

IDN Workshop

Marrakech Meeting

27 June 2006

IDN Issues



**IDN Issues: What IDN means to me
and to my fellow African brothers ?**

Mouhamet Diop
NEXT

Agenda

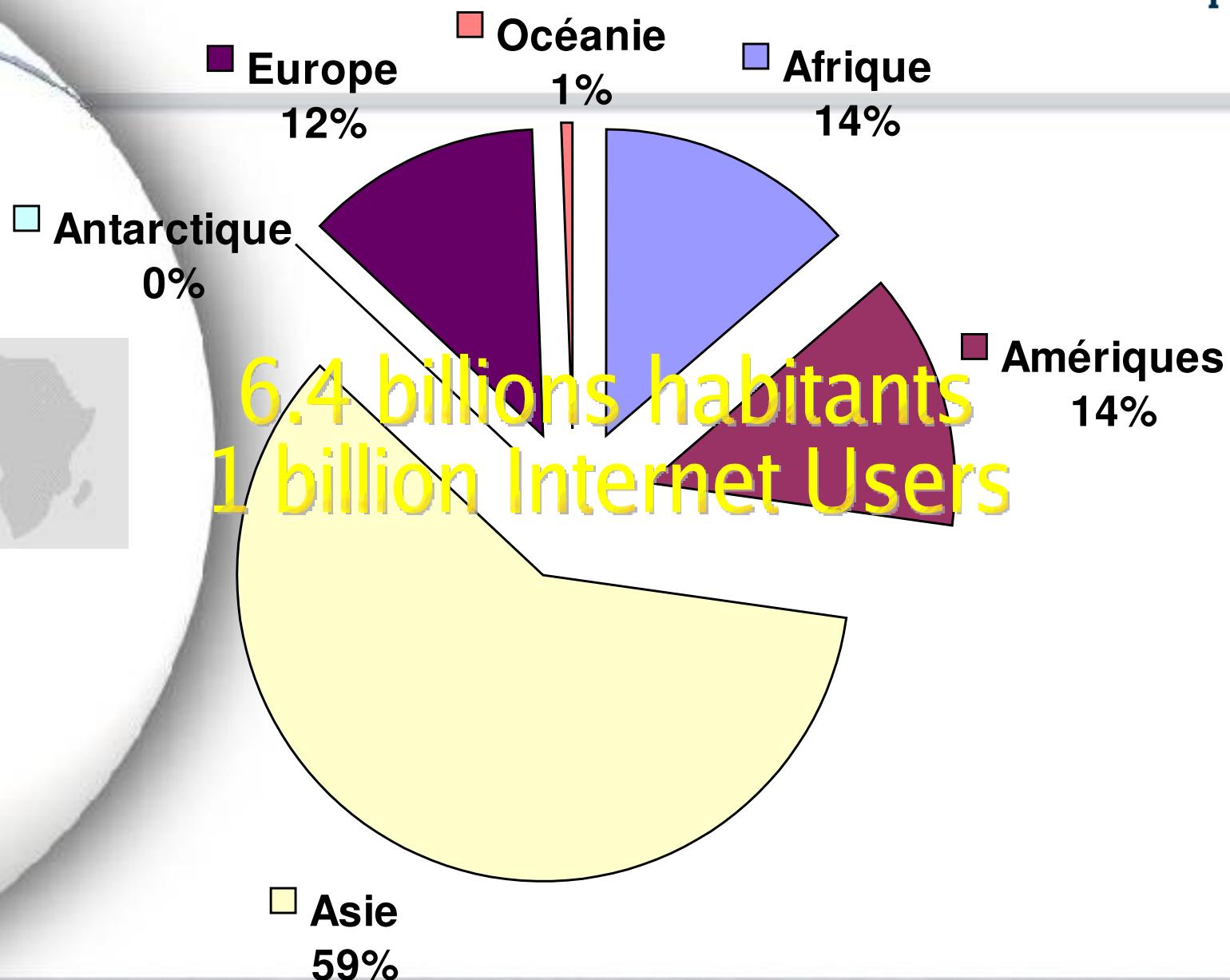
- I. ICANN Founding Core Principles
- II. World Diversity
- III. Languages in the World
- IV. DNS Layout Architecture
- V. IDN standards : IETF input
- VI. Experimentation and the way forward

Founding Principles for ICANN

- ◆ Stability
- ◆ Competition
- ◆ Choice ...
- ◆ Private, bottom-up coordination
- ◆ Representation (inclusive approach)
- ◆ Internationalization



POPULATION DU GLOBE



Languages in the world ...

