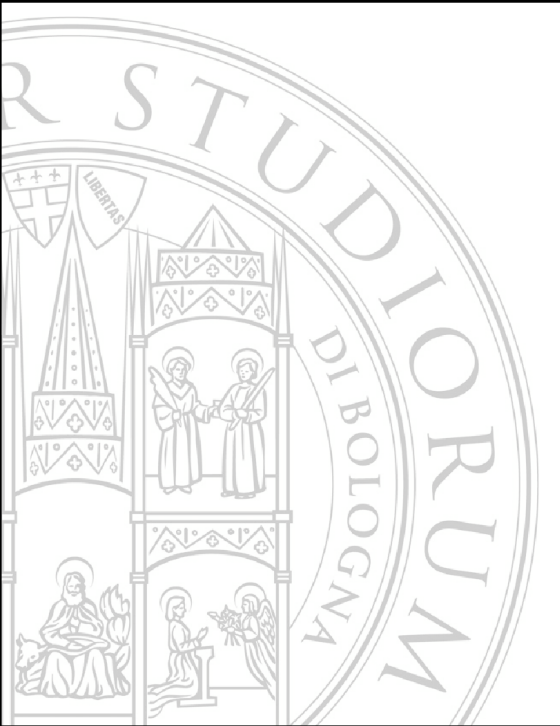
- Per secoli abbiamo avuto il problema di **come** e **cosa ricordare**
- Oggi abbiamo il problema di come **dimenticare!**
- C'è un **diritto all'oblio** che non sappiamo più garantire



Ireneo aveva diciannove anni; [...] Pensai che ciascuna delle mie parole (ciascuno dei miei movimenti) durerebbe nella sua implacabile memoria; mi gelò il timore di moltiplicare inutili gesti.
[J. L. Borges, loc. cit.]

Outline

- 1 Cos'è l'informatica ?
- 2 Impronte su supporto
- 3 Impronte in rete
- 4 (Impronte illeggibili)



In negativo...

Impronte che ~~non~~
si cancellano

e che vorremmo invece
mantenere!

Quando ero giovane...

Una vecchia tesi: 1987

Simone Martini

**MODELLI NON ESTENSIONALI DEL
POLIMORFISMO IN
PROGRAMMAZIONE FUNZIONALE**

Dottorato di Ricerca in Informatica
Primo Ciclo

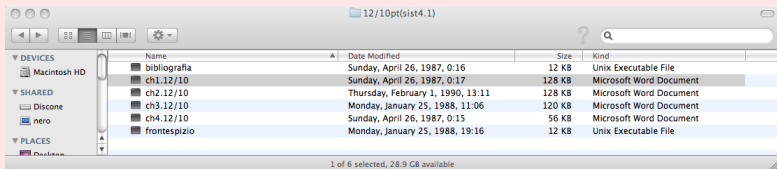
Quando ero giovane...

Una vecchia tesi: 1987

Simone Martini

**MODELLI NON ESTENSIONALI DEL
POLIMORFISMO IN
PROGRAMMAZIONE FUNZIONALE**

E i suoi files



The screenshot shows a Macintosh file browser window titled '12/10pt(sist4.1)'. The window displays a directory listing with columns for Name, Date Modified, Size, and Kind. The files listed are:

| Name | Date Modified | Size | Kind |
|--------------|-----------------------------------|--------|-------------------------|
| bibliografia | Sunday, April 26, 1987, 0:16 | 12 KB | Unix Executable File |
| ch1.12/10 | Sunday, April 26, 1987, 0:17 | 128 KB | Microsoft Word Document |
| ch2.12/10 | Thursday, February 1, 1990, 13:11 | 128 KB | Microsoft Word Document |
| ch3.12/10 | Monday, January 25, 1988, 11:06 | 120 KB | Microsoft Word Document |
| ch4.12/10 | Sunday, April 26, 1987, 0:15 | 56 KB | Microsoft Word Document |
| frontespizio | Monday, January 25, 1988, 19:16 | 12 KB | Unix Executable File |

The status bar at the bottom indicates '1 of 6 selected, 28.9 GB available'.

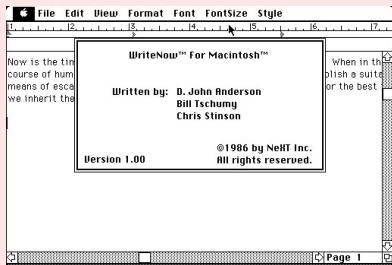
Quando ero giovane...

