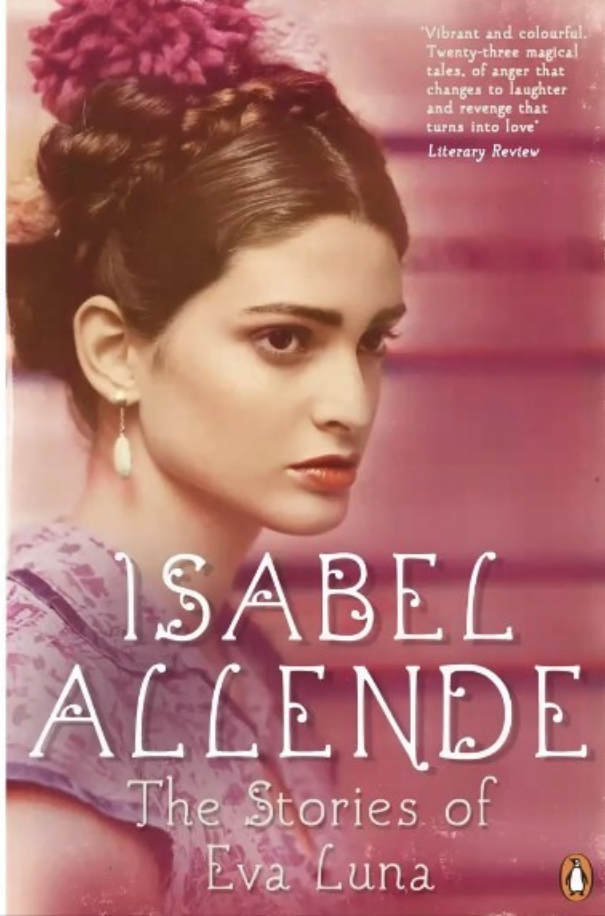


Names

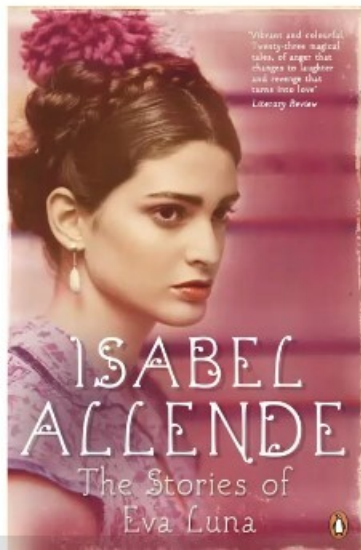
◆ "Two words" tale
(read it if you can)

'Vibrant and colourful.
Twenty-three magical
tales, of anger that
changes to laughter
and revenge that
turns into love'
Literary Review



Names

- ◆ Let's talk about names on the web, both sides: ***social*** and ***technical***



Web names: social side

- ◆ How to choose a ***good web address*** ?
- ◆ There are some rules that can maximize the potential success of the site
- ◆ How much? (read as: should we care about a name?? Can it really impact a website??)
- ◆ Answer: on average **10-20%**, with variance up to **50%** (!!)

Rule 1

- ◆ Shorter names work much better than longer names

Rule 2

- ◆ The name should be unique, sufficiently separatable from other names (that is: it should not be confounded with other names!)

Rule 3

- ◆ Take the .COM (!!!!)
- ◆ Impact: about **+4.5%**

Rule 4

- ◆ It should be easy to memorize and write

Rule 5

- ◆ (overlapping with rule n.4):
better to choose existing words rather than creating new ones
- ◆ In any case, if new words, or acronyms, are used, the important factor is the ratio between the «standard» words and the new ones: the higher the better
- ◆ Range: from +1.5% to **-5%** (!)

Rule 6

◆ Beware the sound...!

The beginning of a name....

- ◆ Names that start with a vowel work well (about **+3.7%**)
- ◆ Semi-vowel (r, j, y, w) work well (about +2.9%)
- ◆ Surprise: consonants like f, v, s, z work even better (**+3.3%**)

Sound (cont.)...

- ◆ Sounds associated with bad words in English (like the phonem «uh», like in «yuck») damage the site up to a – **44%** (!)
- ◆ In other contexts though, for instance adult material (!), they give advantages up to **+7%** (!!)

Ruke 8

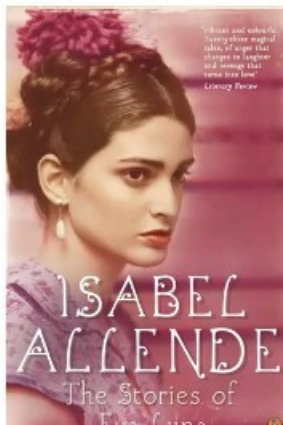
- ◆ No numbers
(rule that is also present in the Internet)
- ◆ **FALSE!**
- ◆ In fact: use numbers!
(Impact: about **+8.2%** !!!!)

