




Objectifs de l'atelier et programme de la journée

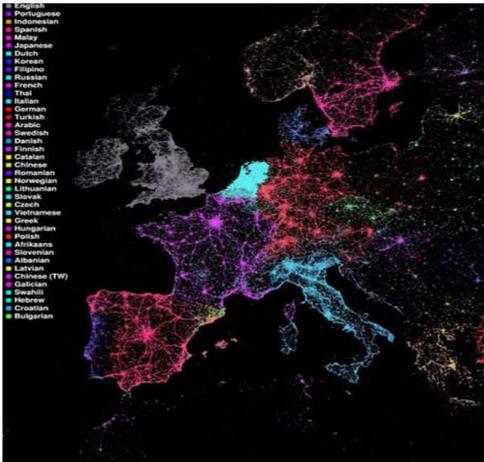
Khalid Choukri
Consortium "Coordination Européenne de Ressources Linguistiques"
(European language Resource Consortium, ELRC)

ELRC Atelier France 2016/05/11 Languages and Language Technologies 1


Le contexte


- L'Europe est multilingue
- 24 langues officielles de l'UE + 2 CEF
- Les langues sont au cœur de notre richesse, de nos cultures et des identités européennes
- L'UE soutient activement le multilinguisme

- English
- Portuguese
- Indonesian
- Spanish
- Malay
- Japanese
- Dutch
- Korean
- Filipino
- Russian
- French
- Thai
- Italian
- German
- Turkish
- Arabic
- Swedish
- Danish
- Finnish
- Catalan
- Chinese
- Romanian
- Norwegian
- Lithuanian
- Slovak
- Czech
- Vietnamese
- Greek
- Hungarian
- Polish
- Albanian
- Slovenian
- Albanian
- Latvian
- Chinese (TW)
- Galician
- Danish
- Hebrew
- Croatian
- Bulgarian



ELRC Atelier France 2016/05/11 Languages and Language Technologies 2



Le contexte





ELRC Atelier France 2016/05/11
Langages and Language Technologies
3

- Les frontières créées par les langues sont-elles infranchissables ?
- Elle limitent la circulation des personnes, des informations, des idées, et sont un obstacle au commerce
- Aucune discrimination par la langue, afin de garantir l'égalité des chances
- Les langues vues comme un atout, un enrichissement, et non un problème



Le contexte





ELRC Atelier France 2016/05/11
Langages and Language Technologies
4

- Quelles solutions pour gérer un espace multilingue ?
 - la traduction !
- Traduction humaine: des professionnels extrêmement qualifiés, ayant suivi un long parcours de qualification !
 - Mais il y a beaucoup de données à traduire... et les besoins de traduction sont considérables
- Le recours aux nouvelles technologies pour traduire :
 - Traduction "soutenue" par la technologie => Traduction Automatique (TA)
 - la traduction assistée par ordinateur (TAO)
 - Dans tous les cas, la technologie au service de la traduction



Le contexte



- Les nouvelles technologies au service des infrastructure en Europe
 - On vient de voir l'ambition du programme MIE (ou CEF)

- L'UE et la vision d'un marché unique numérique
(*DSM*) pour *Digital Single Market* (*une vision et un plan*)

- Le marché unique numérique ne peut être que multilingue

- Un programme dédié à la traduction automatique : CEF-AT

ELRC Atelier France 2016/05/11
Languages and Language Technologies
5



Le contexte



- Traduction automatique à la CE
 - depuis de nombreuses années ...
 - Actuellement la traduction automatique à la CE → MT@EC
 - nouvelle génération CEF-AT

- La nouvelle plateforme CEF-AT a pour objectif de mieux prendre en compte les besoins :
 - de mettre l'accent sur les besoins des citoyens, des services publics, des administrations, etc.
 - des ONG, ...
 - Bref, de nos tous

- Mais pour mieux répondre aux besoins
 - Il faut "personnaliser" la TA

- La gratuité de CEF-AT

ELRC Atelier France 2016/05/11
Languages and Language Technologies
6



Données



- La TA est aujourd'hui:
 - basée sur un processus d'apprentissage à partir de données
 - apprend à partir de traductions humaines = données / Ressources Linguistiques
 - besoins de ressources linguistiques
- Pour mieux répondre à vos besoins, nous devons utiliser les meilleures données
- L'adaptation de l'outil TA à vos besoins ne peut s'envisager sans vos données
- Vos données permettront d'améliorer les résultats de la TA pour vous et pour CEF-AT
- Sans les bonnes données, la TA ne fonctionne pas!

ELRC Atelier France 2016/05/11

Languages and Language Technologies

7



Les objectifs de la journée



Nous sommes ici aujourd'hui ...

- Pour établir des partenariats avec vous afin de trouver les bonnes données pour CEF-AT
- Pour recenser et partager des données à intégrer dans CEF-AT, au bénéfice de l'UE (et de ses citoyens)
- Renforcer notre soutien au multilinguisme dans l'UE
- Et pour soutenir le Français dans l'UE, ainsi que les 25 autres langues

ELRC Atelier France 2016/05/11

Languages and Language Technologies

8



Les acteurs de cette expérimentation



- Les principaux acteurs:
 - Les services publics, les administrations, les ONG, ...
 - La Commission européenne (CEF, MT@EC, ...)
 - Les pays de l'UE, La Norvège et l'Islande
 - Le consortium "Coordination Européenne de Ressources Linguistiques"
 - un consortium mandaté par la CE pour cette mission (4 partenaires)
 - dans chaque pays, un quatuor :
 - un point d'ancrage "Technique" et un point d'ancrage "Public",
 - le représentant local de la DGT,
 - un représentant du consortium (ELDA pour la France)
- Objectif : identifier les besoins et structurer un réseau d'acteurs publics locaux.