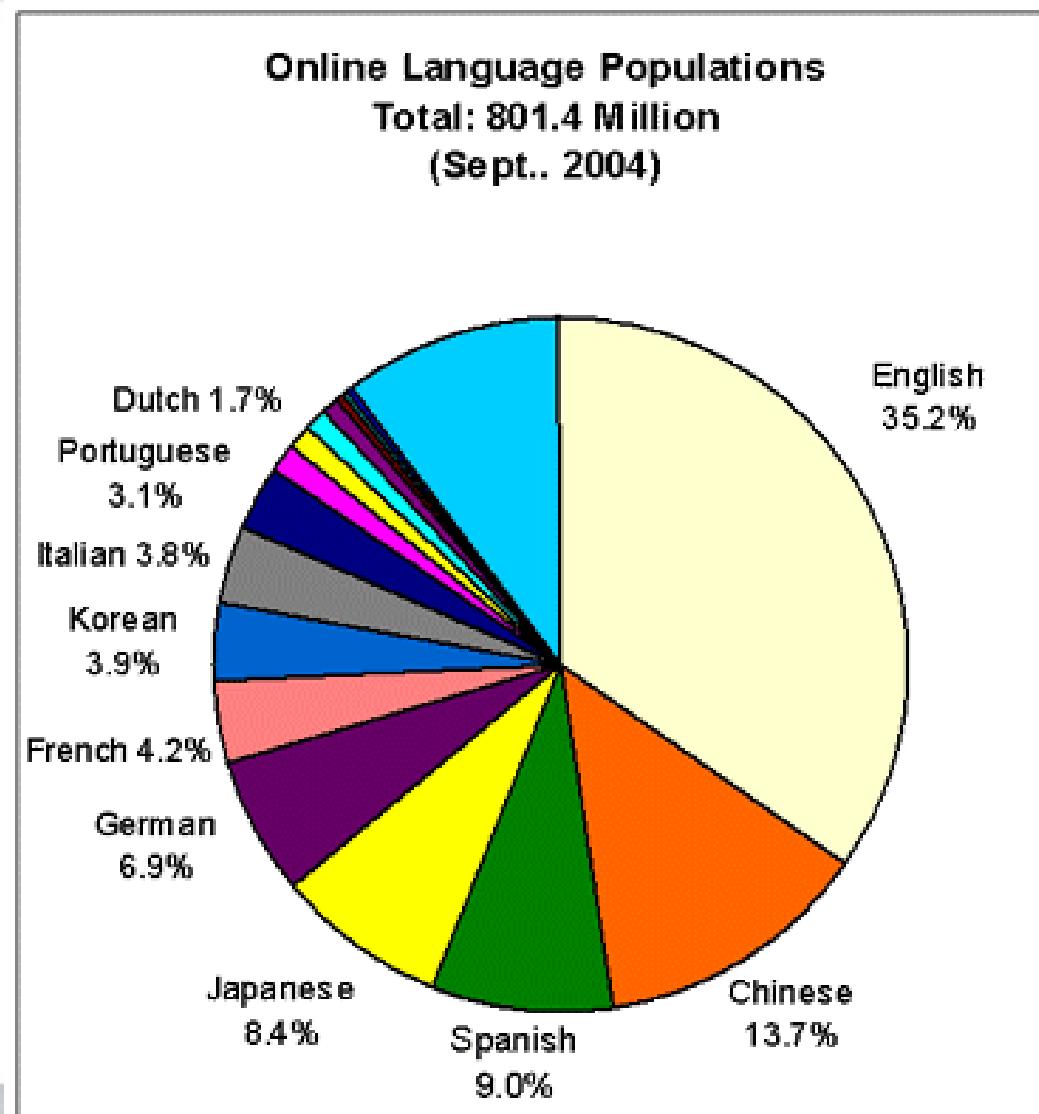
Languages

- + 6000 languages
- 97 % of the world population speak 4 % of the languages
- 96 % of the languages are spoken only by 3 % of the population

Global Internet users

- 64.8% are not English Speakers
- 70 % of web content are in English

<http://glreach.com/globstats>



Languages distribution – Top 10



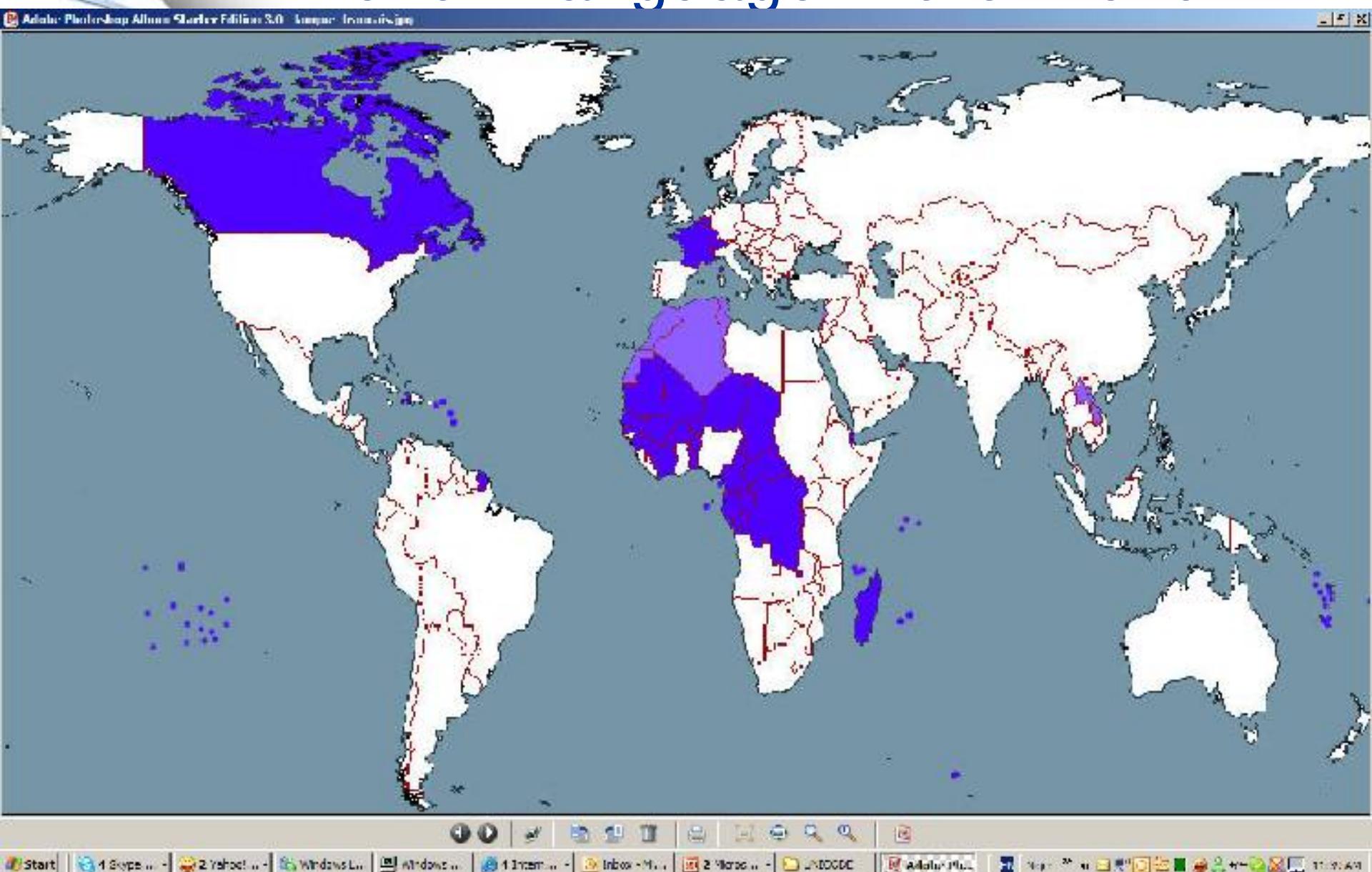
Language	Ranking	Localisations (principales)	Numbers millions (2000)
Chinois Mandarin	1	Chine	885
Anglais	2	Etats-Unis , Royaume-Uni , Canada , Inde , Australie ...	450
Français	3	France , Canada , Belgique , Suisse , Afrique , Asie du Sud-Est , Antilles , Océanie ...	285
Espagnol	4	Espagne , Amérique Latine ...	268
Arabe	5	Afrique du Nord , Moyen-Orient	200
Bengali	6	Bangladesh , Inde	189
Hindi	7	Inde , Pakistan	182
Portugais	8	Brésil , Portugal , Mozambique , Angola	175
Russe	9	Russie , CEI	170
Japonais	10	Japon	126

