



# The Italian approach to Internationalized Domain Names (IDNs)

**Dr. Domenico Laforenza**

Director of the Institute for Informatics and Telematics  
Italian National Research Council, Pisa, Italy

**W3C Workshop**

**Content on the Multilingual Web**

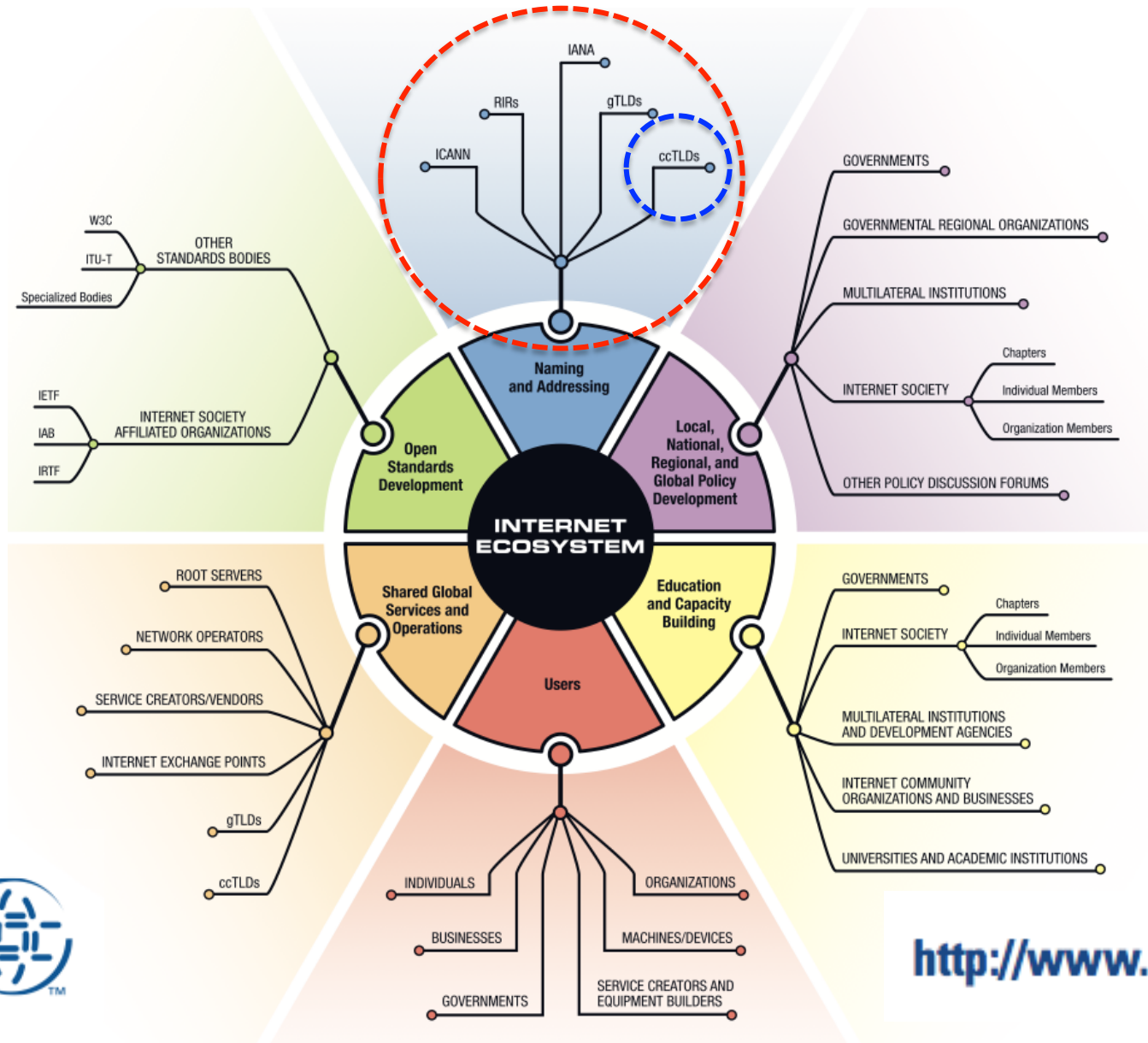
4 - 5 April 2011, Pisa, Italy



Consiglio Nazionale delle Ricerche

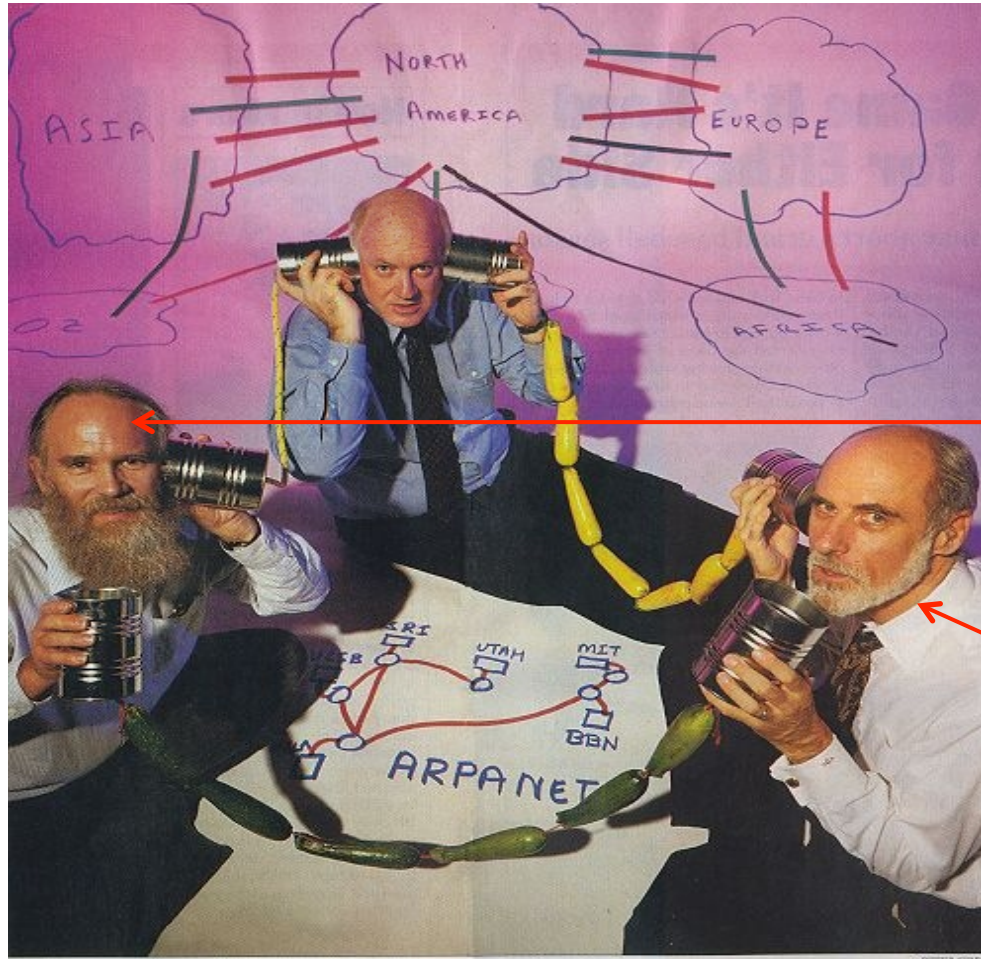
- ❖ The Internet Ecosystem
- ❖ Domain Name System
- ❖ gTLDs and ccTLDs
- ❖ .it Registry
- ❖ Internationalized Domain Names (IDNs)
- ❖ **IDNs under .it**

# The Internet Ecosystem



<http://www.isoc.org>





This photo was published in the **August 8, 1994** issue of Newsweek and commemorates the 25th anniversary of the ARPANET. **Jon Postel, Steve Crocker** and Such was the state of networking in the primitive 1960s...

*Picture from Vint Cerf*

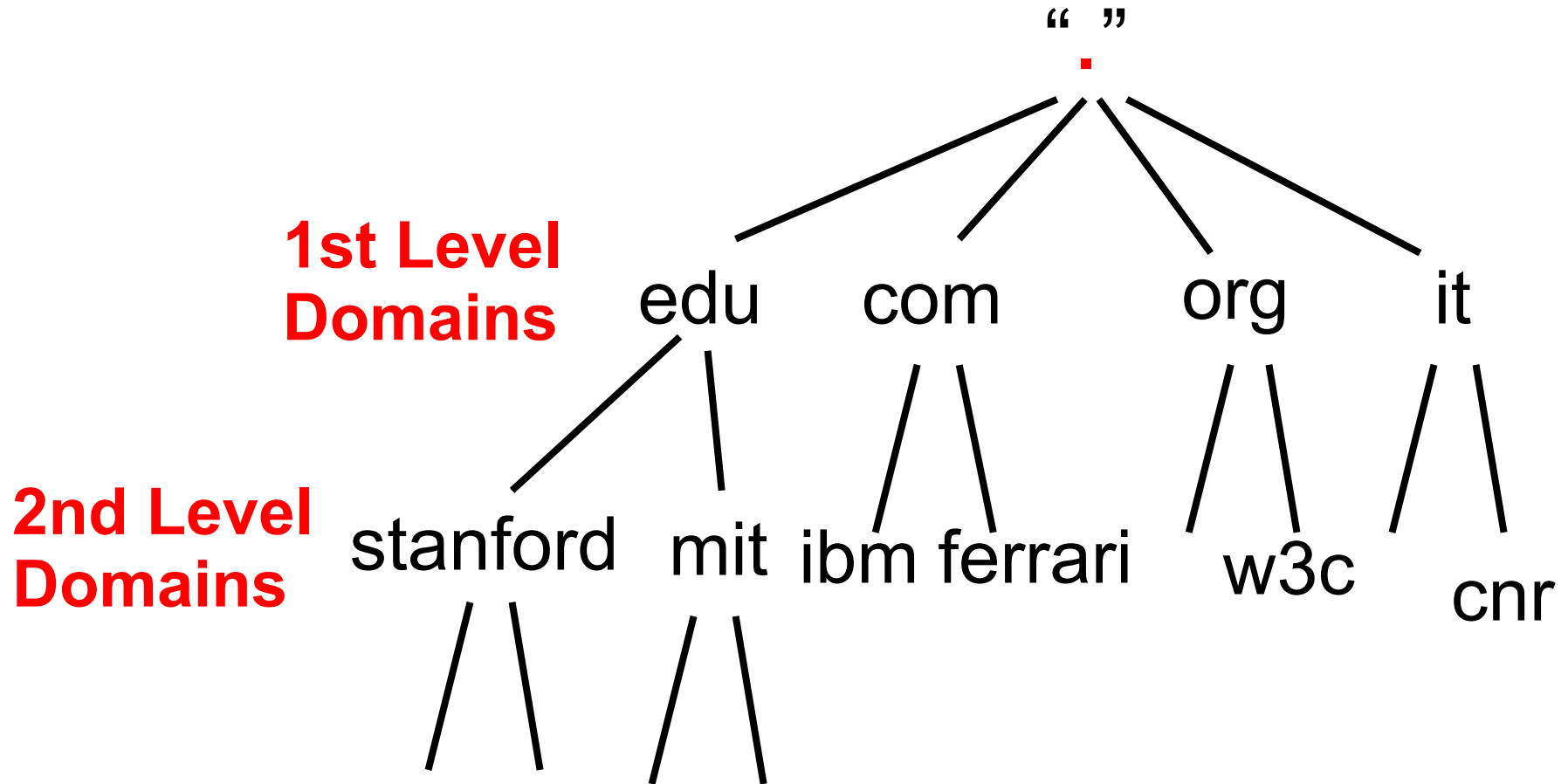


- Every TCP/IP resource can be assigned with a name (*symbolic name*)
- A mechanism to associate the symbolic name of a host to its IP address(es): *direct resolution*
- A mechanism to associate the IP address of a host to its symbolic name: *inverse resolution*
- **Domain Name System (DNS)**
  - Defined by ISI - USC 1984
  - RFC 882, RFC 883, RFC 973 (obsolete)
  - RFC 1034, RFC 1035, RFC 1123, RFC 1537, RFC 1912, RFC 2182 and their updates

