

### Multilingual content management and standards with a view on AI developments

Laurent Romary

### ▶ To cite this version:

Laurent Romary. Multilingual content management and standards with a view on AI developments. AI4EI - Conference Artificial Intelligence for European Integration, Oct 2020, Turin / Virtual, Italy. hal-02961857

### HAL Id: hal-02961857 https://inria.hal.science/hal-02961857v1

Submitted on 8 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Multilingual content management and standards with a view on AI developments

**Laurent Romary** 

Directeur de Recherche, Inria, team ALMAnaCH ISO TC 37, chair

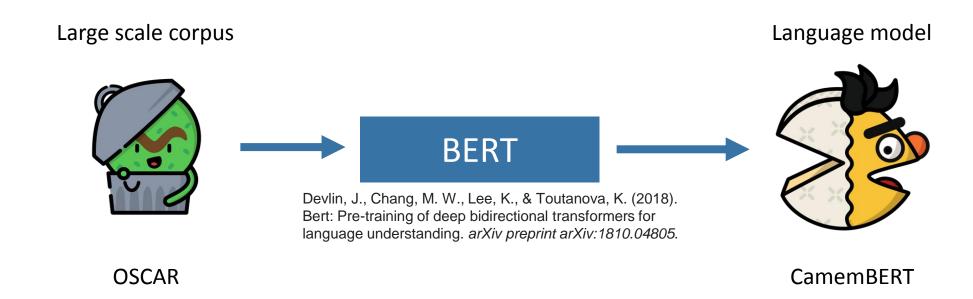




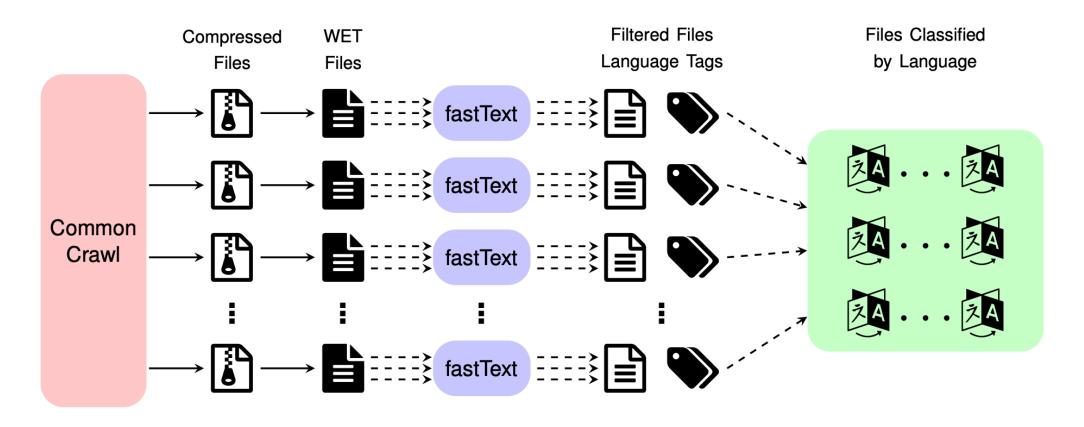
### Language and Al

- Central role of language in the revival of AI (machine-learning based models)
  - Applications: document management and understanding, chatbots, machine translation
  - Information sources: public (web, cultural heritage repositories) and private (Siri, Amazon Alexa) linguistic information
  - European context: cf. Europe's Languages in the Digital Age, META-NET White Paper Series
- Variety of linguistic forms
  - Spoken, written, chats and forums
  - Multilingualism, accents, dialects, technical domains, registers, language learners
    - General notion of language variety
- Classifying and referencing the relevant features
  - Role of standards and standards developing organization (SDO)

### A concrete example for a start



## The infrastructure to create OSCAR - goclassy



Pedro Javier Ortiz Suárez, Benoît Sagot, Laurent Romary. Asynchronous Pipeline for Processing Huge Corpora on Medium to Low Resource Infrastructures. 7th Workshop on the Challenges in the Management of Large Corpora (CMLC-7), Jul 2019, Cardiff, United Kingdom. (10.14618/IDS-PUB-9021). (hal-02148693)