Source © PopulationData.net 2004

25 June 2006

Page 6

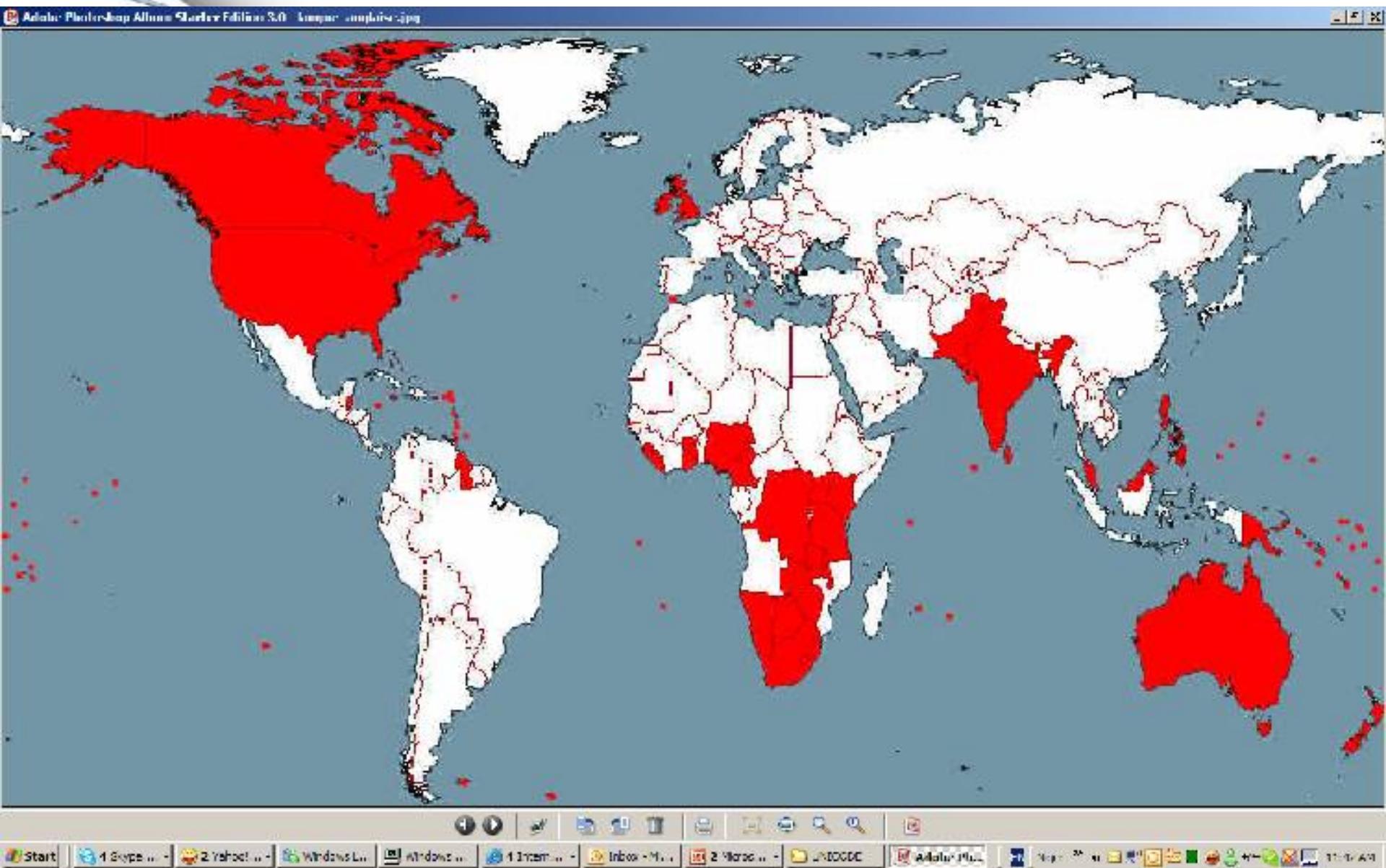
French Language in the World





English Language in the world

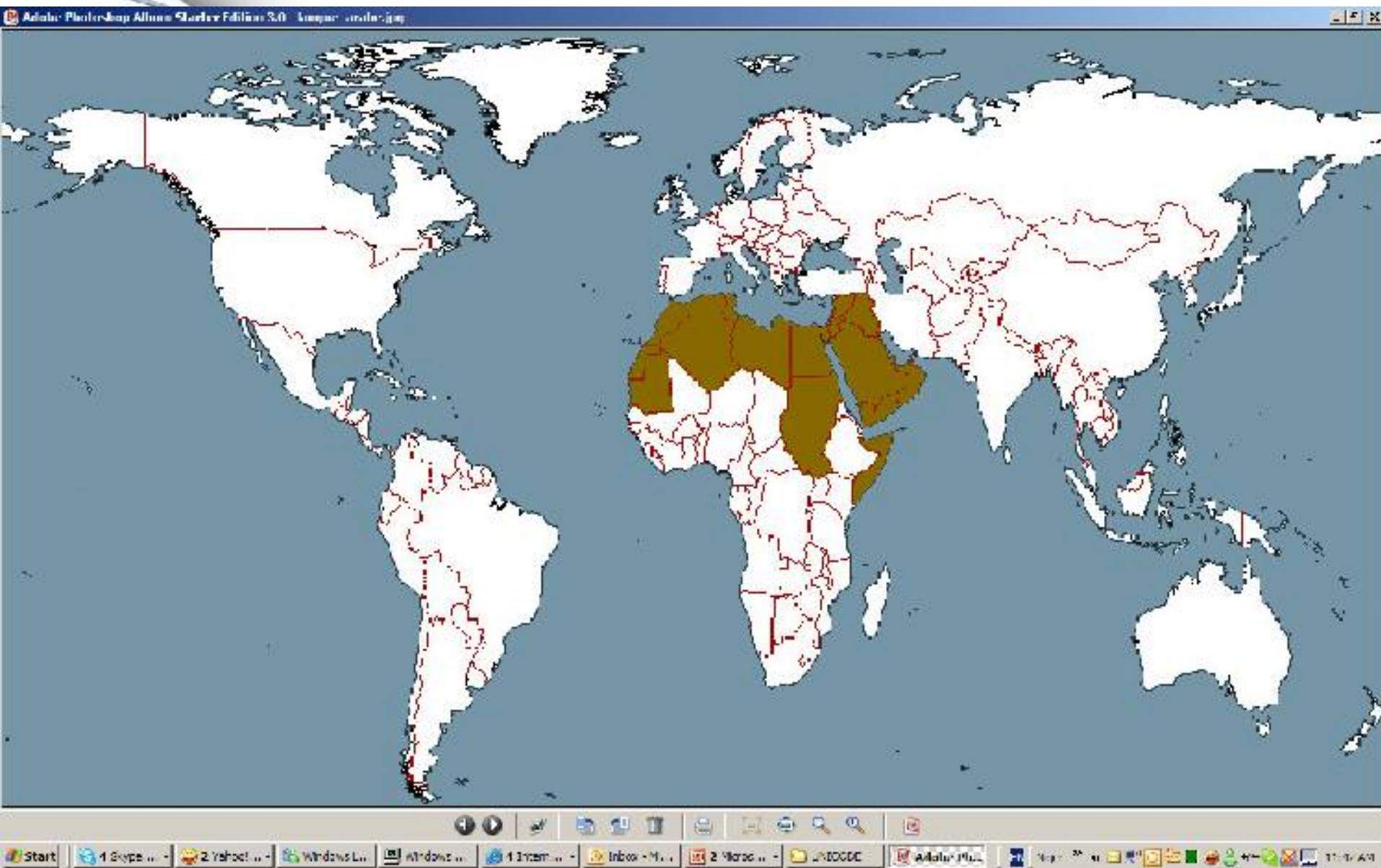
ICANN



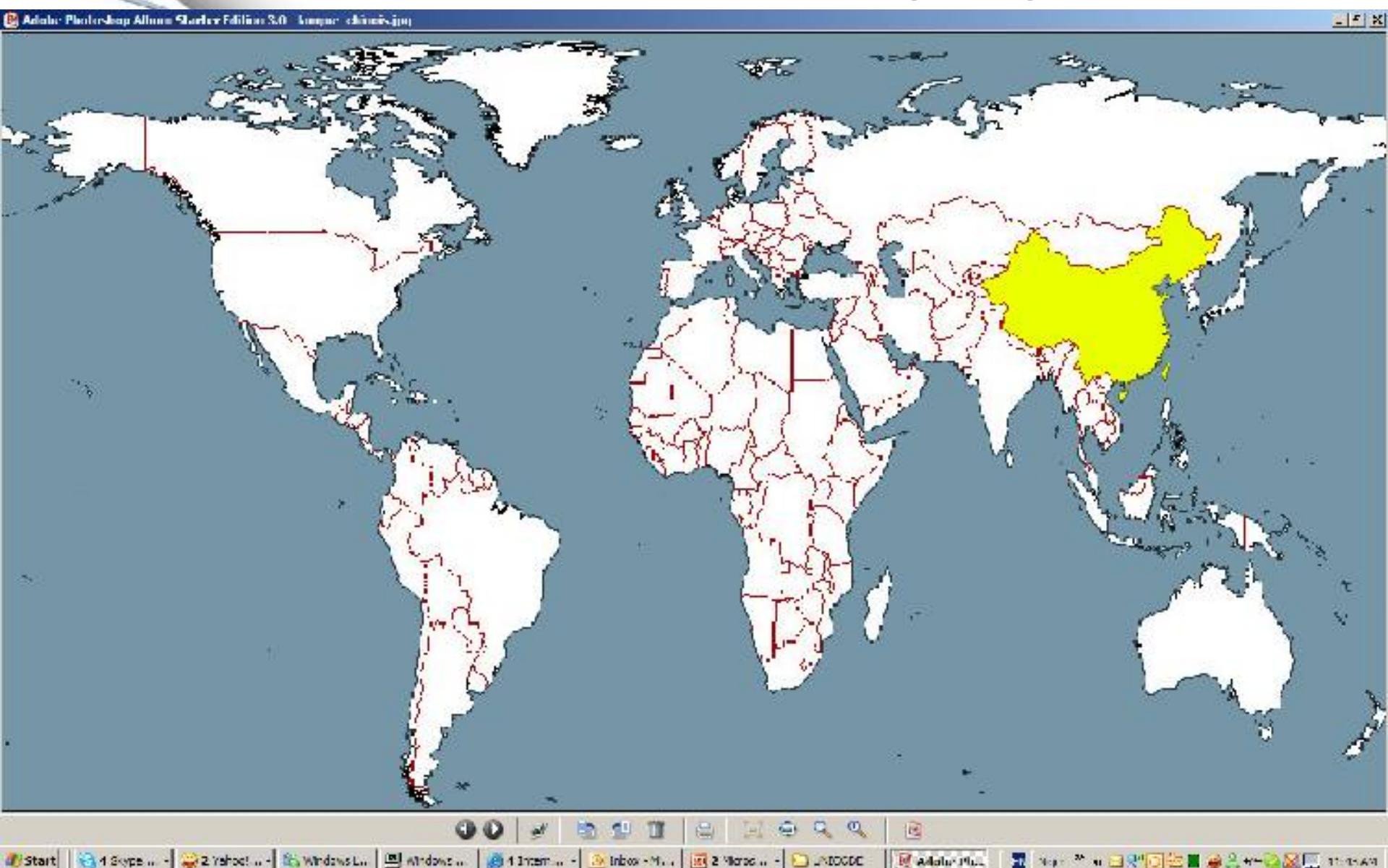


Arab Language in the World

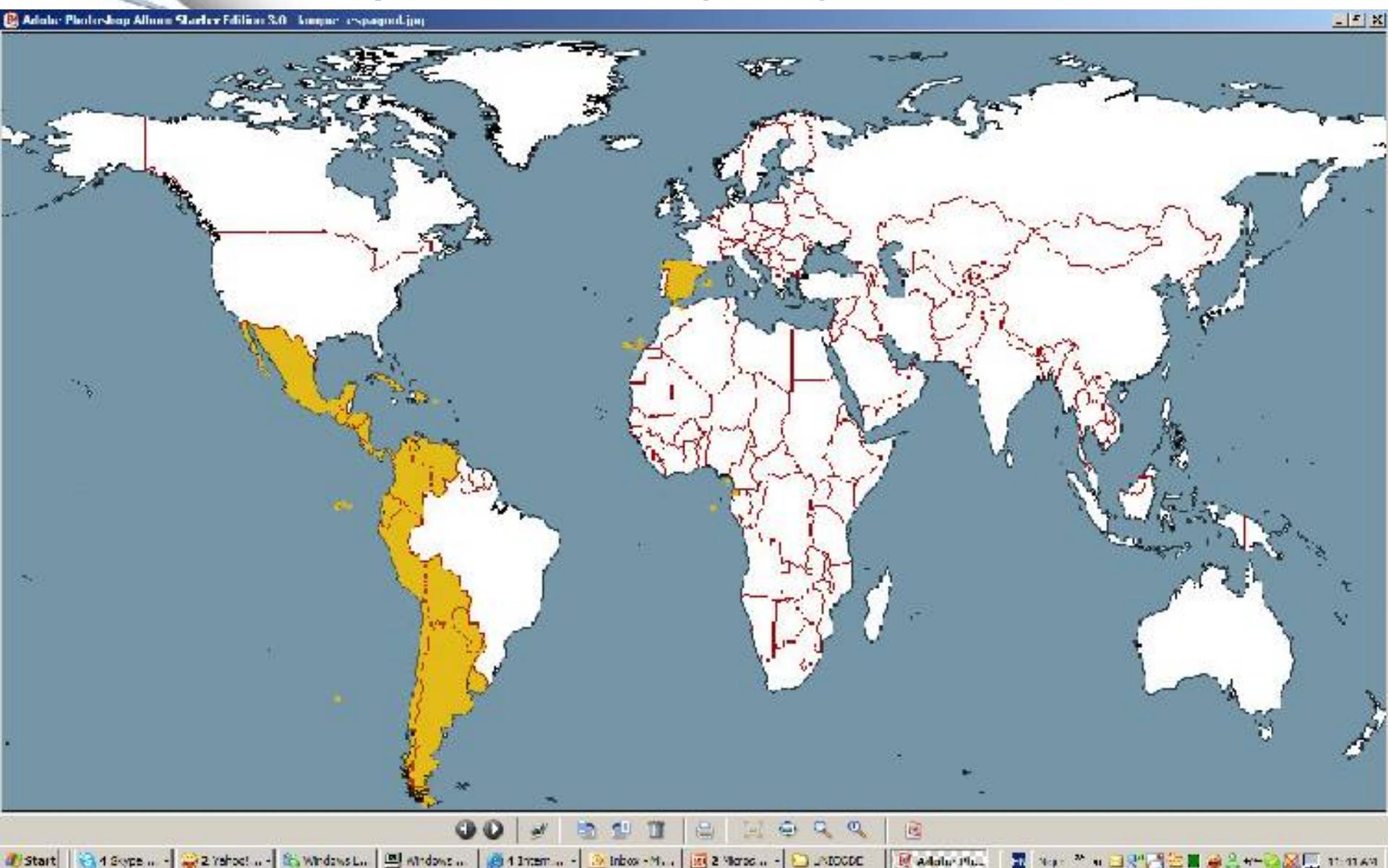