- Namespace organized according to a hierarchical model:
  - the DNS database logical structure is like a “**overturned tree**”
  - Each tree node represents a **domain**
  - Each domain can have children: the **sub-domains**
  - Each node is identified by a **label**

The root of the tree is unique and its label is empty. Generally the root is identified by a “.”

- **Namespace structure (RFC 1591)**:
  - **g**eneral **T**op **L**evel **D**omains (**gTLDs**)
  - **c**ountry **c**ode **T**op **L**evel **D**omains (**ccTLDs**)
  - Domains for the inverse resolution (arpa)





The first level domains are classified in gTLDs and ccTLDs

## Historical **g**eneric **T**op **L**evel **D**omains:

- ✓ **com** commercial organizations
- ✓ **edu** USA university and research
- ✓ **gov** USA governative organizations
- ✓ **mil** USA military organizations
- ✓ **net** open to everyone
- ✓ **org** open to everyone
- ✓ **int** international organizations, treaties

<http://www.w3c.org>

Second Level

First Level

# Registro | gTLDs

## **aero**

- ❖ for airline companies - Societe Internationale de Telecommunications Aeronautiques SC, (SITA (<http://www.nic.aero>))

## **biz**

- ❖ for companies - NeuLevel, Inc. (<http://www.nic.biz>)

## **info**

- ❖ “unrestricted use” - Afilias Lt (<http://www.nic.info>)

## **museum**

- ❖ for museums - Museum Domain Management Association, (MDMA) (<http://musedoma.museum>)

## **name**

- ❖ for individuals - Global Name Registry, Ltd (<http://www.name>)

## **pro**

- ❖ for freelances - RegistryPro, Ltd (<http://www.registry.pro>)

## jobs

- ✓ for companies dealing with human resources management - Employ Media LLC: <http://www.goto.jobs>

## mobi

- ✓ for mobile and related services companies - mTLD Top Level Domain, Ltd.: <http://pc.mtld.mobi>

## travel

- ✓ for travel agencies - Tralliance Corporation (<http://www.tralliance.info>)

## cat

- ✓ for the Catalan cultural and linguistic community - Fundació puntCat: <http://www.domini.cat>

## xxx

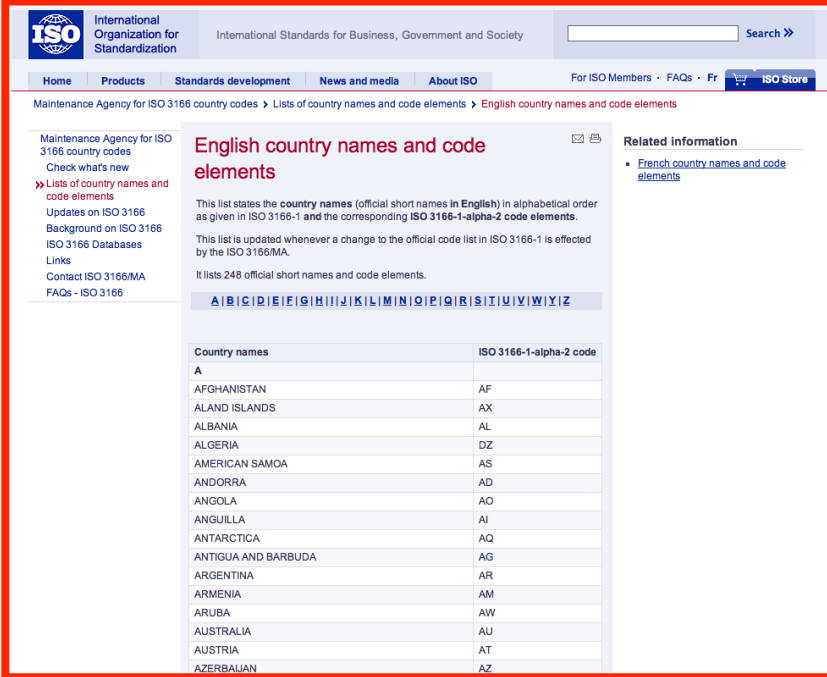
- ❖ for pornographic sites (ICANN approved on 18 March 2011)



A **country code Top-Level Domain** is an Internet top-level domain used or reserved for a country, a sovereign state, or a dependent territory

National domains, represented by the ISO 3166-1 codes:

- ✓ **.al** Albania
- ✓ **.de** Germany
- ✓ **.it** Italy
- ✓ **.uk** Great Britain
  
- ✓ .....



Maintenance Agency for ISO 3166 country codes > Lists of country names and code elements > English country names and code elements

### English country names and code elements

This list states the country names (official short names in English) in alphabetical order as given in ISO 3166-1 and the corresponding ISO 3166-1-alpha-2 code elements.

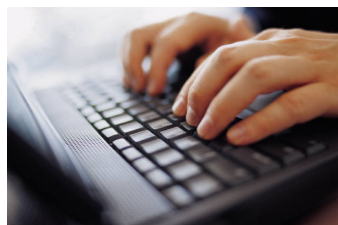
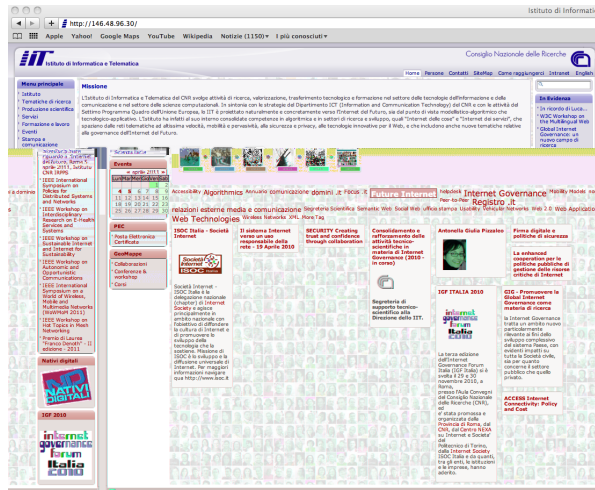
This list is updated whenever a change to the official code list in ISO 3166-1 is effected by the ISO 3166/MA.

It lists 248 official short names and code elements.

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [Y](#) [Z](#)

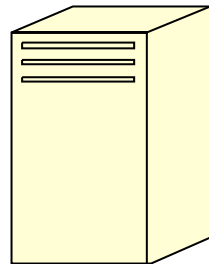
Country names	ISO 3166-1-alpha-2 code
<b>A</b>	
AFGHANISTAN	AF
ALAND ISLANDS	AX
ALBANIA	AL
ALGERIA	DZ
AMERICAN SAMOA	AS
ANDORRA	AD
ANGOLA	AO
ANGUILLA	AI
ANTARCTICA	AQ
ANTIGUA AND BARBUDA	AG
ARGENTINA	AR
ARMENIA	AM
ARUBA	AW
AUSTRALIA	AU
AUSTRIA	AT
AZERBAIJAN	AZ

## DNS Query: An example

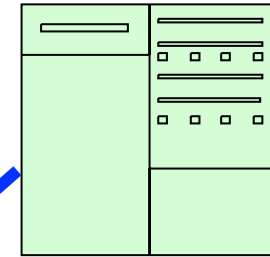
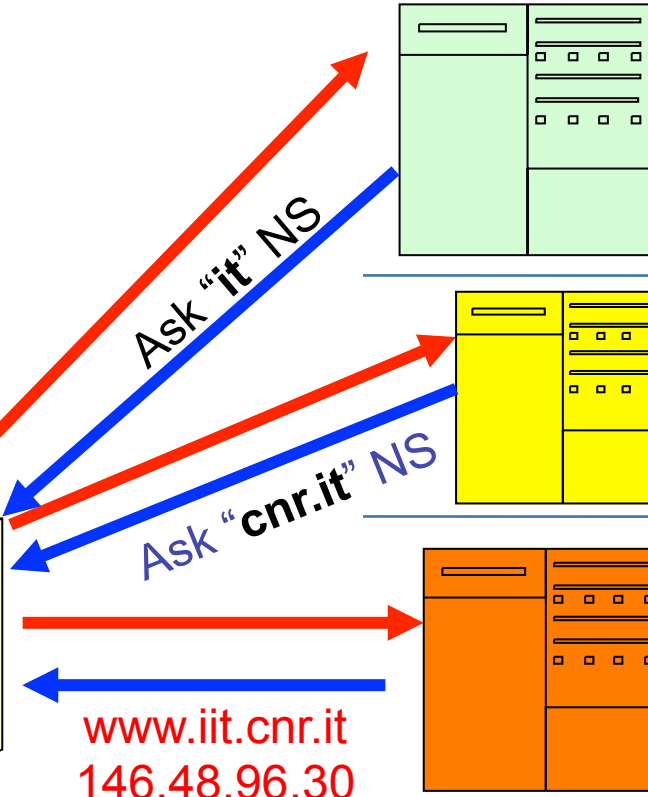


www.iit.cnr.it

www.iit.cnr.it  
146.48.96.30



Recursive Resolver



Root Server

“ “  
.”