Rule 0

- ◆ Be careful on how you check whether a domain is free or not... (!!!!!!!!!)
- ◆ Suggestion: **internic**

URI!!

◆ **Uniform Resource Identifier**



URI and Web addresses

◆ URI == web addresses http://... ?

◆ **NO!**

◆ URIs are a ***superset*** of the web addresses http://....

Web addresses

- ◆ Common use for instance to say `www.mysite.com`
- ◆ This is not a URI!
- ◆ And in fact... the form above doesn't exist in any standard (!)
- ◆ It's the browsers that, for convenience, translate this "web address" (for the people) into `http://www.mysite.com/`

Examples of URIs

- ◆ `http://corsi.math.unipd.it/wim`
- ◆ `news:it.culture`
- ◆ `telnet://example.net:453`

URLs

- ◆ Uniform Resource **LOCATOR**
- ◆ URLs are those URI that identify the resource via a representation of their primary ***access mechanism***
- ◆ So, yes, for instance web addresses like "http://...." are URLs

URNs

- ◆ Uniform Resource **NAMES**
- ◆ URNs are those URIs that stay unique and persistent even when the resource doesn't exist any more or it not available any more

The Schema

- ◆ Defines the semantics (the meaning) of the URI
- ◆ For instance
- ◆ In `http://corsi.math.unipd.it/wim`
- ◆ ... the schema is **http**
- ◆ → This is a URI whose semantics is set by the formal specification describing the http schema

Hierarchical URIs

◆ Typical general form:

◆ **Schema :// authority path ? query**

Authority

- ◆ Literally, the authority!!
- ◆ The element that states that URI is under control of a certain authority



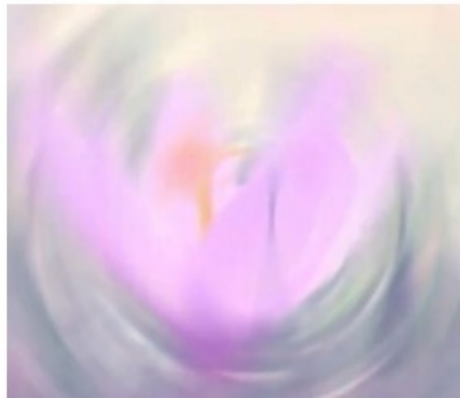
Query

- ◆ Information interpreted by the resource ("input parameters")



Example

- ◆ `mailto:director@cnn.com`
- ◆ `mailto` → `schema`
- ◆ `director@cnn.com` → `opaque_part`



More URIs

◆ **TELEPHONE** (/FAX etc etc)

◆ Rfc2806

◆ Examples:

◆ **tel:+358-555-1234567**

Beware...



- ◆ To use relative URIs that are completed into something we don't want
- ◆ Example of one of the most common web errors: in our `http://www.mysite.com` we insert a link to "`www.othersite.com`"
- ◆ → completed into "`http:www.mysite.com/www.othersite.com`"

Example (cont.)...



- ◆ Within `http://www.mysite.com`, we insert in a web page the email link to `"shop@mysite.com"`
- ◆ → completed into `"http:www.mysite.com/shop@mysite.com"`

URNs

- ◆ Uniform Resource **NAMES**
- ◆ Typical form:
urn:*NID* :...

ISBN as URNs

◆ Rfc3187

◆ URN:ISBN:0-395-36341-1



New Zealand...

◆ Has been the first state to ask and obtain a URN of its own (!)

◆ **urn:nzl:...**



Beyond URI

- ◆ URI: URLs URNs etc
- ◆ But.... what **encoding**?
- ◆ ASCII...(!) ☹
- ◆ → **IRI**
(Internationalized
Resource Identifiers)

実
神
活
愛
家
音

IRI and security...

- ◆ IDN (Internazionalied Domain Names), even in extensions: example .pф (.ru)
- ◆ ***Homograph attack***
(rnicrosoft.com, G00GLE.COM...)
- ◆ → enormously powered up with IRIs
(es. Cyrillic, Greek etc)

wikipedia.org

Back to URI: problem

◆ What is the data format of
<http://www.sito.it/a/b.html>?
(Italian? English? French?)