ICANN



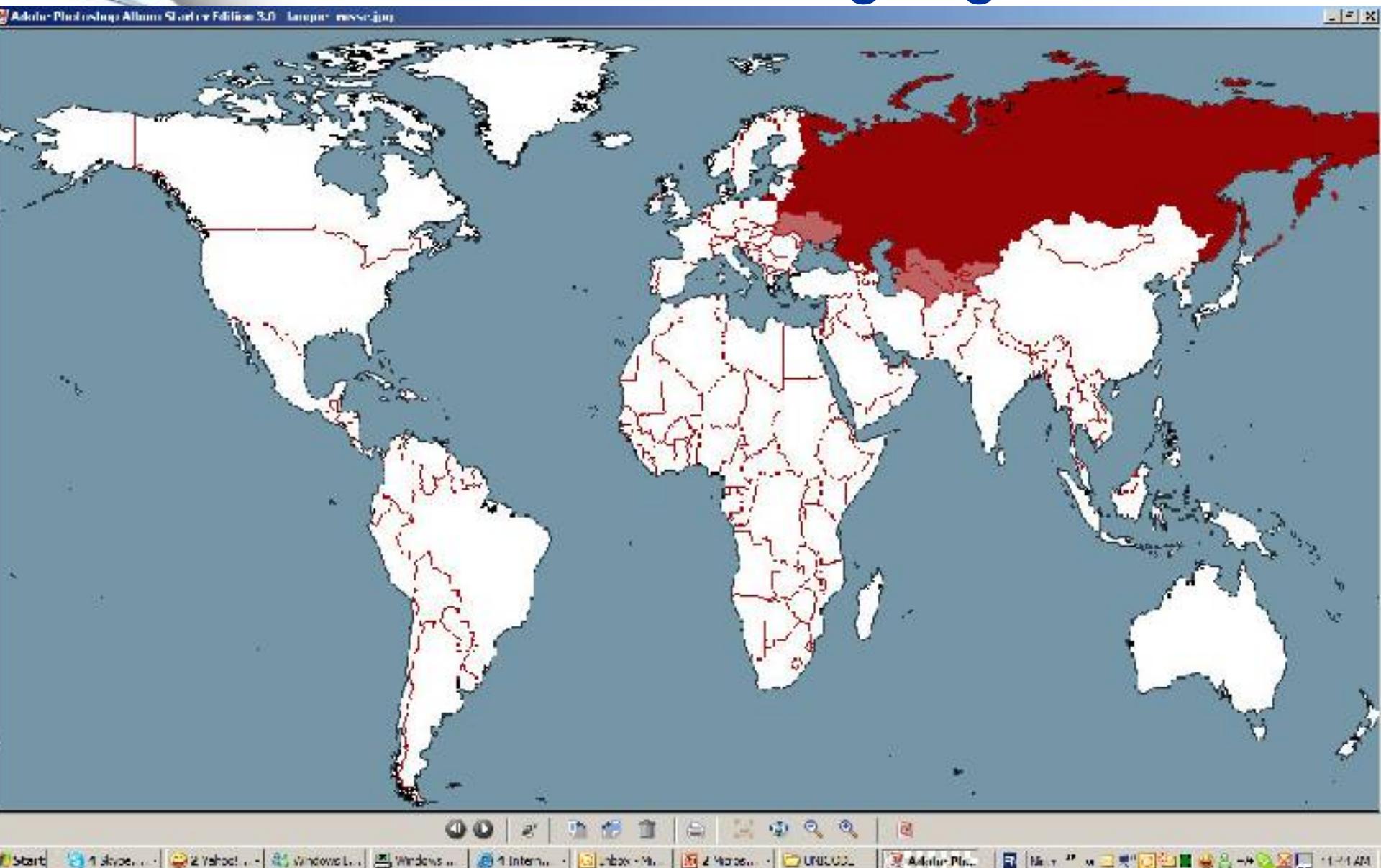
Chinese Language



Spanish language in the World



Russian Language ...



UNESCO – Cultural Diversity

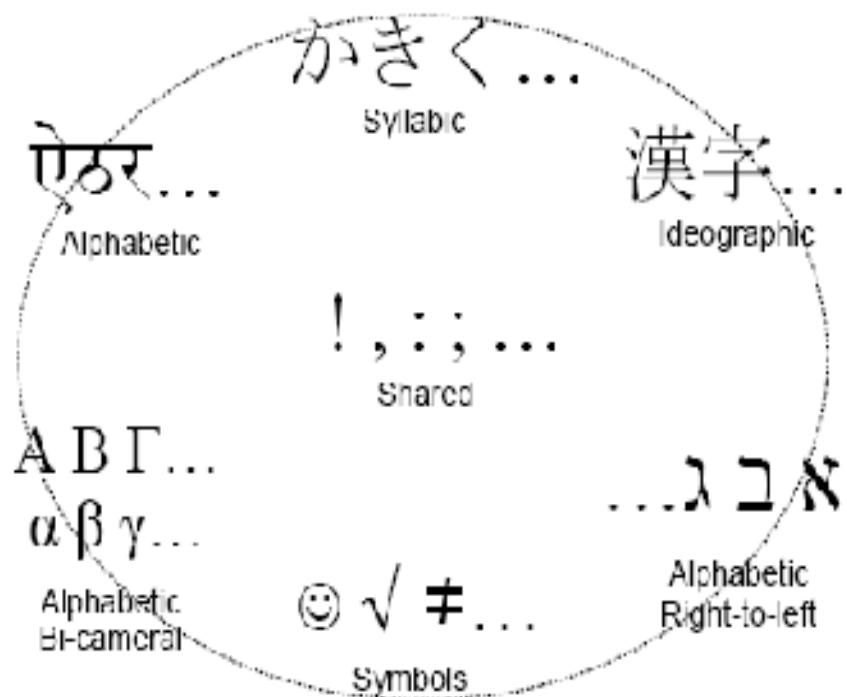
- *Partager un monde de différence : la diversité linguistique, culturelle et biologique de la Terre*, auquel s'ajoute la carte la biodiversité culturelle du monde : peuples, langues et écosystèmes est un matériel éducatif co-publié par l'UNESCO, Terralingua et Fonds Mondial pour la Nature (WWF), destiné aux étudiants et au grand public, et qui introduit le concept de “diversité bioculturelle” en terme de développement durable.