### **OSCAR**

Language	Size		Words	
	Orig	Dedup	Orig	Dedup
English	2.3T	1.2T	418,187,793,408	215,841,256,971
Russian	1.2T	568G	92,522,407,837	46,692,691,520
Spanish	278G	149G	47,545,122,279	25,928,290,729
French	282G	138G	46,896,036,417	23,206,776,649
German	308G	145G	44,878,908,446	21,529,164,172
Italian	137G	69G	22,248,707,341	11,250,012,896
Portuguese	124G	64G	20,641,903,898	10,751,156,918
Chinese	508G	249G	14,986,424,850	6,350,215,113
Japanese	216G	106G	4,962,979,182	1,123,067,063
Polish	109G	47G	15,277,255,137	6,708,709,674
Total OSCAR	6.3T	3.2T	844,315,434,723	425,651,344,234

### Language and standards — why?

- A general issue of reproducibility
  - Can we compare two linguistic descriptions
    - E.g. two etymological assumptions in a dictionary entry
- An acuter context created by the machine learning revival
  - Comparing results of AI components
    - Issues related to segmentation, categorisation (noun, verbs, feminine, dual, etc.), dependencies towards other linguistic descriptions (lexica, lower linguistic levels)
  - Reproducibility of results
    - Can we compare two training campaigns?
    - How do we deal with cross linguistic comparison?
    - Precise qualification (e.g. coverage) and identification of training corpora

### Standards

- What they are
  - Reference background for the management of a technical process
  - The three pillars of a standard: consensus, publicity, maintenance
  - A constraint for each, a benefit for all!
    - Reading it, understanding it, implementing it...
- What they aren't
  - Something coming from nowhere... participating is always an option
  - Mandatory documents: rather a common ground for a transaction
  - Regulations: unless a state or an international organisation takes it up
- The need for standards developing organizations (SDO)
  - Ensure that the three pillars of standardisation are in place
- Standards are like the EU: you need to understand how they work to appreciate them...

### Standards and Al

- A new committee has been set up
  - IT context: ISO/IEC JTC 1/SC 42 Artificial intelligence
  - Work program 3 publications so far:
    - ISO/IEC 20546:2019 Information technology Big data Overview and vocabulary
    - ISO/IEC TR 20547-2:2018 Information technology Big data reference architecture Part 2: Use cases and derived requirements
    - ISO/IEC TR 20547-5:2018 Information technology Big data reference architecture Part 5: Standards roadmap
  - Numerous ongoing activities:
    - Concepts and terminology, Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML), Risk Management, Assessment of the robustness of neural networks, Overview of ethical and societal concerns, Governance implications of the use of artificial intelligence by organizations
- Overview:
  - Very high level endeavour
  - Requires domain-specific standardisation activities => language

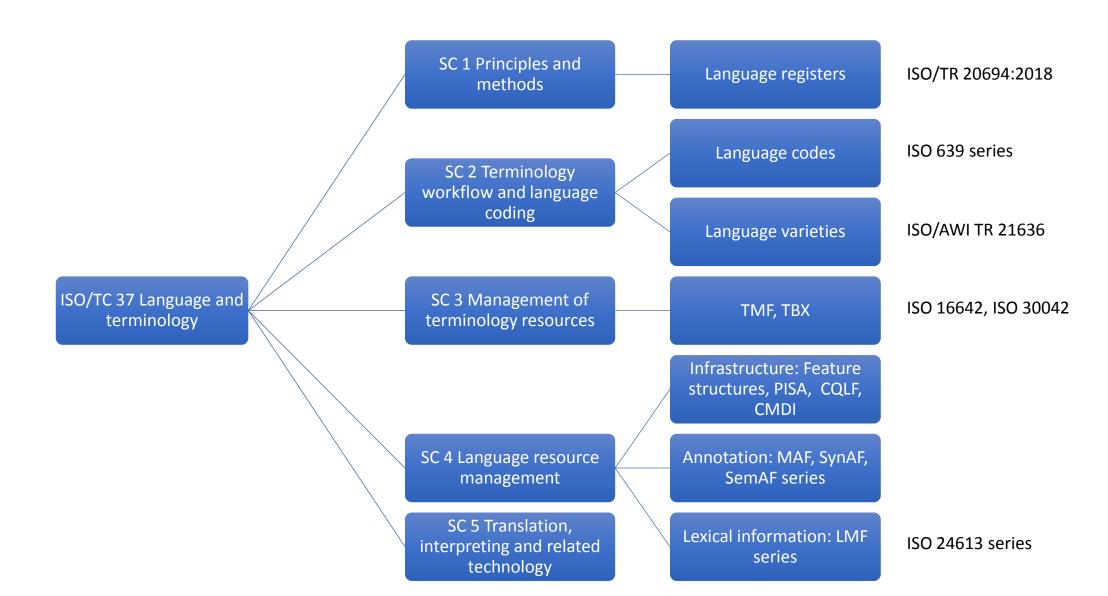
### Language and standards — where?