IT

COM

“it” Name Server

CNR

UNIPI

“cnr.it” Name Server

IIT

BO

WWW

## Domain Name Resolution Process

[HOME](#)[ABOUT US](#)[EVERYTHING ON .IT](#)[CREATE AND CHANGE .IT](#)[LEGAL](#)[MAINTAINERS](#)[REGISTRARS](#)

You are here: [Home](#)

## ABROAD .IT ON THE UP



Demand for dot it domains in on the up and up, even abroad.

Growth in the demand for dot it domains: in the space of a years the number of dot it domains registered by companies outside Italy has more than doubled.

Download [Focus .it](#)

[CREATE YOUR DOMAIN](#)[WHOIS](#)[FIND A MAINTAINER OR A REGISTRAR](#)[FAQ](#)[FOCUS](#)[.IT REGISTERED DOMAINS](#)

Apr 03, 2011 12:00 AM

2.149.332

[NEWS](#)[RSS](#)

From 1 July 2010 it will no longer be possible to draw up a new Maintainer contract »

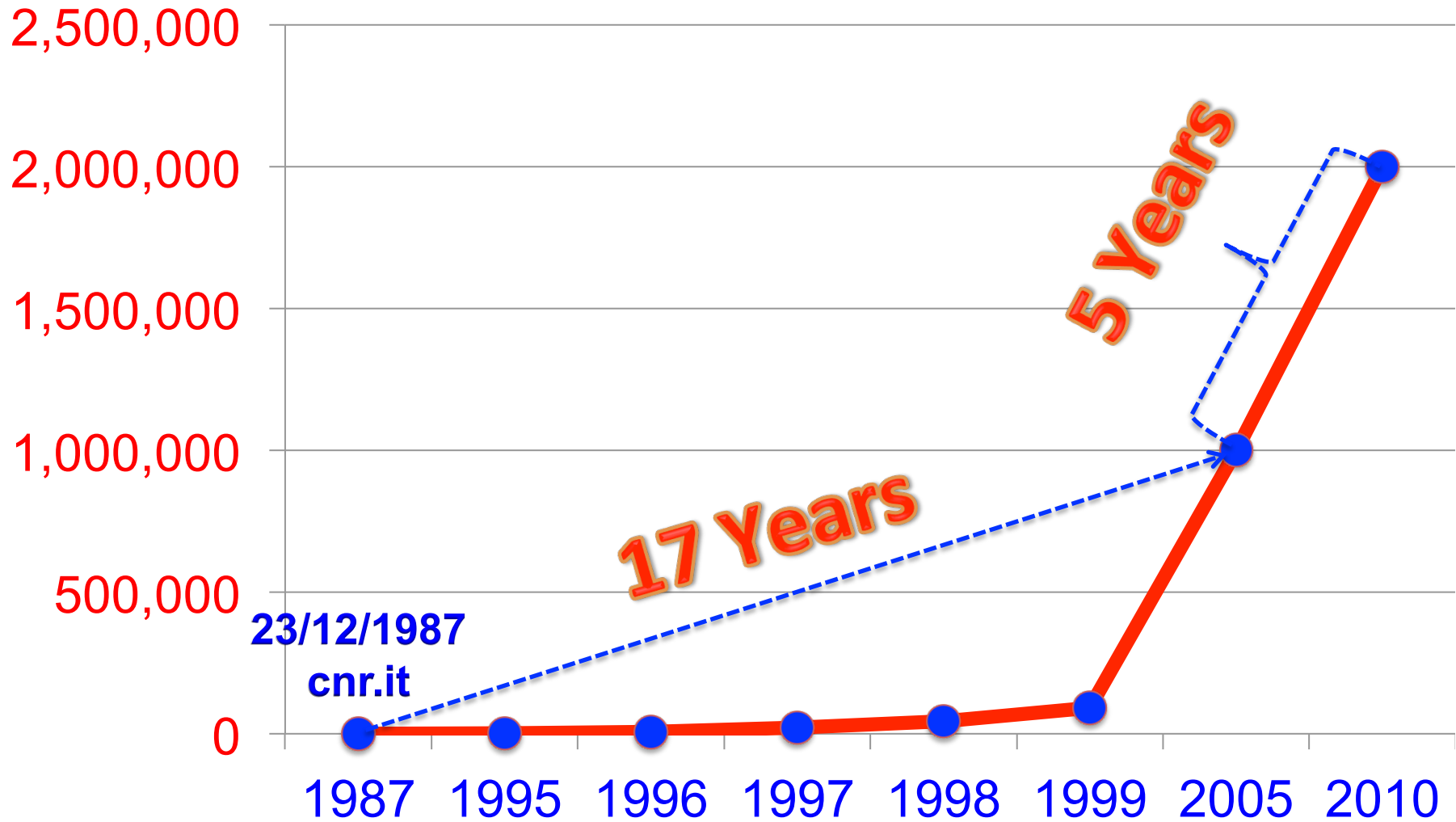
2009 Yearbook of .It Registry online »

[DIGITAL NATIVES](#)





# .it Domain Names



# Pisa, October 22nd, 2010



## TOP CCTLD REGISTRIES BY DOMAIN NAME BASE, SECOND QUARTER 2010

1. .de (Germany)
2. .uk (United Kingdom)
3. .cn (China)
4. .nl (Netherlands)
5. .eu (European Union)
6. .ru (Russian Federation)
7. .br (Brazil)
8. .ar (Argentina)
9. .it (Italy)
10. .tk (Tokelau)

Source: Zooknic, August 2010

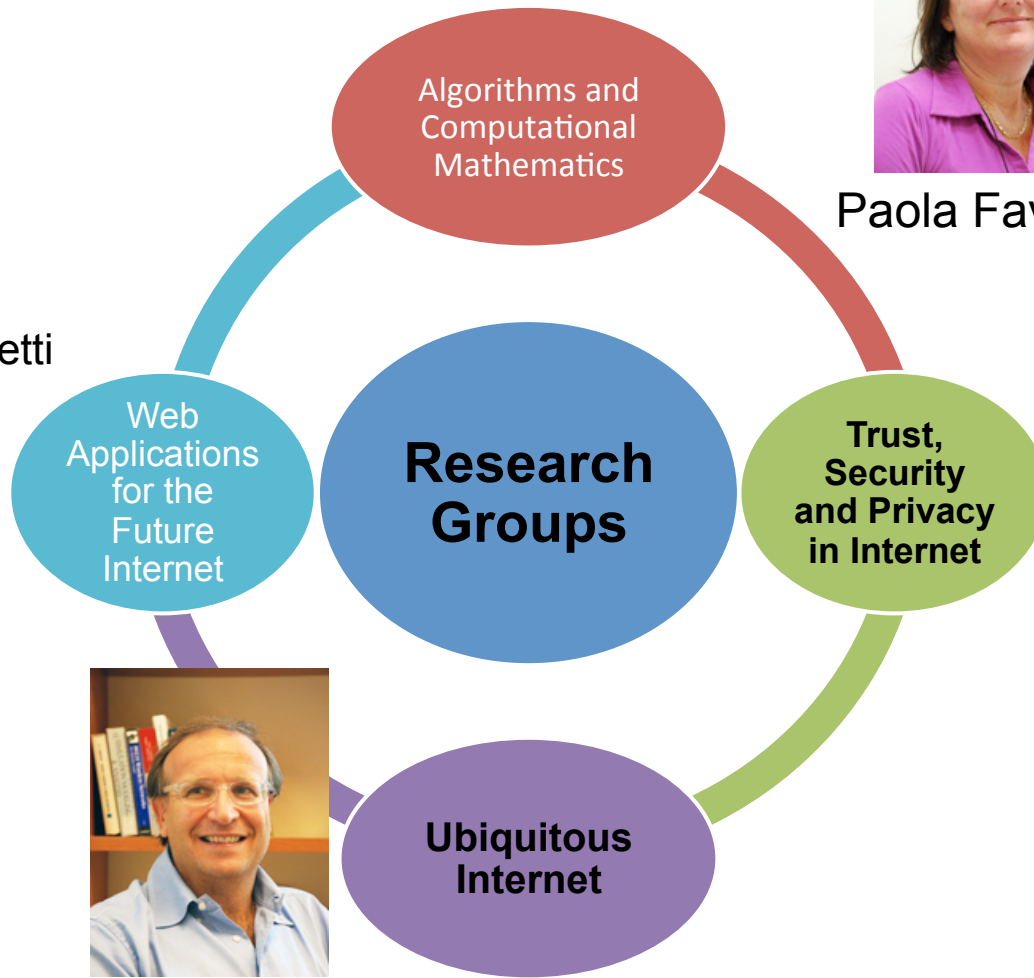
- Carry out research, development, technology transfer and training in the fields of information and communication technologies and computational science
- In line with the strategy of the:
  - ❖ CNR ICT Department
  - ❖ EU FP7
- IIT is naturally and concretely projected towards the **Internet of the Future**