Ethnologue , un des catalogues les plus utilisés des langues du monde, dénombre **6 809 langues utilisées** (la plupart parlées, mais en incluant aussi **114 langues des signes**) dans **228 pays** en l'an 2000.

- *Cependant, moins de 300 de ce grand nombre de langues parlées de par le monde avait des locuteurs de plus d'un million, les langues les plus parlées étant le chinois mandarin, l'hindi, l'espagnol et l'anglais.*

RATIONALE FOR THE PROJECT

- Promoting multilingualism and an African presence on the Net



华人 .公司 .cn

華

高島屋 .会社 .jp

人 .商業 .tw

삼성 .회사 .kr

三星 .회사 .kr

viagénie.qc.ca

الاهرام .م
ישראל .קום

現代 .com

ヤフー .com



THE LAUNCHING OF 2006 AS THE YEAR OF AFRICAN
LANGUAGES AND THE INAUGURAL SESSION OF THE INTERIM
GOVERNING BOARD OF THE AFRICAN ACADEMY OF
LANGUAGES (ACALAN) 20-21 JUNE 2006,
ADDIS ABABA



• LANCEMENT DE 2006, ANNEE DES
LANGUES AFRICAINES
&
SESSION INAUGURALE DU CONSEIL
D'ADMINISTRATION INTERIMAIRE DE
L'ACADEMIE AFRICAINE DES LANGUES
(ACALAN)
20 JUIN 2006
Addis Abeba , Ethiopie



Different african meetings on languages ...

- Addis Ababa African Union Meeting 2006
- Dakar Sept 2005
- Casablanca June 2005
- Bamako May 2005
- Accra February 2005
- Kinshasa 2003
- Bamako 2002
- Etc....

African Initiatives ... www.Bisharat.net

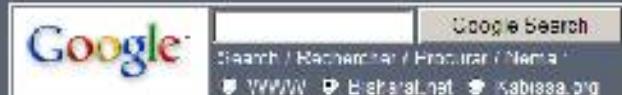


Bisharat A language technology and development initiative

Bisharat Initiative langues - technologie - développement

Bisharat Initiative línguas - tecnologia - desenvolvimento

Bisharat! Tsira aiki bisa halsunan mama - hizak'ar kimiva - kyaatata helin rayawa



NKo Alphabet ...

FontLab Font Table

FontL. Conakry

à	á	â	ã	ä	å	æ	ç
è	é	ê	ë	ì	í	î	ï
ð	ñ	ò	ó	ô	õ	ö	÷
ø	ù	ú	û	ü	ý	þ	ÿ
□			≠	∞	≤	≥	∂
Σ	Π	π	∫	Ω	√	◊	ı
„	„	„	„	„	„	„	„
ẽ	Ѐ	Ѐ	Ѐ	Ѐ	Ѐ	Ѐ	Ѐ

FontLab Font Table

FontL. Conakry

í	Ł	ł	ń	ń	ć	ć	đ
ś	Σ	ſ	ū	ū	ž	ž	ȝ
ȝ	fi	fl	¢	○	/	≈	-
·	Δ	μ	Ω	ο	ř	ř	đ
‡	†	‡	‡	‡	‡	‡	‡
Ƴ	ѧ	ѧ	՚	՚	՚	՚	՚
՚	՚	՚	՚	՚	՚	՚	՚
՚	՚	՚	՚	՚	՚	՚	՚

Casablanca Statement

- African localisation experts met in Casablanca in a workshop organised by Kabissa with Bisharat under IDRC funding, and in collaboration with MTDS and the Casablanca Technopark centre. The event benefitted from contributions from the Moroccan Minister-Delegate to the Prime Minister in Charge of General and Economic Affairs, the Canadian Ambassador to Morocco, and experts from other continents.
- After three days of work, the participants in the meeting reached the following conclusions:
- Limiting people to the use of information and communication technology (ICT) in a foreign language tends to exacerbate the digital divide; makes ICT adoption long, difficult, and expensive; and impoverishes local culture.
- Localisation makes ICT more accessible to everybody, including users from rural areas and young students, reinforcing the importance of our culture and helping us preserve our identity.
- Localisation of ICT into indigenous African languages is therefore key to rapid and fair development in Africa.
- For localisation to succeed and have its maximum impact in society, collaboration among governments, civil society, educators, linguists, computer professionals, standards organisations and development agencies is necessary.
- We, the participants, commit ourselves to promoting this vision and working towards social development in Africa through ICT localisation.

- Casablanca, 15 June 2005

Initiatives ...www.Evertype.com



 Evertype Home Page



Cloghrafolír • Typographer

- Michael Everson
- Evertype
- Typesetting Portfolio
- Typography & scripts

Clófhórlne • Fonts

- CelticScript
- Excursion Mono
- Celtic typeface history
- Celtic keyboard layouts
- NKO font development

Leabhair • Books

- The Unicode Standard
- An Íobad (The Book in Irish)
- English-Cornish Dictionary Second edition!
- I astall den Scathan (Printing Glass in Irish)
- Eschtrai Ellso (Alice in Irish)
- Testament Knwell
- Breton Grammar



Calgheáin • Standards

- Unicode & ISO/IEC JTC1/SC2/WG2
- Codes for Internationalization
- ISO TC46/WG3
- I-SAMU / SC23/WG4
- The Script Encoding Initiative

Téangacha • Languages

- The Alphabets of Europe
- List of Language Lists
- Celtic and other languages
- Linguistic rights

Bogearrai • Software

- Alighairtach text editor requirements
- Celtic spelling checkers
- Mac OS software

Nascanna eile • Other links

- The plural of euro is euros!
- Essays and interviews
- Poems, stories, & other things

Copyright © 1993-2000 Evertype
All Rights Reserved

Ethiopian Initiative



Amharic Spelling is Flexible

Symbol Redundancy / Phonemic Decay
The Ge'ez Syllabary

	ə	u	i	a	e	(i)	o
h	v	v̄	v̄̄	γ	γ̄	v̄	v̄̄
l	ʌ	ʌ̄	ʌ̄̄	ʌ	ʌ̄	ʌ̄	ʌ̄̄
ḥ	d̄	d̄̄	d̄̄̄	d̄	d̄̄	d̄̄	d̄̄̄
m	əm	əm̄	əm̄̄	əŋ	əŋ̄	ŋ̄	ŋ̄̄
s	v̄	v̄̄	v̄̄̄	v̄l	v̄̄l	z̄	z̄̄
r	č	č̄	č̄̄	č̄	č̄̄	č̄̄̄	č̄̄̄̄
f	ñ	ñ̄	ñ̄̄	ñ̄	ñ̄̄	ñ̄̄̄	ñ̄̄̄̄
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Multilingualism

- As the Internet continues to grow, many people around the world are eager to go online using their native languages
- Multilingualism has been one of the key-issues through the WSIS process
- Paragraph 53 of the Tunis Declaration has ensured the commitment towards the multilingualization of the Internet
- It is not only about domain names, it is more about making the Internet adapted to the “local” needs
- This involves not only domain names, but also content, user interfaces, e-mail addresses, search engines, etc...
- But domain names have received more global attention than other items

The International Multi-stakeholder Organisation of the 21st Century:



- Transnational
- All stakeholders represented
 - Including governments with choice of relevant agency or agencies
- Flexible in organisational management
- No capture by individuals, groups, or organisations
- Reflective of its own regime – that is, 35 years of the Internet's players at work
- Focus on effectiveness and relevancy
- Cooperation, Coordination and Collaboration



What is IDN?