The Answer

- ◆ That information is not part of the URL/URI: the web address is a black-box string, and every property depends on the schema, not on the string (!)
- ◆ HTTP provide methods (***content negotiation***) to transmute the correct data format. Every other algorithm (like using the final extension of the URL, like .txt, .gif etc) is not 100% reliable dell'URL non è attendibile.

“TLD overkill”

- ◆ Create a new TLD to solve a problem

Example of problem:

- ◆ In the Internet it's full of crazy perverts, lots of sex and vulgar people!!



■ ■ ■
VH1 #DATINGNAKED

THURS JULY 17 + 9/8C



And dually...

◆ Proposal to add a **.kids** (“secure” site)



Other proposal

- ◆ **".adult"**

- ◆ → more generic, "adult website"



Proposal... by whom?

- ◆ ICM Registry, California-based company
(Palm Beach Gardens)



So we can implement this two ways...

- ◆ [“light”] Adding domains with these new TLDs, or
- ◆ [“strong”] Force all the content of a specific kind into the corresponding TLD



Costs?

- ◆ Technological cost: **SUPER LOW!**
- ◆ Social cost: ... (!?)

Pro/cons (“light” case)

◆ **Pro:** ???

◆ **Con:** companies will (maybe) buy another .sex domain, keeping the existing one

◆ *Question:* But that's just a transition, eventually they will all migrate, no?

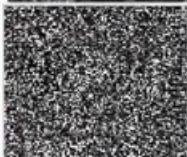
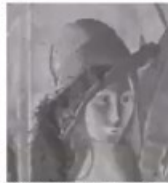
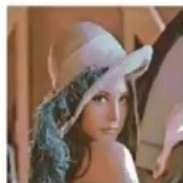
◆ *Answer:* no! (tracking/filtering)

Example...



What is this?

- ◆ The most famous test image used in computer graphics (great mixture of details, "flat" regions, shading, textures)
- ◆ Time ago it even came included in every graphic program (like for instance within Windows with Paint)



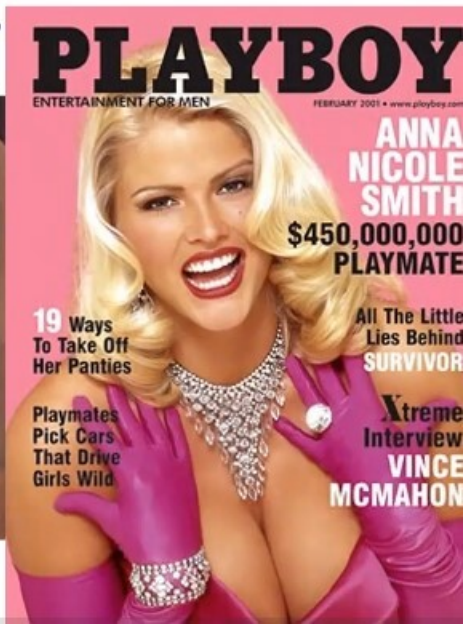
The lady?

- ◆ Lena Soderberg (ne Sjööblom), Swedish, happily living there, married with three children, working in the state monopoly and helping as a volunteer for disabled people



Playboy!

◆ ... scanned from the big centerfold page of Playboy (1972!!) and ..



... “cropped”...



And so...



- ◆ In 1996 someone curious finds out
- ◆ → petition to ban forever that (for instance from all articles in the **IEEE TRANSACTIONS ON IMAGE PROCESSING**)
- ◆ Formally never “forbidden”, but “strongly suggested” not to use it → people started to use it less and less → taken out from all graphic programs

Boundaries...

- ◆ So what is porno and what is not? (Renoir, Picasso, Dali', Botticelli, a Fellini movie?)



Douglas Crockford

- ◆ Co-inventor of Javascript, JSON etc... (!)



JSMin

- ◆ The reference library to minimize Javascript
- ◆ License, free software with an added sentence:
- ◆ "The Software shall be used for Good, not Evil"



Result

- ◆ Infinite legal battles and argumentations... (!), whole operating systems incorporating that library at risk... (!!!)



SQLite

- ◆ “May you do good and not evil
May you find forgiveness for yourself
and forgive others
May you share freely, never taking more
than you give.”

SQLite

Back to URL: the moral

- ◆ Using URLs for scopes that go beyond, forcing information that doesn't belong there, is just stupid!



JSMin

- ◆ The reference library to minimize Javascript
- ◆ License, free software with an added sentence:
- ◆ "The Software shall be used for Good, not Evil"



And in fact...

- ◆ 18 March 2011...
- ◆ **.xxx** has been approved....
- ◆ Real effect?
- ◆ 200 milion dollars a year more in the bank account of **ICM Registry** (...!!)