# The Institute for Informatics and Telematics (IIT-CNR) **Research Groups**



Andrea Marchetti



Paola Favati



Paolo Mori

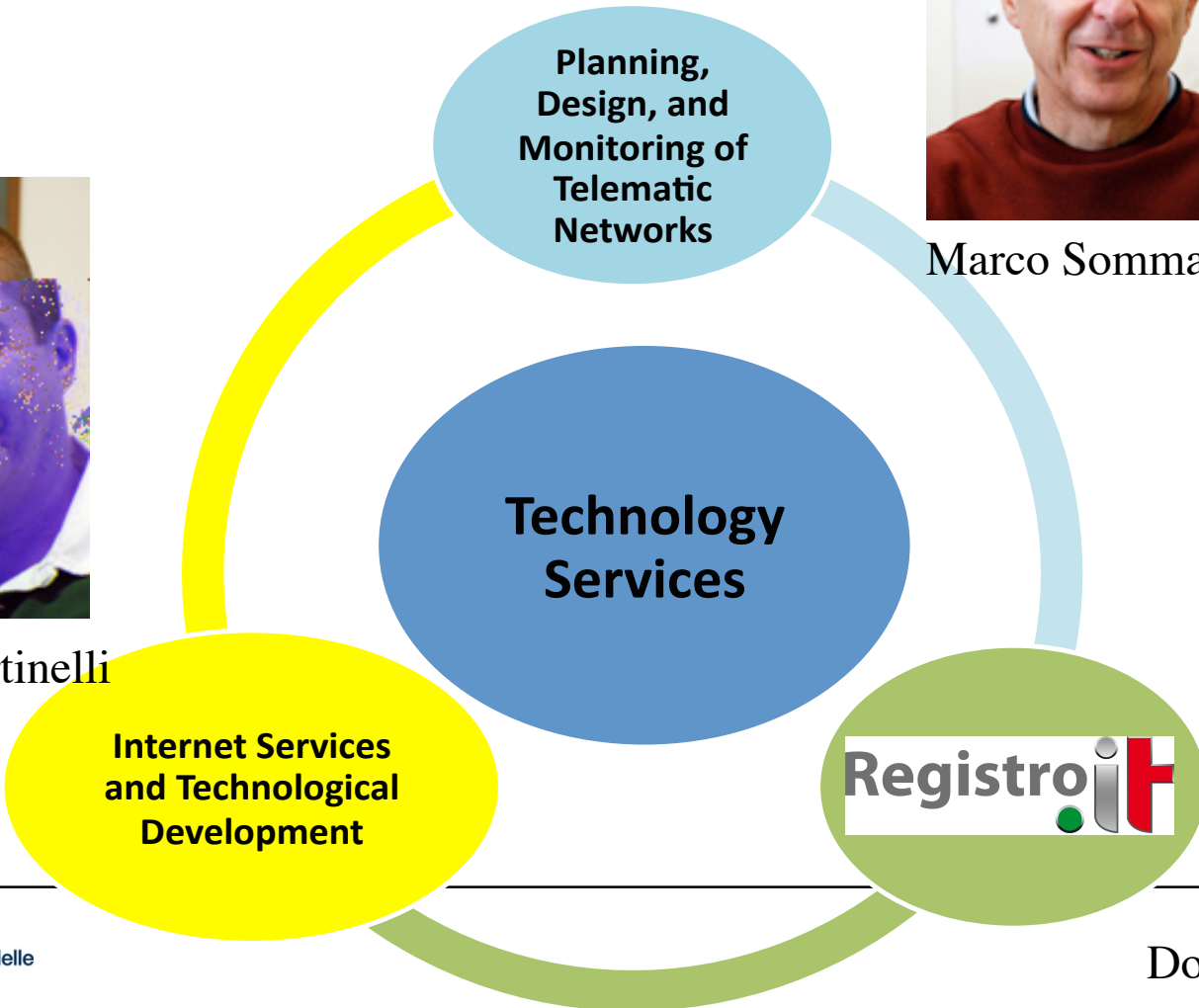


Marco Conti





Maurizio Martinelli

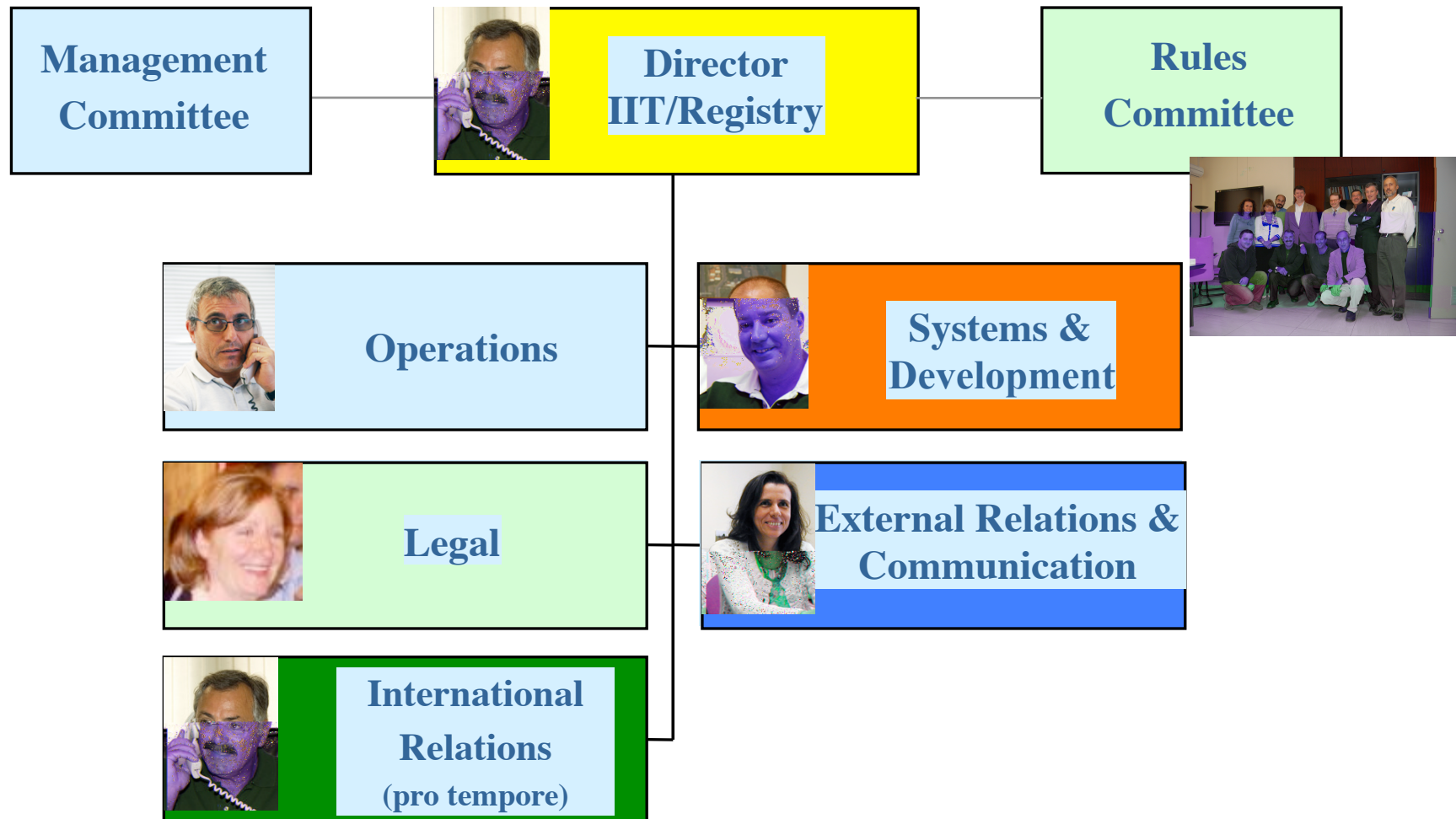


Marco Sommani



Domenico Laforenza

- Delegated to CNR on December 23rd, 1987
- More than 2.140.000 domain names
- New “synchronous” (**E**xtensible **P**rovisioning **P**rotocol - **EPP**) registration system from September 28<sup>th</sup>, 2009
- Coexistence of the “asynchronous” and “synchronous” systems until August 2011
- About 1,200 Registrars
- Open to EU juridical and natural persons



*About 70 people including staff-operators, administrative persons and technicians*

## 2010 Internet World Statistics Internet Users by World Regions - Preliminary

Region	Population	Internet Users	P. R.	% Users
Asia	3,834,792,852	872,526,978	22.8 %	43.0 %
Europe	813,319,511	475,123,735	58.4 %	23.4 %
North America	344,124,450	271,330,900	78.8 %	13.4 %
Latam / Carib.	592,556,972	209,874,973	35.4 %	10.3 %
Africa	1,013,779,050	115,631,340	11.4 %	5.7 %
Middle East	212,336,924	63,708,386	30.0 %	3.1 %
Oceania, Aust.	34,700,201	21,272,470	61.3 %	1.1 %
Total World	6,845,609,960	2,029,468,782	29.6 %	100.0 %

Source: Internet World Stats, preliminary estimates for December 31, 2010. Notes: a) Population data comes from the US Census Bureau; b) Internet user estimates come from ITU, Nielsen Online, GfK, official regulating agencies and other trustworthy surveys; c) Mexico is included in Central America and Turkey is included in Europe, according to the United Nations country classification; d) P.R. means the Penetration Rate (Internet users times 100 divided by population). Copyright © 2011, Miniwatts Marketing Group.

Internet Users in the European Union					
EUROPEAN UNION	Population ( 2010 Est. )	Internet Users, Latest Data	Penetration (% Population)	User Growth (2000-2010)	Users % Table
<a href="#">Austria</a>	8,214,160	6,143,600	74.8 %	192.6 %	1.3 %
<a href="#">Belgium</a>	10,423,493	8,113,200	77.8 %	305.7 %	1.7 %
<a href="#">Bulgaria</a>	7,148,785	3,395,000	47.5 %	689.5 %	0.7 %
<a href="#">Cyprus</a>	1,102,677	433,800	39.3 %	261.5 %	0.1 %
<a href="#">Czech Republic</a>	10,201,707	6,680,800	65.5 %	568.1 %	1.4 %
<a href="#">Denmark</a>	5,515,575	4,750,500	86.1 %	143.6 %	1.0 %
<a href="#">Estonia</a>	1,291,170	969,700	75.1 %	164.5 %	0.2 %
<a href="#">Finland</a>	5,255,695	4,480,900	85.3 %	132.5 %	0.9 %
<a href="#">France</a>	64,768,389	44,625,300	68.9 %	425.0 %	9.4 %
<a href="#">Germany</a>	82,282,988	65,123,800	79.1 %	171.3 %	13.7 %
<a href="#">Greece</a>	10,749,943	4,970,700	46.2 %	397.1 %	1.0 %
<a href="#">Hungary</a>	9,992,339	6,176,400	61.8 %	763.8 %	1.3 %
<a href="#">Ireland</a>	4,622,917	3,042,600	65.8 %	288.1 %	0.6 %
<a href="#">Italy</a>	58,090,681	30,026,400	51.7 %	127.5 %	6.3 %
<a href="#">Latvia</a>	2,217,969	1,503,400	67.8 %	902.3 %	0.3 %
<a href="#">Lithuania</a>	3,545,319	2,103,471	59.3 %	834.9 %	0.4 %
<a href="#">Luxembourg</a>	497,538	424,500	85.3 %	324.5 %	0.1 %
<a href="#">Malta</a>	406,771	240,600	59.1 %	501.5 %	0.1 %
<a href="#">Netherlands</a>	16,783,092	14,872,200	88.6 %	281.3 %	3.1 %
<a href="#">Poland</a>	38,463,689	22,450,600	58.4 %	701.8 %	4.7 %
<a href="#">Portugal</a>	10,735,765	5,168,800	48.1 %	106.8 %	1.1 %
<a href="#">Romania</a>	21,959,278	7,786,700	35.5 %	873.3 %	1.6 %
<a href="#">Slovakia</a>	5,470,306	4,063,600	74.3 %	525.2 %	0.9 %
<a href="#">Slovenia</a>	2,005,692	2,003,136	64.8 %	332.8 %	0.3 %
<a href="#">Spain</a>	46,505,963	29,093,984	62.6 %	440.0 %	6.1 %
<a href="#">Sweden</a>	9,074,055	8,397,900	92.5 %	107.5 %	1.8 %
<a href="#">United Kingdom</a>	62,348,447	51,442,100	82.5 %	234.0 %	10.8 %
<b>European Union</b>	<b>499,671,847</b>	<b>337,779,055</b>	<b>67.6 %</b>	<b>257.8 %</b>	<b>100.0 %</b>

NOTES: (1) The European Union Internet Statistics were updated for June 30, 2010. (2) Population is based on data from the [Census Bureau](#). (3) The usage numbers come from various sources, mainly from data published by [Nielsen Online](#) , [ITU](#) , [GfK](#), local NICs and private sources. (4) Data may be cited, giving due credit and establishing an active link to [Internet World Stats](#) . Copyright © 2010, Miniwatts Marketing Group. All rights reserved.