- Internationalized Domain Name is a domain name that contains non-Latin scripts or Latin scripts with diacritics
- Transition from an ASCII-based DNS to a Unicode (ISO 10646)-based DNS would have been too complex
- An algorithm is needed to transform Unicode into ASCII

Background

- Two core resources of the Internet: IP Addressing and Domain Names
- A Domain Name Server (DNS) is used to map names to numbers and visa-versa
- DNS was originally restricted to case-insensitive ASCII letters (a-z), digits (0-9) and Hyphen (LDH)
- Should the DNS support more characters than the basic Latin?

Some IDN Milestones

- 1996 – 1998: Research activities at the University of Zurich and at the National University of Singapore
- 1998 – 2000: National and regional groups in Asia Pacific, deploying test beds and demonstrating prototypes in international conferences
- January 2000: IETF IDN working group formed
- June 2000: Inaugural Launch of the Multilingual Internet Names Consortium (MINC)
- March 2001: ICANN IDN working group formed
- April 2001: Arab Internet Name Consortium formed
- March 2003: Publication of RFC 3490, RFC 3491 and RFC 3492
- June 2003: Publication of ICANN IDN Guidelines version 1.0
- March 2004: Arabic Domain Name Pilot Project
- November 2005: Publication of ICANN IDN Guidelines version 2.0
- November 2005: ICANN President Advisory Committee for IDNs
- February 2006: Publication of ICANN IDN Guidelines version 2.1
- March 2006: IDN PAC announced a proposed timetable for IDN technical tests

IETF Contribution

- RFC 3490: Internationalizing Domain Names in Applications (IDNA)
- RFC 3491: Nameprep: A Stringprep Profile for Internationalized Domain Names (IDN)
- RFC 3492: Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA)

ICANN Guidelines

- A list of general standards for IDN registration policies and practices
- Registries seeking to deploy IDNs under their agreements with ICANN have been authorized to do so on the basis of the Guidelines
- IDN TLD Working Group
 - ccNSO and GNSO constituency membership
 - Provide regular updates to the Presidents IDN Advisory Committee
- Version 1.0
 - Commitment by .cn, .jp, .tw, .info, and .org
 - Further authorized and implemented by .info, .museum, .biz, .org, .name, .com and .net
- Version 2.0
 - Opened for script based implementation in addition to language based
 - No mixing of scripts within labels
- Version 2.1
 - Included the provision of the “UseSTD3ASCIIRules” flag

Technical Test Project Items

- Project Timeline (posted)
 - March: Initial RSSAC and SSAC consultation
 - April: development of draft objectives
 - May: public comment period
 - June: Revision and approval
 - July: initiate test launch pending technical upgrades
- Project Objectives (underway)
 - Selection criteria and required commitment for applicants
 - Life time for test labels
 - Definition and certification of test labels restrictions
 - Legal considerations
 - Definition of IANA process for insertion of IDN TLDs
 - Test results display and evaluation

Two Approaches

- NS-records - permit the insertion an internationalized label (in punycode) in the root zone:

<domain>.<tld> does not necessarily resolve to the same Internet address as
<domain>.<idn-tld>

- DNAMEs – as defined in RFC 2672 DNAMEs is the equivalent mapping of one domain into another:

<domain>.<tld> resolves to the same Internet address as <domain>.<idn-tld>

« UNICODE & IDN in Africa »

Thème : « Informatisation des Langues Africaines sur Internet »
 Novotel-Dakar Mercredi 07 septembre 2005

FEUILLE DE PRESENCE

N°	Nom & Prénoms		Organisation service	Contacts Tél.-Fax	E-mail
1	NIANG	Aly Abdoulaye	Direction de l'Alphabétisation et des langues Nationales	P :655 82 75 D : 842 24 82	aliniang@hotmail.com
2	SARRY	Ouleymatou sarr	Direction de l'Alphabétisation et des langues Nationales	P :573 57 32 Fax :842 24 83	uleysarr@yahoo.fr
3	NDIAYE	Alassane	Direction de l'Alphabétisation et des langues Nationales	P : 658 50 03	alassnd@yahoo.fr
4	SYLLA	Yéro	IFAN cheikh Anta DIOP SAFEOFOD	Tél :824 44 06 Por :644 58 58 F : 824 14 59	yero@safefod.org
5	FAYE	Assane	Président du MELAS	Tél :825 50 89	
6	Domingo	Madeleine	Commission Nationale de la Francophonie	tél. :579 57 90	mado_domingo@yahoo.fr
7	DON	Osborn	Bisharat , LTD	Tél :+86 28 405561	dzo@bisharat.net
8	LISHOU	Claude	ESP UCAD L.T.I	Tel:825 6164 Por:638 41 85	claudelishou@yahoo.fr
9	CORENTHIN	ALEX	ESP/UCAD	Port:667 07 75	
10	GUEYE	Fatim	Commission Nationale pour la Francophonie	Tél :849 76 95 Fax :821 40 87	fatingueye@yahoo.com cnfs@refer.sn
11	DIAGNE	Abdoulaye	Netplus	Tél :864 48 59 Fax :868 48 59	netplus@sentoosn
12	LEHENTO	Ven	Panos AO/CIPACO	Tél :849 16 65	klohento@panos-ao.org
13	DANDJINOU	Pierre	PNUD Projet UNICODE IDN Afrique	Tél :869 06 22	dandjinoupierre@undp.org
14	DIOP	Mouhamet	NEXT	Tél :821 42 02	mouhamet@next.sn
15	SOMé	Maxime	MODYCO/P.X	Tél : +331 4025849	maxsome@club-internet.fr
16	OUEDRAGO	Pierre	AIF /INTIF	Tél :+335560159 23	po@francophonie.org
17	AIDARA	Nana	ART Du Sénégal	Tél :869 03 69 Fax : 869 03 70	Nana.aidara@art.sn