- Various levels of specialisation
  - Horizontal vs. Vertical standards
- W3C World Wide Web Consortium
  - Development of (horizontal) standards for the Internet
    - Ex. HTML, XML, Xpath, RDF, OWL, SKOS
    - The basic layer upon which language resources are being represented
- ISO International Organization for Standardization
  - General purpose standardisation consortium
  - Organized in technical committees targeted at specific applications
    - ISO TC 37: Language and terminology (language codes, language resources, translation and interpretation)
- TEI Text Encoding Initiative
  - Consortium dedicated to textual content
    - Drama, Transcription of spoken language, dictionaries, manuscripts, etc.
  - Based on XML defines a vocabulary of around 600 elements

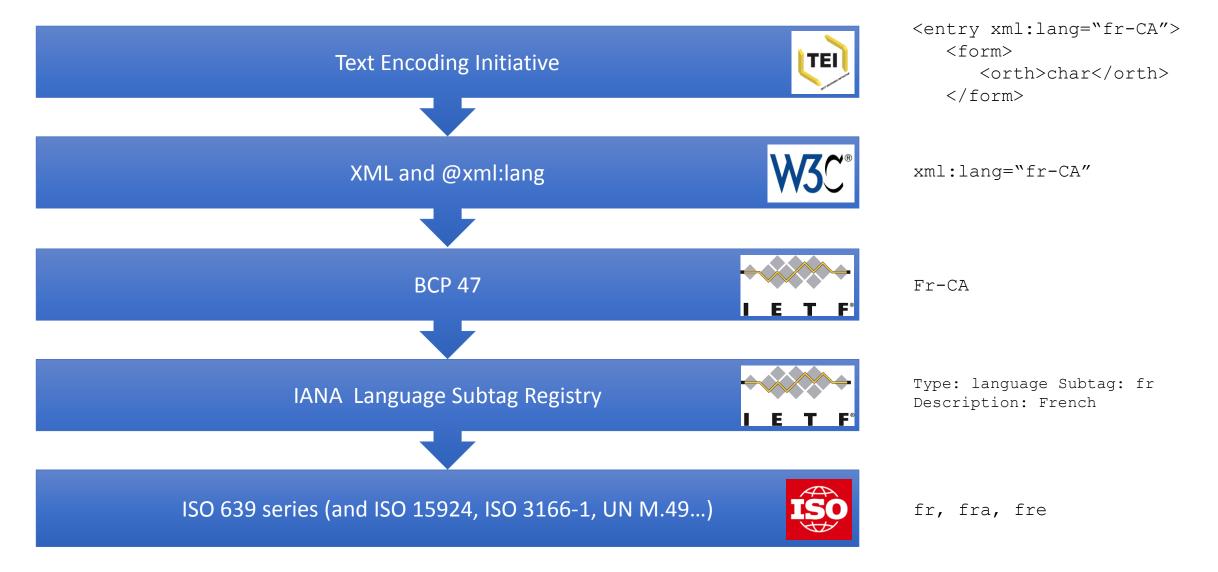
### Language and standards — what and how?

- All levels of linguistic documentation, description, representation
  - Character encoding
  - Structure of documents
  - Linguistic annotation
  - Lexical resources
  - Metadata and bibliography
- The need for a variety of complementary standards
  - Standards piling