Today's Internet user is increasingly likely to be a non-English speaker who does not use Latin script characters in every day life.

- ❖ 60% of Internet users are non-English speakers, while the dominant language used on the Internet is English

Therefore, it is important to depict domain names in **non-Latin scripts**, called Internationalized Domain Names (IDNs), to enhance the ability of all users to use the Internet in their own languages.

<http://www.icann.org/en/topics/idn/>



## Scripts

European Scripts	African Scripts	South Asian Scripts	East Asian Scripts
<b>Armenian</b>	<b>Bamum</b>	<b>Bengali</b>	<b>Bopomofo</b>
Armenian Ligatures	Bamum Supplement	Brahmi	Bopomofo Extended
<b>Coptic</b>	<b>Egyptian Hieroglyphs (1MB)</b>	<b>Devanagari</b>	<b>CJK Unified Ideographs (Han) (28MB)</b>
Coptic in Greek block	<b>Ethiopic</b>	Devanagari Extended	CJK Extension-A (6.3MB)
<b>Cypriot Syllabary</b>	Ethiopic Supplement	<b>Gujarati</b>	CJK Extension B (30MB)
<b>Cyrillic</b>	Ethiopic Extended	<b>Gurmukhi</b>	CJK Extension C (2.8MB)
Cyrillic Supplement	Ethiopic Extended-A	<b>Kaithi</b>	CJK Extension D
Cyrillic Extended-A	<b>N'Ko</b>	<b>Kannada</b>	(see also Unihan Database)
Cyrillic Extended-B	<b>Osmanya</b>	<b>Kharoshthi</b>	<b>CJK Compatibility Ideographs (.5MB)</b>
<b>Georgian</b>	<b>Tifinagh</b>	<b>Lepcha</b>	CJK Compatibility Ideographs Supplement
Georgian Supplement	<b>Vai</b>	<b>Limbu</b>	<b>CJK Radicals / KangXi Radicals</b>
<b>Glagolitic</b>	<b>Middle Eastern Scripts</b>	<b>Malayalam</b>	CJK Radicals Supplement
<b>Gothic</b>	<b>Arabic</b>	<b>Meetei Mayek</b>	CJK Strokes
<b>Greek</b>	Arabic Supplement	<b>Oj Chiki</b>	Ideographic Description Characters
Greek Extended	Arabic Presentation Forms-A	<b>Oriya</b>	<b>Hangul Jamo</b>
<b>Latin</b>	Arabic Presentation Forms-B	<b>Saurashtra</b>	Hangul Jamo Extended-A
Latin-1 Supplement	<b>Aramaic, Imperial</b>	<b>Sinhala</b>	Hangul Jamo Extended-B
Latin Extended-A	<b>Avestan</b>	<b>Siyoti Nagri</b>	Hangul Compatibility Jamo
Latin Extended-B	<b>Carian</b>	<b>Tamil</b>	<b>Halfwidth Jamo</b>
Latin Extended-C	<b>Cuneiform (1MB)</b>	<b>Telugu</b>	<b>Hangul Syllables (.7MB)</b>
Latin Extended-D	Cuneiform Numbers and Punctuation	<b>Thaana</b>	<b>Hiragana</b>
Latin Extended Additional	Old Persian	<b>Vedic Extensions</b>	<b>Katakana</b>
Latin Ligatures	Ugaritic	<b>Southeast Asian Scripts</b>	Katakana Phonetic Extensions
<b>Fullwidth Latin Letters</b>	<b>Hebrew</b>	<b>Batak</b>	Kana Supplement
<b>Linear B</b>	Hebrew Presentation Forms	<b>Balinese</b>	<b>Halfwidth Katakana</b>
Linear B Syllabary	<b>Lycian</b>	<b>Buginese</b>	<b>Kanbun</b>
Linear B Ideograms	<b>Lydian</b>	<b>Cham</b>	<b>Lisu</b>
<b>Ogham</b>	<b>Mandaic</b>	<b>Javanese</b>	<b>Yi</b>
<b>Old Italic</b>	<b>Old South Arabian</b>	<b>Kayah Li</b>	Yi Syllables (.5MB)
<b>Phaistos Disc</b>	<b>Pahlavi, Inscriptional</b>	<b>Khmer</b>	Yi Radicals
<b>Runic</b>	<b>Parthian, Inscriptional</b>	Khmer Symbols	<b>American Scripts</b>
<b>Shavian</b>	<b>Phoenician</b>	<b>Lao</b>	<b>Cherokee</b>
<b>Phonetic Symbols</b>	<b>Samaritan</b>	<b>Myanmar</b>	<b>Deseret</b>
<b>IPA Extensions</b>	<b>Syriac</b>	Myanmar Extended-A	<b>Unified Canadian Aboriginal Syllabics</b>
<b>Phonetic Extensions</b>	<b>Central Asian Scripts</b>	<b>New Tai Lue</b>	UCAS Extended
Phonetic Extensions Supplement	<b>Mongolian</b>	<b>Rejang</b>	<b>Other</b>
<b>Modifier Tone Letters</b>	<b>Old Turkic</b>	<b>Sundanese</b>	Alphabetic Presentation Forms
<b>Spacing Modifier Letters</b>	<b>Phags-Pa</b>	<b>Tai Le</b>	Halfwidth and Fullwidth Forms
<b>Superscripts and Subscripts</b>	<b>Tibetan</b>	<b>Tai Tham</b>	ASCII Characters
<b>Combining Diacritics</b>		<b>Tai Viet</b>	
<b>Combining Diacritical Marks</b>		<b>Thai</b>	
Combining Diacritical Marks Supplement		<b>Philippine Scripts</b>	
<b>Combining Half Marks</b>		<b>Buhid</b>	
		<b>Hanunoo</b>	
		<b>Tagalog</b>	
		<b>Tanhanua</b>	



# Script for everyone.....

(<http://www.unicode.org/charts/#scripts>)

Arabic Supplement

	075	076	077
0	ب	پ	تش
1	ث	پ	ژ
2	ب	ک	ط
3	ت	ک	ا
4	ن	ک	آ
5	ب	م	ی
6	ب	م	ی
7	خ	ی	ی
8	ح	ن	ف
9	ط	ن	ف
A	د	ل	ط
B	ر	ز	ط
C	ش	ز	ح
D	ع	س	س
E	ع	ح	س
F	غ	ح	ک

AC00

Hangul Syllables

ACFF

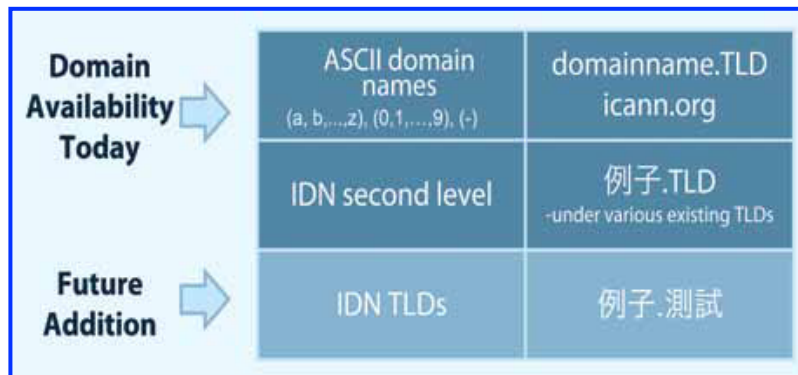
	AC0	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	ACA	ACB	ACC	ACD	ACE	ACF
0	가	감	갸	갹	갈	각	갚	거	검	겐	갹	결	격	겉	고	곰
1	각	갑	갸	갹	갹	갹	갹	거	검	겐	갹	결	격	겉	곡	곰
2	갸	갹	갹	갹	갹	갹	갹	거	검	겐	갹	결	격	겉	구	굶
3	갹	갹	갹	갹	갹	갹	갹	거	검	겐	갹	결	격	겉	국	굶
4	간	갸	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
5	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
6	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
7	간	갸	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
8	갈	각	갸	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
9	갸	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
A	갸	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
B	갸	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
C	갸	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
D	갸	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
E	갸	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹
F	갸	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹	갹

C1 Controls and Latin-1 Supplement

	008	009	00A	00B	00C	00D	00E	00F
0	XXX	DCS	NB SP	o	À	Đ	à	đ
1	XXX	PU1	i	±	Á	Ñ	á	ñ
2	BPH	PU2	¢	²	Â	Ò	â	ò
3	NBH	STS	£	³	Ã	Ó	ã	ó
4	IND	CCH	¤	´	Ä	Ô	ä	ô
5	NEL	MW	¥	µ	Å	Õ	å	õ
6	SSA	SPA		¶	Æ	Ö	æ	ö
7	ESA	EPA	§	·	Ç	×	ç	÷
8	HTS	SOS	¨	¸	È	Ø	è	ø
9	HTJ	XXX	©	¹	É	Ù	é	ù
A	VTS	SCI	ª	º	Ê	Ú	ê	ú
B	PLD	CSI	«	»	Ë	Û	ë	û
C	PLU	ST	¼	½	Ì	Ü	ì	ü
D	RI	OSC	½	¾	Í	Ý	í	ý
E	SS2	PM	®	¾	Î	Þ	î	þ
F	SS3	APC	—	¿	Ï	ß	ï	ÿ

# Internationalized Domain Names (IDNs)

- IDNs are domain names represented by local language characters (non-ASCII characters). Such domain names could contain characters with diacritical marks (ñ, é) as required by many European languages, or characters from non-Latin scripts (for example, Arabic or Chinese)
  - are domain names with characters other than a, b..., z; 0, 1, ..., 9; and “-“
- IDNs on the second and third levels exist in some gTLDs and in some ccTLDs



Script	Language	SLD, TLD U-labels	SLD A-label	TLD A-label
Arabic	Arabic	مثال اختبار	xn--mgbh0fb	xn--kgbechtv
Arabic	Persian	مثال آزمائشی	xn--mgbh0fb	xn--hgbk6aj7f53bba
Chinese, simplified	Chinese	例子.测试	xn--fsqu00a	xn--0zwm56d
Chinese, traditional	Chinese	例子.測試	xn--fsqu00a	xn--g6w251d
Cyrillic	Russian	пример.испытание	xn--e1afmkfd	xn--80akhbyknj4f
Devanagari	Hindi	उदाहरण.परीक्षा	xn--p1b6ci4b4b3a	xn--11b5bs3u9aj6g
Greek	Greek	παράδειγμα.δοκιμή	xn--hxajbhq2az3al	xn--jxalpdlp
Hangul	Korean	실례. 테스트	xn--9n2bp8q	xn--9t4b11yi5a
Hebrew	Yiddish	דאָס.לעבן	xn--fdbk5d8ap9b8a8d	xn--deba0ad
Kanji Hirigana, and Katakana	Japanese	例え.テスト	xn--r8jz45g	xn--zekzah
Tamil	Tamil	உதாரணம்.பரிட்சை	xn--zke6cc5bi7f6e	xn--hlcj6aya9esc7a

- TLD Registries determine the choice of characters available under these TLDs (see the Italian approach later on)

- ❑ IDN TLDs will be made available through two separate processes, initially through [the New gTLD Program](#) and the [IDN ccTLD Fast Track Process](#)
- ❑ **IDN SLDs** Usually a reference for domain names with local characters at the **S**econd **L**evel, while the top level remains in ASCII-only characters.
  - ❖ For example: [*παράδειγμα.test*] (“*example.test*” in Greek)

Second Level



- ❑ **IDN TLDs** Usually the short reference for internationalized **T**op **L**evel labels referring to the entire domain name being represented by local characters.