ELRC Atelier France 2016/05/11

Languages and Language Technologies

9



Objectifs des ateliers



- Sensibiliser sur la valeur des données détenues par les acteurs publics et l'importance de leur mise à disposition pour surmonter les barrières linguistiques ;
- Identifier les données partageables pour adapter CEF-AT:
 - Pour répondre aux besoins des services publics dans les pays de l'UE +CEF
 - Et établir un partenariat avec vous et vos équipes en vue d'un recensement régulier des données partageable avec CEF-AT
- Présenter les futurs services CEF.AT (qui seront gratuits) ;

ELRC Atelier France 2016/05/11

Languages and Language Technologies

10



Objectifs des ateliers



- Présenter le "Guichet" d'assistance
 - mis en place pour répondre aux questions juridiques et techniques
 - questions associées aux données détenues par les administrations publiques
 - questions liées à la collecte et au partage des données.

ELRC Atelier France 2016/05/11
Languages and Language Technologies
11



Programme de la matinée



08:30 – 09:00	Inscription
09:00 – 09:05	Accueil des participants
09:05 – 09:25	Allocation de bienvenue : Commission européenne et DGT (Direction générale de la traduction) Isabelle Jégouzo, Chef de la Représentation (sous-réserve) Isabelle Tranchant (direction générale de la traduction, cheffe du département de langue française)
09:25 – 09:40	Le Programme "Mécanisme d'interconnexion européen" (MIE) ou Connecting Europe Facility, (CEF) , Jean Jacques LEANDRI, Secrétaire Général à la Modernisation de l'Action publique, SGMAP Représentant de la France au comité MIE
09:40 – 09:50	Objectifs de l'atelier et programme de la journée Khalid Choukri, (Consortium ELRC, ELDA)
09:50 – 10:05	L'Europe et le multilinguisme Kimmo Rossi, Commission européenne (DG CONNECT, Unit G3 – Data Value Chain)
10:05 – 10:25	Les Langues et les technologies en France Loïc Depecker, Délégué général, DGLFLF, Ministère de la culture
10:25 – 10:45	Les différentes technologies de la langue Pierre Zweigenbaum/Joseph Mariani, LIMSI-CNRS
10:45 – 11:15	Pause café
11:15 – 12:05	Table ronde: "Le multilinguisme et les services publics en France" Modérateur : Loïc Depecker, DGLFLF, Ministère de la culture Intervenants : Alain Repaux, chef du centre de traduction des ministères économiques et financiers, Jean-Michel Thivel, Secrétaire général des affaires européennes (SGAE)
12:05 – 12:35	Traduction automatique : Principes, état de l'art et usages (Comment ça marche ?) François Yvon, LIMSI-CNRS
12:35 – 13 :00	Comment les services publics peuvent-ils bénéficier de la plateforme CEF.AT ? Kimmo Rossi, Commission européenne (DG CONNECT, Unit G3 – Data Value Chain)

ELRC Atelier France 2016/05/11
Languages and Language Technologies
12

Déjeuner
13:00 – 14 :00

NOVOTEL PARIS BERCY
85, Rue de Bercy
75012 Paris
France
Téléphone : 01.43.42.30.00
FAX : 01.43.45.30.60

ELRC Atelier France 2016/05/11
Languages and Language Technologies
13

Programme de l'après-midi

13:00 – 14 :00	Déjeuner Hôtel NOVOTEL PARIS BERCY, 85, Rue de Bercy, 75012 Paris, Téléphone : 01.43.42.30.00
14 :00 – 14:20	De quelles données la Commission a-t-elle besoin ? Aspects techniques et pratiques Khalid Choukri, Consortium ELRC, ELDA
14:20 – 14:50	Quel cadre juridique pour la mise à disposition des données publiques concernées ? (Directive européenne sur la réutilisation des données publiques) Danielle Bourcier, CNRS, Responsable du groupe «Droit gouvernance et technologies » au CERSA
14:50 – 15:40	Table ronde : "Données linguistiques en France" Modérateur : M. E. Geoffroy (Agence Nationale pour la Recherche) Intervenants : Mme Nadia Amellah-Chikh, Légifrance M. Stéphane Cottin, Secrétariat général du gouvernement M. Brian Stacy (Agence Nationale pour la recherche)
15:40 – 16:10	Pause café
16:10 – 16:45	Session interactive : Comment contribuer en pratique au CEF-AT ? Animateur: Khalid Choukri, Consortium ELRC, ELDA
16:45 – 17:00	Conclusion : Quel suivi ? Quels engagements ?

ELRC Atelier France 2016/05/11
Languages and Language Technologies
14



En fin de journée



- Recenser et partager des données pour CEF-AT
- Les données issues des administrations, des services publics et des ONG dans toute l'UE et dans les pays affiliés CEF.
- Fournir les meilleurs services adaptés à vos besoins et à vos textes (CEF-AT)
- Soutenir l'infrastructure CEF
- Petite contribution au multilinguisme et au marché unique numérique multilingue

ELRC Atelier France 2016/05/11
Languages and Language Technologies
15



Soutenir nos langues c'est soutenir l'Europe
et soutenir l'Europe c'est soutenir nos langues.

ELRC Atelier France 2016/05/11
Languages and Language Technologies
16