### A selection of some ISO/TC 37 activities



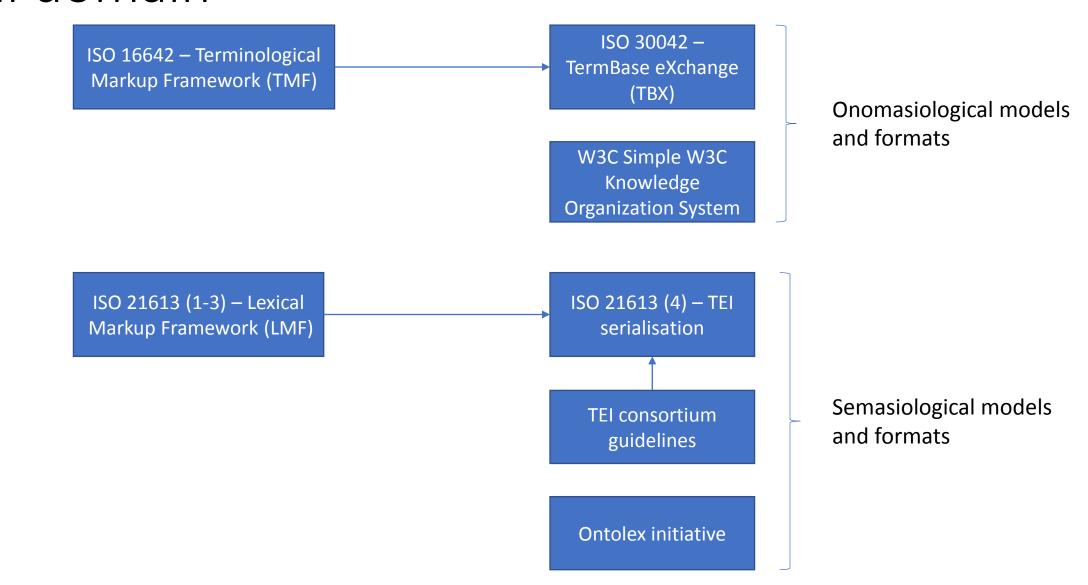
## Piling up standards



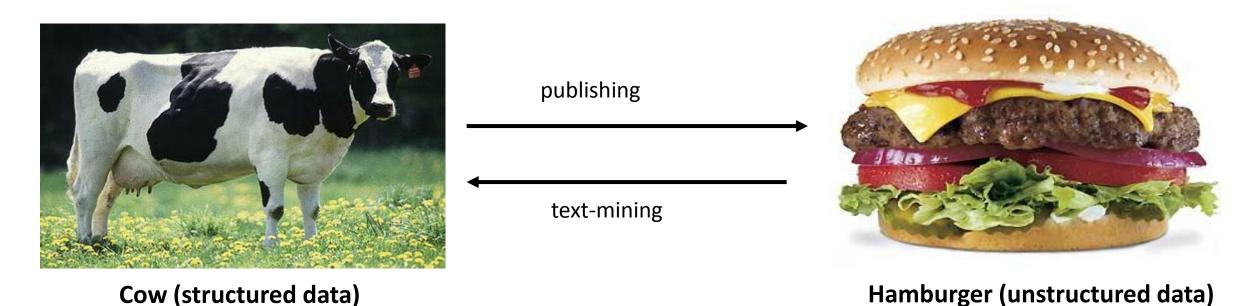
### Dealing with variety: lexical resources

- A variety of contexts and forms
  - Legacy dictionaries, dialectological studies, NLP lexica
  - Onomasiological (concept to term) vs. semasiological (word to sense) structures
  - Lexical (machine processable) vs. Editorial (human readable) views
  - Full forms, etymology, corpus based examples
  - Word document, database, shoebox, XML...
- Why standardizing all this?
  - Defining methods, models and format to facilitate
    - Exchange of lexical data
    - Pooling heterogeneous lexical data
    - Interoperability between software components
      - Search engines, layout, extraction of linguistic properties
    - Comparability of results
      - E.g. Linguistic coverage of lexical databases

# Complementary of standardisation activities in the lexical domain



# Back to machine learning: converting legacy dictionaries into TEI compliant structures



"Converting PDF to XML is a bit like converting hamburgers into cows. You may be best off printing it and then scanning the result through a decent OCR package." Michael Kay (http://lists.xml.org/archives/xml-dev/200607/msg00509.html)

Inspired from: Duncan Hull

### **GROBID-Dictionaries**

- Automatic extraction of TEI structures from digitised dictionaries (Khemakhem et al. 2017)
  - Input: PDF (or OCR/HTR formats such as ALTO)
  - Output: TEI compliant lexical resource
- Spin-off from GROBID (Romary and Lopez 2015)
  - Initiated in 2007
  - Automatic extraction of structural data from scholarly papers
    - Metadata (author, title, affiliations, keywords, abstract), bibliography, ... full text
  - And open source...
- Uses Conditional Random Fields (CRF) (Lavergne et al. 2010)
  - Probabilistic models for sequence labelling tasks

# Cascading CRF models, an intuition

CON

Eugêne IV à Ferrare en 1438-1439, puis à | concordant, e adj. Qui s'accorde : témoi-Florence de 1439 à 1442], de Latran (1512-1517), de Trente (1545-1563) (où fut décidée la réforme générale de l'Eglise catholique en face de la Réforme protestante], de Vati-can I (1870) [où fut défini le dogme de l'infaillibilité pontificale], de Vatican II (1962-1965) (où fut définie l'attitude de l'Eglise romaine à l'égard du monde modernel. conciliable adj. Qui peut se concilier avec

conciliabule [kūsiljabyl] n. m. (lat. conci-

gnages concordants.

concordat [kokorda] n. m. (lat. concordatum). Traité entre le pape et un gouvernement sur les affaires religieuses. | Dr. Accord entre le commerçant qui, ayant déposé son bilan, a été admis par le tribunal de commerce au règlement judiciaire et ses créanciers.