- ❖ For example: [*실례.테스트*] (“*example.test*” in Hanguk (Korean script))

Second Level



First Level



## The First Four


- Special recognition of the first four IDN ccTLD:
  - United Arab Emirates, Saudi-Arabia, Russian Federation, Egypt
  - simultaneous: delegation: الإمارات السعودية and مصر.
  - .pφ followed shortly after
- Milestones include:
  - Jun 2003: IDNA protocol, IDN Guidelines, initiation browser implemented
  - Feb 2008: First report from IDNC leading to IDN ccTLD Fast Track Process
  - Oct 2008: First draft implementation plan for IDN ccTLD Process
  - Oct 2009: ICANN Board approves IDN ccTLD Fast Track Process
  - Nov 2009: Launch of IDN ccTLD Fast Track Process
  - Jan 2010: First four IDN ccTLD labels approved
  - Apr 2010: First four IDN ccTLD approved to be delegated
  - May 2010: Introduction of the first four

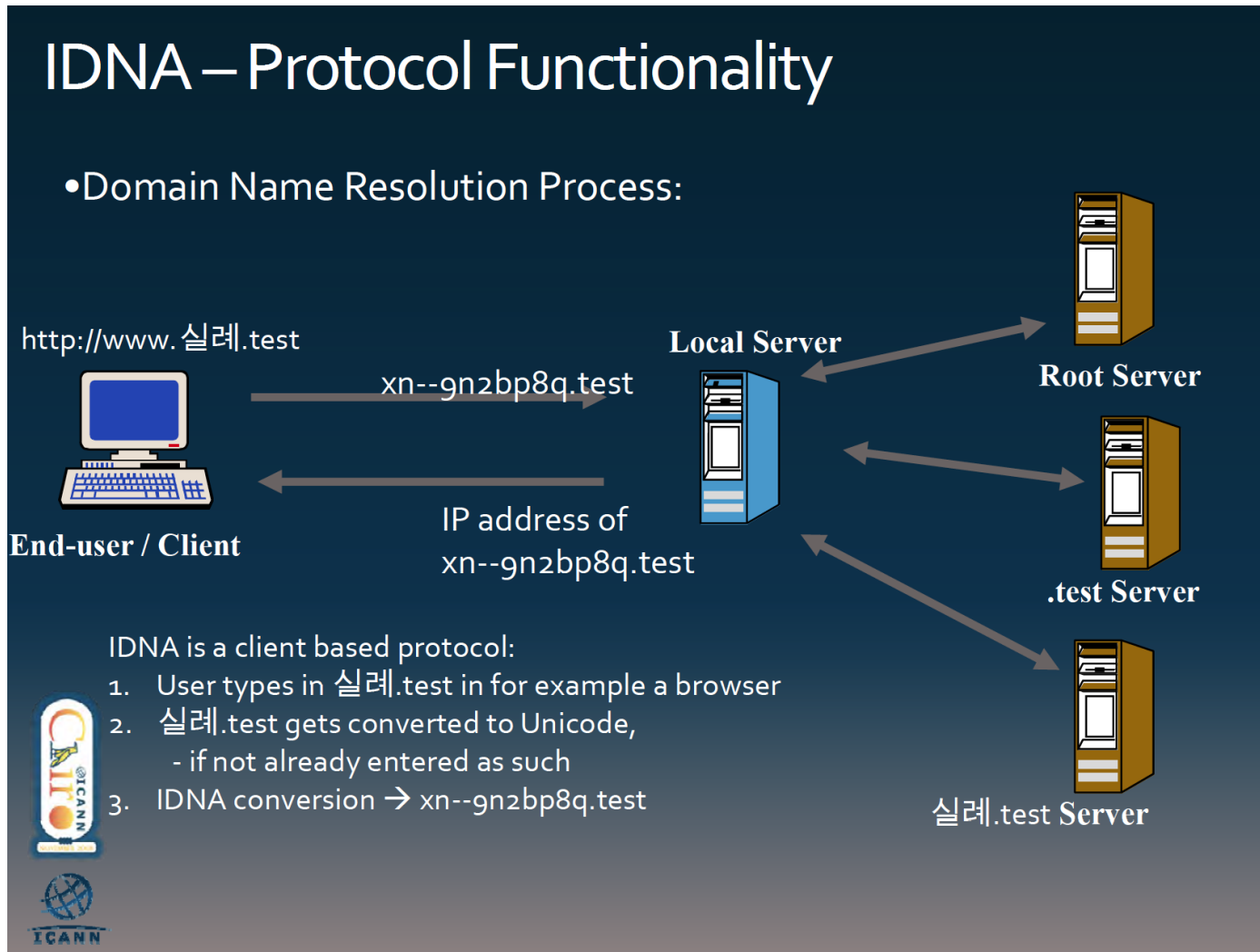
- IDNs will impose minimal changes to current DNS system
- IDNs must not break existing structure and hierarchy
- What is the solution?

## ✓ Punycode (RFC3492)

- IDNs are stored in the Domain Name System as ASCII strings using Punycode transcription



- ❑ Sequence of ASCII characters all IDNs will be encoded into in order for the Domain Name System (DNS) to understand and manage the names.
  - ❑ The intention is that domain name registrants and users will never see this decoded form of a domain name.
- ❑ The sole purpose is for the DNS to be able to resolve for example a web address containing local characters.
- ❑ The DNS is only capable of handling ASCII characters.
- ❑ For example, the punycode version of  [उदाहरण.परीक्षा] this is the Hindi, in Devanagari script, version of “example.test”) is:
  - ❖ xn--p1b-6ci4b4b3a.xn--11b5bs3a9aj6g
- ❑ The prefix for the Punycode version of the domain names is always “xn--”. Hence this prefix is often reserved at the registry level to avoid confusion in registration of IDNs.



- A mechanism defined in 2003 for handling IDNs containing non-ASCII characters.
  - An IDNA-enabled application is able to convert between the internationalized and ASCII representations of a domain name.
  - It uses the ASCII form for DNS lookups but can present the internationalized form to users who presumably prefer to read and write domain names in non-ASCII scripts such as Arabic or Hiragana.
  - Applications that do not support IDNA will not be able to handle domain names with non-ASCII characters, but will still be able to access such domains if given the (usually rather cryptic) ASCII equivalent.
  - Although the Domain Name System supports non-ASCII characters, applications such as email and web browsers restrict the characters which can be used as domain names for purposes such as a hostname.

## Why are IDNs important?

- The most important is the growing number of Internet users around the world for whom it is difficult to use ASCII characters.
  - The fact is, the Internet is accessed by more people who do not use Latin languages and scripts than those who do.
    - This means that it is difficult for them to recognize ASCII characters and reproduce them on keyboards or use software to enter website addresses in browsers.

The screenshot shows the Chinese homepage of the Chinese Academy of Sciences. At the top, there is a navigation bar with links for English, Simplified Chinese, RSS, Site Map, Favorites, Email, and Contact Us. The main header features the CAS logo and the slogan "唯实 求真 协力 创新". Below this, there are several news articles and sections. A prominent article is titled "周永康在天津考察时视察中科院天津工业生物技术研究所". Other sections include "走进中国科学院", "科技动态", "院领导", "院机关", "院属机构", "院长信箱", "科研创新单元", "资源服务", "学术活动", "会议信息", "招生招聘", and "视频直播". The layout is organized into multiple columns with various icons and text links.

The screenshot shows the English homepage of the Chinese Academy of Sciences. The top navigation bar includes links for Home, About CAS, CAS Institutes, Newsroom, Administration, and Join Us. The main content area is divided into several sections: "Events" (CAS President Bai Chunli Answers Science's Questions on Innovation), "Quick Links to" (listing various departments like Science & Technology, Scientists, etc.), "Survey & Feedback" (inviting users to improve the site), "Photo & Video" (featuring a duck), "Research Progress" (listing scientific achievements), "International Cooperation News" (reporting on various international events), "Upcoming Conferences" (listing academic meetings), and "China News" (reporting on archaeological findings). The right sidebar contains a search bar, "Annual Reports", "Highlights" (featuring the Chinese Science Bulletin), and "Journals" (listing various scientific journals).





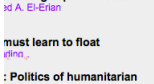
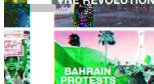
**Invest in one of the world's most violent cities!**

Al Jazeera's Chris Arsenault reports from Juarez, where low-wage manufacturing is increasing



[News](#) | [In Depth](#) | [Programmes](#) | [Video](#) | [Blogs](#) | [Business](#) | [Weather](#) | [Sport](#) | [Watch Live](#)

[Africa](#) | [Americas](#) | [Asia-Pacific](#) | [Central & South Asia](#) | [Europe](#) | [Middle East](#)



### Libyan rebels lay down terms for ceasefire

Opposition offers ceasefire if Gaddafi halts attacks against rebel-held cities but battles rage for Brega and Misrata.

Last Modified: 01 Apr 2011 14:43 GMT

[Read More](#)

### Libyan rebels lay down terms for ceasefire

More 'defections' from Gaddafi inner circle

Libya Live Blog - April 1

Deaths in Syrian protest march for freedom

Fighting intensifies around Gbagbo home

### UN staff killed in Afghanistan Quran protests

UN staff killed in Afghanistan Quran protests

Last Modified: 01 Apr 2011 12:53 GMT

### Fears of clashes as Yemenis demonstrate

Thousands of protesters for and against the president rally in Sanaa, the capital, raising fears of bloodshed.