18	NDIAYE	Amadou	Ministère de la communication DLL	Tél : 823 32 75 Fax : 842 34 09	ndiayeama@hotmail.com
19	LOPIS SYLLA	Jeanne	IFAN- C hAnta Labo de Linguistique	Tél :820 93 83 Fax :644 56 41	jeannelopis@yahoo.fr
20	DIOP	Bagley Amina	Digital Freedom Initiative BOKK JANK	Tél :820 60 79	aminatty@gmail.com
21	DOUMBOUYA	Mamady	NIKO INSTITUTE	Tél :0012152480 111	mamady@nkoinstitute.com
22	EVERSON	Michael	EVERTYPE		everson@evertype.com
23	ADEGBOLA	Tunde	African Languages Technology initiative Ibadan Nigeria	Tél:+ 234 34019398	Tadegbda@alt-i.org
24	DIALLO	Issa	CNRST-Ouaga BF	Tél: 226 70251195	issadbble@hotmail.com
25	TCHITCHI	Toussaint Yaouvi	Université d'abomey-calavi BENIN	Tél:229 21 32 1418 Tél:229 90 (90) 67 22	tchitchitoussaint@yahoo.fr
26	MBODJ	Chérif	CLAD/Universté ch Anta DIOP Dakar	Tél:824 52 06 Por: 631 47 47	chembodj@refer.sn
27	FALL	Arame	Linguiste à la retraite	Tél: 825 39 10 Port: 637 63 65	abd@sentoosn
28	FAYE	Souleymane	CLAD-UCAD	Port: 642 10 95	karuseer@yahoo.fr
29	RAMAVONIRINA	Oliva	Centre des Langues de l'Academy Malgache	Tél :00 261 2255997	Olivia.ramavonirina.r@free.fr
30	DIKI-KIDIRI	Marcel	CNRS LLCAN	Tél :0549301769	madkidiri@free.fr
31	WANE	Mohamedoune	Dept Langues Nationales et de Linguistiques Université de Nouakchott	Tél :00 222 529 233	mamahamath@yahoo.fr douwahn@yahoo.fr
32	SAMASSEKOU	Adama	Acalan	Tél :2232238447 Tél :2232238450	acalan@timbagga.com.ml
33	INNE	Anne-Rachel	ICANN	Tél :+132330434 20	inne@icann.org annerachel@gmail.com



IDN Experimentation

- Part III - Experimentation with IDN - Implementation at the second level with (two) ccTLDs
- Multilingualism in the Cyber space.
- ICANN - IDN Guideline 2.0
- IETF publish a series of RFCs:
 - namely RFC3490, RFC3491, RFC3492, RFC3454 and
 - RFC3743 for the risk of “phishing” in IDNs

IDN implementation with working groups

IDN implementation with working group



Scaling Up -

3.1 Second level Implementation ccTLDs (two countries)

3.2 African perspective on IDN implementation

4.1 Pilot generalized

- Identification of best practices in ccTLD management

- Variant Table for registration
- Adopt a registration policy
- Impact assessment

- Develop Business Plan

- Advocacy towards African Institutions (AU, NEPAD, etc)
- Funds raising

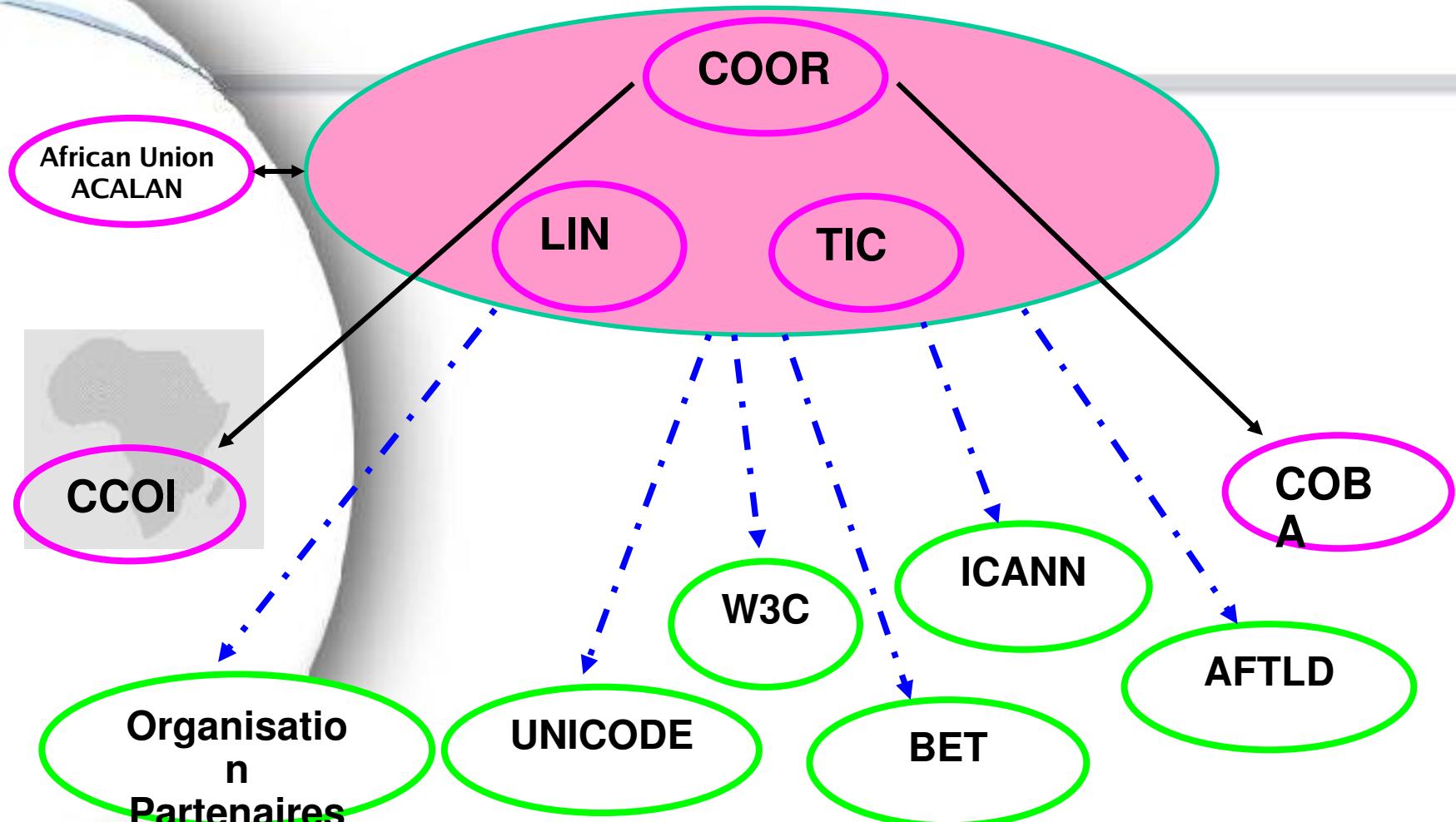
Technical assistance
Technical & Business Partnership Missions Consultancies

=====

=====

Consultations
Missions
Regional Workshop

PROJET “UNICODE-IDN in AFRICA”



COPI: Comité de Pilotage; COOR: Comité de Coordination

COBA: Comité des Bailleurs de Fonds; BCT: Berkeley Encoding Institute

AFTLD: African Top Level Domains; MINC



Questions ?