Una vecchia tesi: 1987

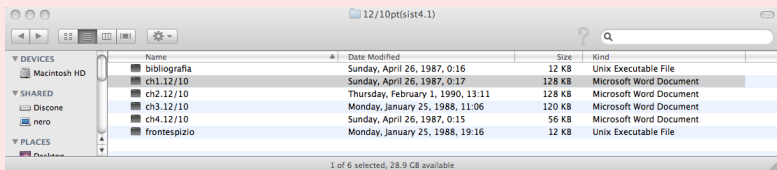
Simone Martini

**MODELLI NON ESTENSIONALI DEL
POLIMORFISMO IN
PROGRAMMAZIONE FUNZIONALE**

In formato WriteNow

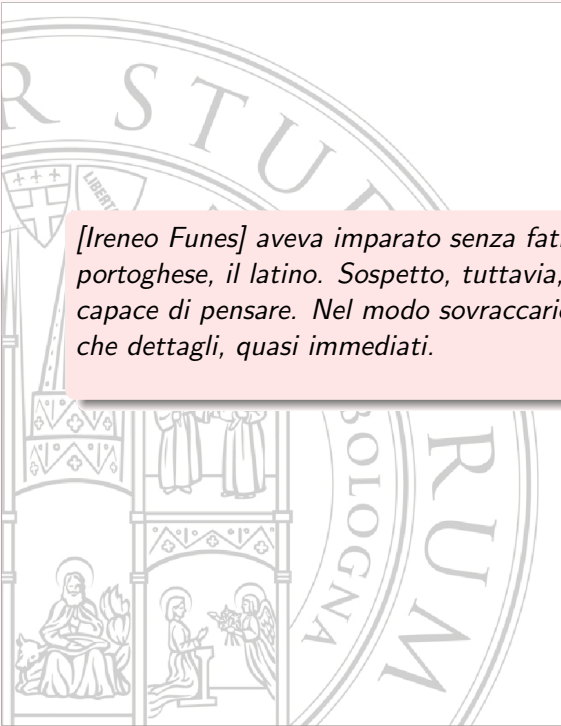


E i suoi files



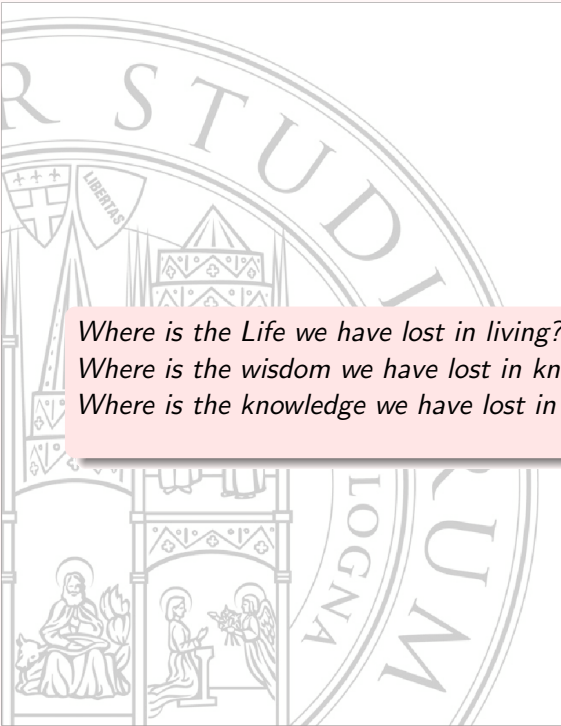
Concludendo

- Come cancellare informazione?
- Risorse con scadenza?
- Garantire il diritto all'oblio
- Garantire anonimato e privacy



[Ireneo Funes] aveva imparato senza fatica l'inglese, il francese, il portoghese, il latino. Sospetto, tuttavia, che non fosse molto capace di pensare. Nel modo sovraccarico di Funes non c'erano che dettagli, quasi immediati.

[J. L. Borges, loc. cit.]



E ancora
T. S. Eliot

Where is the Life we have lost in living?

Where is the wisdom we have lost in knowledge?

Where is the knowledge we have lost in information?

[T.S.Eliot, The Rock, Chorus I]