Last Modified: 01 Apr 2011 12:29 GMT

### Japan nuclear evacuation will be 'long term'

Officials under pressure to extend evacuation zone around Fukushima atomic plant as radiation leaks continue.

Last Modified: 01 Apr 2011 12:00 GMT

### UN court rejects Georgia case against Russia

International Court of Justice says it cannot rule on dispute over alleged atrocities in Abkhazia and South Ossetia.

Last Modified: 01 Apr 2011 15:55 GMT

### In Video

- Women vie for votes in Nigeria
- Question marks over reported Libya defections
- Mexican drug crimes go global
- South Sudan contemplates new capital

### Libyan rebels lay down terms for ceasefire

Opposition offers ceasefire if Gaddafi halts attacks against rebel-held cities but battles rage for Brega and Misrata.

Last Modified: 01 Apr 2011 14:43 GMT

[Read More](#)

### Libyan rebels lay down terms for ceasefire

More 'defections' from Gaddafi inner circle

Libya Live Blog - April 1

Deaths in Syrian protest march for freedom

Fighting intensifies around Gbagbo home

### UN staff killed in Afghanistan Quran protests

UN staff killed in Afghanistan Quran protests

Last Modified: 01 Apr 2011 12:53 GMT

### Fears of clashes as Yemenis demonstrate

Thousands of protesters for and against the president rally in Sanaa, the capital, raising fears of bloodshed.

Last Modified: 01 Apr 2011 12:29 GMT

### Japan nuclear evacuation will be 'long term'

Officials under pressure to extend evacuation zone around Fukushima atomic plant as radiation leaks continue.

Last Modified: 01 Apr 2011 12:00 GMT

### UN court rejects Georgia case against Russia

International Court of Justice says it cannot rule on dispute over alleged atrocities in Abkhazia and South Ossetia.

Last Modified: 01 Apr 2011 15:55 GMT

### In Video

- Women vie for votes in Nigeria
- Question marks over reported Libya defections
- Mexican drug crimes go global
- South Sudan contemplates new capital

**الأخبار** | كثير من التحليلات السياسية للقضايا في المنطقة

الجمعة 27/4/1432 هـ - الموافق 14/4/2011 م (آخر تحديث) الساعة 18:21 (مكة المكرمة)، 15:21 (غرينتش)

[الرئيسية](#) | [الأخبار](#) | [الفضائية](#) | [المعرفة](#) | [الاقتصاد والأعمال](#) | [مركز الدراسات](#) | [حريات وحقوق](#)

[الصفحة الرئيسية](#) | [الصفحة الأخيرة](#) | [الصفحة الأخيرة](#)

### التطورات على مدار الساعة:

- 18:07 إطلاق مدون واعتقالات بالبحرين
- 18:03 احتجاجات العراق هل تمدد لأميركيين؟
- 17:54 آمل بوعي الجيش في التغيير باليمن
- 17:54 انتقادات لإعلان الدستوري بمصر
- 17:54 قتل بتفريق مظاهرات قرب دمشق
- 17:42 الانتقالي الليبي يشترط للهدنة

[أحداث الأسبوع بالصور](#) | [عودة النوار إلى إجابيا الليبية](#) | [عملية "فجر الأوديسا" ضد القذافي](#)

[كارتون](#) | [تقرير: الدبلوماسيون](#)

### ربيع الثورات العربية

[مظنة الإجراء الإجباري حول فوكوشيما النووية](#) | [أحداث الأسبوع](#) | [كل العناوين](#)

### الأخبار الفضائية والاقتصاد والمعرفة مركز الدراسات حريات وحقوق

- مؤيد للتغيير بعيداً عن العنف
- ملوح الشعب اليمني للتغيير
- ثورة شبابية ضد نظام صالح
- تحديث من ليبيا 3
- تحديث من ليبيا 2
- تحديث من ليبيا 1

[تحديثات صفحة الجزيرةت على](#) | [إطلاق مدون واعتقالات بالبحرين: الراجح](#) | [السياسة الخارجية اليوم من المونديسمود](#) | [اليوسف الذي كانت الخارجية الأميركية-](#) | [http://bit.ly/gEOaPK](#) | [minutes ago · reply · retweet · favorite 41](#)

[احتجاجات العراق هل تمدد لأميركيين؟](#) | [بحر](#) | [خبراء قانونيون صعوبات كبيرة أمام الحكومة](#) | [http://bit.ly/gOaPK](#) | [minutes ago · reply · retweet · favorite 41](#)

[آمل بوعي الجيش في التغيير باليمن: قال](#) | [الجزيرة](#) | [شاركت](#)





Da [giusuecarducci.it](http://giusuecarducci.it) alla [festadelpapa.it](http://festadelpapa.it): i domini .it guadagnano l'accento e, dal gennaio prossimo, cominceranno ad apparire in rete per offrire nuove possibilità di personalizzare la propria presenza sul web. Il Registro .it si appresta a introdurre gli Idn (Internationalized Domain Names):

ta non propria. Per quanto riguarda l'Italia, l'impatto degli accenti sarà sicuramente più contenuto ma darà a tutti la possibilità di registrare domini esattamente uguali al nome della persona, della società o del marchio commerciale prescelto. Gli Idn rappresenteranno sicuramente un vantaggio per le regio-

il Registro di domani / via libera agli Idn

## e a gennaio arrivano i domini con l'accento

di Stefania Fabbri

espressione complessa - ma con un acronimo facile da ricordare - che indica il sistema tramite il quale si può "scrivere" su Internet in danese piuttosto che in cinese, utilizzando lettere accentate o caratteri non latini. Sul web, fino a poco tempo fa, la scelta dei nomi era limitata dai ventisei caratteri dell'alfabeto latino impiegati in lingua inglese (oltre ai dieci caratteri numerici e al trattino "-"). L'uso degli Idn, introdotta da Iann (Internet Corporation for Assigned Names and Numbers) nel 2009 durante il meeting di Seoul, rappresenta una svolta soprattutto per le popolazioni di lingua araba, cinese o cirillica: centinaia di milioni di utenti Internet del mondo che fino a ieri erano di fatto costretti a usare un alfabe-

ni italiane bilingue (Valle d'Aosta e Trentino Alto Adige), dove è più sentita la necessità di registrare domini in lingua francese o tedesca. I nomi con gli "accenti" potranno essere registrati solo in modalità sincrona. I Registrar potranno farlo semplicemente specificandolo nella richiesta di registrazione. Da un punto di vista tecnico, i nuovi domini Idn per essere utilizzabili in rete devono essere convertiti dal Registro in un formato apposito, chiamato Punycode, che corrisponde allo standard americano per lo scambio di informazioni. Per i Registrar è già disponibile, sul portale loro riservato, un tool per la conversione del nome in Idn: in pratica il nome a dominio "papà.it" viene tradotto in "xn-pap-ola.it".



18

[www.registro.it](http://www.registro.it)

À, Â, Ã, È, É, Ê, Ë, Ì, Î, Ï, Ò, Ô, Ö, Ù, Ú, Ü, Æ, Ç, ÿ

From [giusuecarducci.it](http://giusuecarducci.it) to [festadelpapa.it](http://festadelpapa.it): domains can now contain accented letters. From next January, users will have

a new way to personalize their presence on the web. The Italian Registry will be introducing IDNs (Internationalized Domain Names)

the Registry of tomorrow / green light to IDNs

## and coming in January domains with an accent

by Stefania Fabbri



- a complex concept but with an easy to remember acronym. Basically this is a system through which you can "write" on the Internet, for example in Danish or Chinese, using accented letters or non-Latin characters. Until recently, the choice of domain names was limited by the twenty-six Latin characters used in English (in addition to the ten digits and the hyphen "-"). The use of IDN was introduced by ICANN (Internet Corporation for Assigned Names and Numbers) in 2009 during a meeting in Seoul. It represents a breakthrough, especially for people of Arabic, Chinese or Cyrillic: hundreds of millions of Internet users in the world that until now were forced to use an alphabet that was not their own. With regard to Italy, the impact of accents will certainly be less marked, but it will give everyone the opportunity to register

domains which completely match the name of the person, company or brand name chosen. IDNs will be an advantage for bilingual Italian regions (Valle d'Aosta and Trentino Alto Adige), where they may need to register domains in French or German. Names with accents can only be registered in synchronous mode. Registrars can do so simply by specifying as such in their request for registration. From a technical point of view, the new IDN domains have to be converted by the Registry into a special format, called Punycode, which corresponds to the American standard for information exchange.

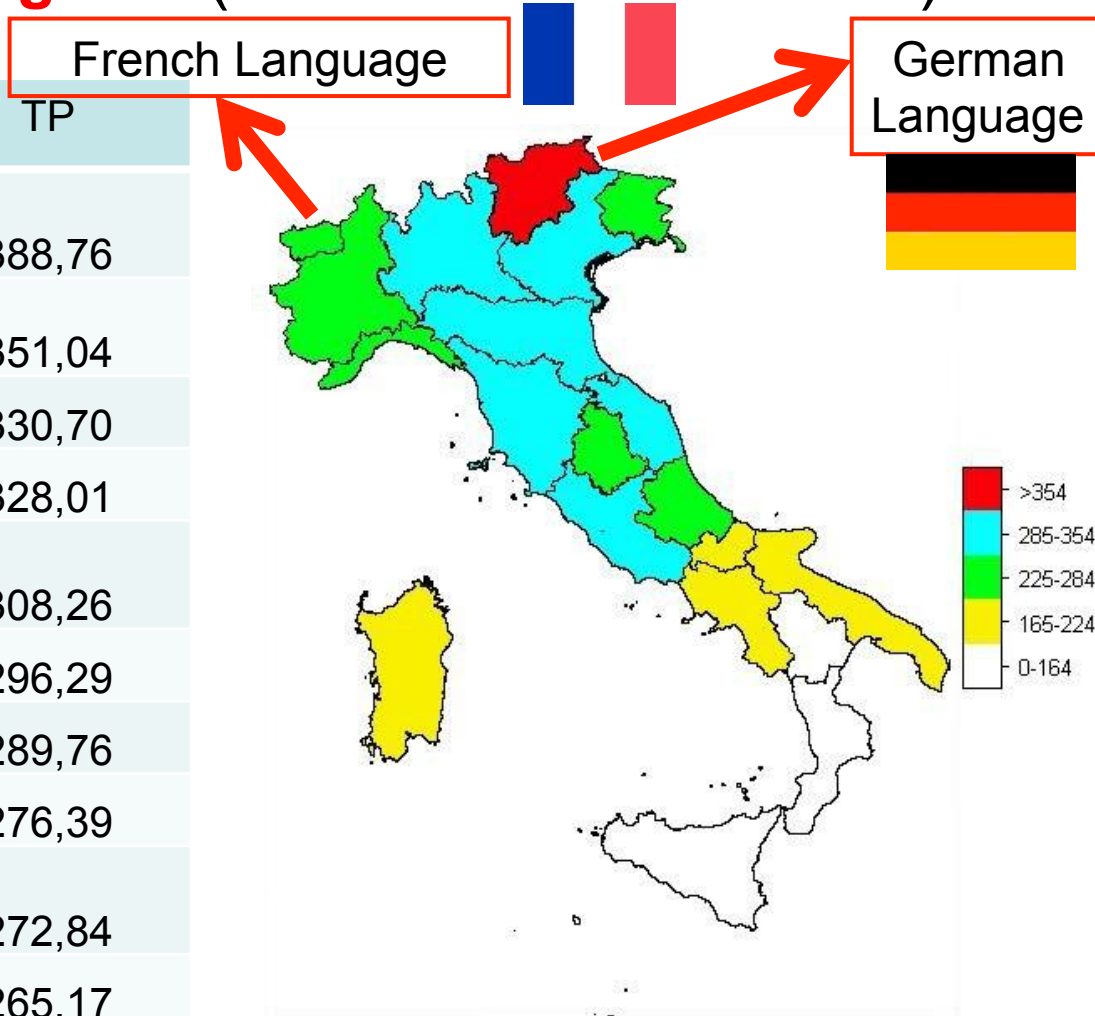
On their dedicated section on the Registry website, Registrars can use a tool for converting the name to IDN: for example, the domain name "papà.it" is translated into "xn-pap-ola.it".

[icann@registro.it](mailto:icann@registro.it)

19

## Internet Diffusion in Italy: The Top 10 Regions (TP = Penetration Rate)

Regione	% domini	TP
Trentino Alto Adige	2,54%	388,76
Lombardia	21,92%	351,04
Toscana	7,83%	330,70
Lazio	11,85%	328,01
Emilia Romagna	8,61%	308,26
Marche	2,93%	296,29
Veneto	9,03%	289,76
Umbria	1,58%	276,39
Valle d'Aosta	0,22%	272,84
Friuli V. G.	2,07%	265,17



**TP out of 10.000 inhabitants**

- Most of the words and names, in the Italian language are represented by ASCII characters (a-z A-Z 0-9) and, optionally, hyphen (-)
- IDNs will be an advantage for bilingual Italian regions (Valle d'Aosta and Trentino Alto Adige), where they may need to register domains in French or German.

**www.papa.it**

- papa (The Pope)
  - From Greek: πάππας (pappas)
  - **www.papa.it**
- papà (The Father)
  - **www.papà.it** (**www.xn--pap-cla.it**)



- ❑ Possibility of registering a domain name exactly equal to the real name of the person, company, brand, etc..
  - giosuècarducci.it (xn--giosucarducc-1db)
  - cittàicastello.it (xn--cittdicastell-sdb)
  - nestlè.it (xn--nestl-vqa)
- ❑ Allow the registration of domain names in languages other than English in order to cater to the needs of those Italian Regions where there is Italian bilingualism (Valle d'Aosta e Trentino Alto Adige)



- ❑ Reduction in the future, any disputes on domain names
  - [nestle.it](#) IS NOT EQUAL TO [nestlè.it](#)
- ❑ Increase in registrations of domain names



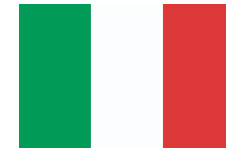
- The introduction of IDNs could generate, in some cases, confusion to the end user:
  - [www.agora.it](http://www.agora.it) e [www.agorà.it](http://www.agorà.it)
  - Web sites can target different
- Development and maintenance of many applications used in the Registry

- 1) Registration of the domain names in ACE (ASCII Compatible Encoding– es. xn--pap-cla.it) or IDN (es. papà.it) formats
  - ACE will be the format, however, necessary for the generation of the .it zone and of the of geographic domain names zones
- 2) Simultaneous starting of IDNs for the Italian language, for the German and French

### 3) Definition of the Unicode character set to support:

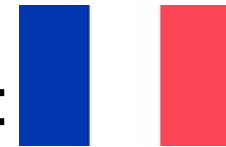
- ❖ The Latin-1 supplement and Latin Extended-A sets are more than enough to support accented characters of Italian, French and German

– characters to support in the Italian language:



- à, è, é, ì, ò, ù

– characters to support in the French language:



- à, â, æ, ç, è, é, ê, ë, î, ï, œ, ô, ù, û, ü, ÿ

– characters to support in the German language:



- ß, ä, ö, ü

- In essence, combining the non-ASCII characters of the three alphabets, we get the following set of 21 characters (and similar capital letters):
- à, â, ä, è, é, ê, ë, ì, î, ï, ò, ô, ö, ù, û, ü, æ, œ, ç, ß, ÿ
  - ß was introduced in the encoding Unicode (in July 2008) +1E9E LATIN CAPITAL LETTER SHARP S.

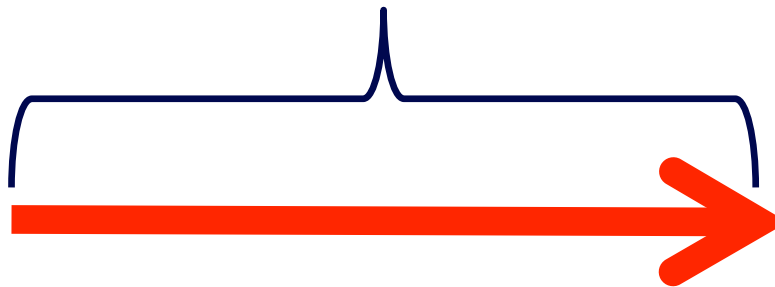
- Registration of the domain name in IDN format
- Whois query both in IDN and Punycode
- Output both in IDN and Punycode
- Authoritative nameservers for a domain name in the Punycode format
- Generation of DNS zones in Punycode format

- "Landrush" is a term used to define the stampede of registrations when a registry opens to registrations.
- We will adopt the Landrush procedure
- Landrush: Pros
  - First come first served
  - Total lack of organizational and management problems
- Landrush: Cons
  - Risk of cybersquatting



## Test period

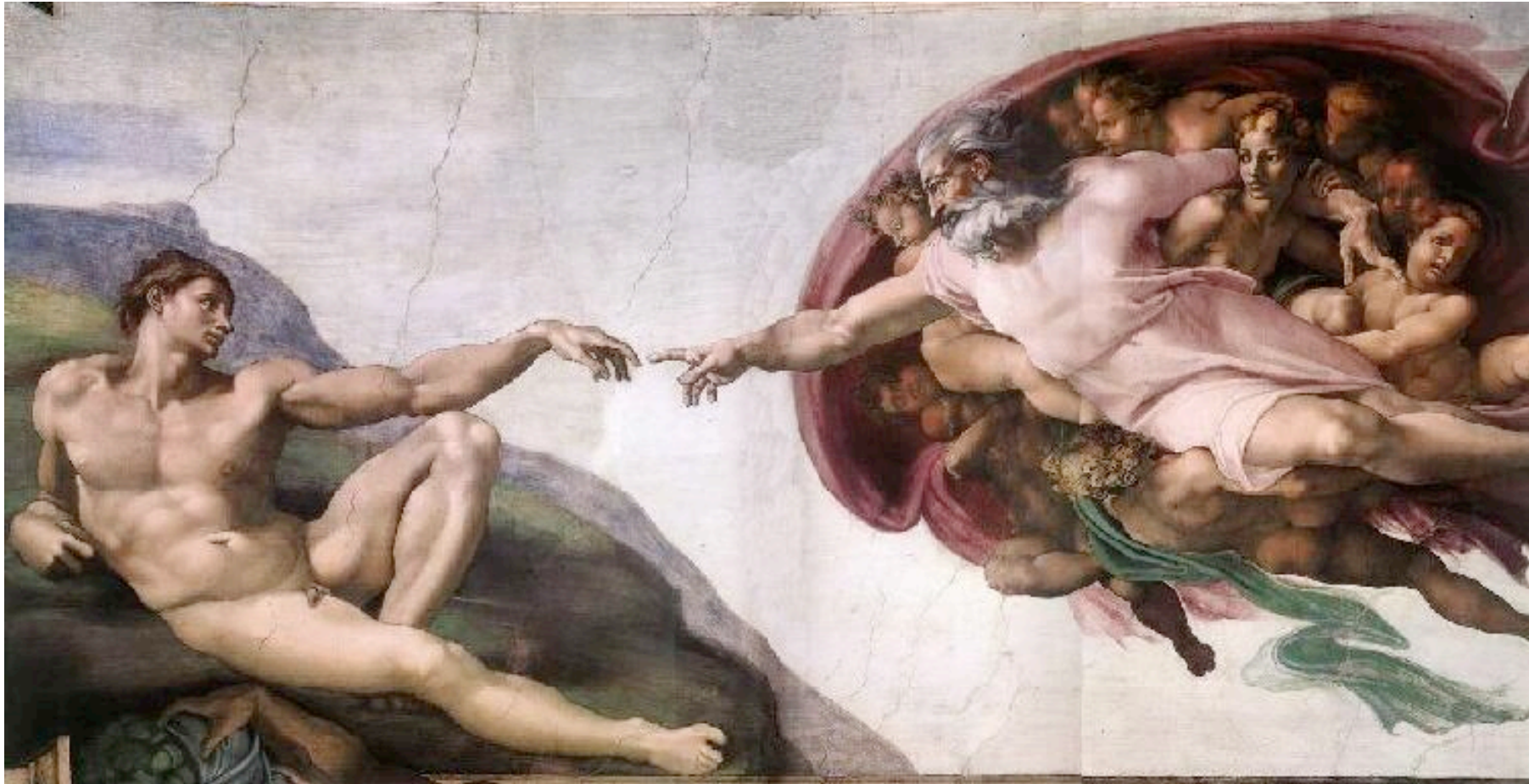
**TEST**



**IDNs: Available on  
the test platform  
January 2011**



**IDNs Official Landrush  
July 2011**



Creation of Adam (1511) by Michelangelo, Sistine Chapel, Vatican



- I would like to express my gratitude to Dr. Maurizio Martinelli, responsible for the Internet Services and Technological Development Department, IIT-CNR

Registro **it**

**GRAZIE**

**Danke**

谢谢

**Arigato** おかげで

спасибо

ευχαριστίες תודה

**Spasibo**

Obrigado

**Thank You**

شكرا

**Siyabonga**

**Achiu**

**gracias**

Dr. Domenico Laforenza  
domenico.laforenza@iit.cnr.it