Les plus anciens concordats sont le concordat de Worms (1122), entre Calixte II et Henri V; le concordat de 1516, entre concurrence n. f. Rivalité entre plusieurs

concupiscence n. f. (du lat. concupiscere désirer). Penchant à jouir des biens ter restres, particulièrement des plaisirs sensuels. concupiscent (kökypisä l. e 1-at l adj. Qui exprime la concupiscence : regards concu-piscents. I Attaché aux plaisirs sensuels. concurremment [KOKYrama] adv. Par

concurrence. Par un concours mutuel, de concert : agir concurremment avec quel-

Léan X et François I. Le concordat entre personnes qui visent un même but : entrer

rendre épais). Rendre plus dense, réduire refroidissement ou compression condense la vapeur d'eau. d'une manière concise, en peu de

conclaviste n. m. Personne qui s'enferme au conclave avec un cardinal, pour le servir. concluant, e adj. Qui prouve bien ce qu'on

a avancé : argument concluent. conclure [k5klyr] v. t. (lat. concludere) [conj. 62]. Achever, terminer : conclure une affaire. Tirer une conséquence : conclure une chose d'une autre. | - V. i. Donner son avis, ses conclusions; se prononcer : on vous demande de conclure. | Etre probant

les témoignages concluent contre lui. conclusion n. f. (lat. conclusio). Assangement définitif : conclusion d'un traité. | Fin, résultat final : la conclusion d'un discours. Conséquence d'un argument : la conclusion d'un syllogisme ne doit pas dépasser les prémisses. | - Pl. Prétentions respectives de chacune des parties dans un procès. | Ecrit exposant ces prétentions. | Réquisition du ministère public. | - En conclusion loc. adv. En conséquence, pour conclure,

concecter v. t. Fam. Elaborer avec soin concocter une lettre de réclamation.

concombre [kākābr] n. m. (anc. provenç cocombre). Plante potagère de la famille des cucurbitacées, cultivée pour ses fruits allongés que l'on consomme comme légume ou en salade, | Ce fruit.

concomitamment adv. De façon conco-

concomitance [kokomitas] n. f. Coexistence, simultanéité de deux ou de plusieurs

concomitant, e adj. (lat. concomitant). Qui accompagne, qui se produit en même temps : des faits concomitants. . Variations concomitantes, variations simultanées et proportionnelles de certains phénomènes.

concordance n. f. Conformité, accord concordance de témoignages. | Géol. Disposition parallèle des couches sédimentaires. • Concordance de phases (Phys.), état de plusieurs vibrations sinusoïdales de même nature et de même période, dont la différence de phases est nulle. | Concordance des temps règles de syntaxe d'après lesquelles le temps du verbe d'une subordonnée varie seion celui du verbe de la principale

concours. | Lutte sportive : concours hippique. • Concours général, concours annuel entre les premiers élèves des classes supérieures des lycées, collèges et écoles normales, concret [kbkrz], ete [-zt] adj. (lat. concre-tus). Epais, condensé : huile concrète (vieilli). Qui exprime quelque chose de réel, de positif : obtenir des avantages concrets. Qui a le sens des réalités précises : esprit concret. | Gramm. Se dit d'un terme qui désigne un être ou un objet pouvant être perçu par les sens. . Musique concrète, technique de composition qui utilise les bruits produits par divers objets sonores enregistrés sur bande magnétique et susceptibles de transformation.

concret n. m. Qualité de ce qui est concret. concrétement adv. De façon concrète. concreter v. t. (cong. 3). Rendre concret

concrétion [kökresib] n. f. (de concret). Action de s'épaissir : la concrétion de l'huile. du sang. | Réunion de parties en un corps solide : concrétion saline. || Agrégation solide dans les tissus vivants : concrétions bi-Haires.

concrétiser v. t. Rendre concret ce qui est abstrait : concrétiser une idée, un orantage concubin, e adj. (lat. concubino). Relatif au concubinage. | — N. Personne qui vit en oncubinage

concubinage [kokybina3] n. m. Etat d'un homme et d'une femme qui vivent ensemble sans être mariés. (On dit aussi union LIBRE.)



condamnable adj. Qui mérite d'être condamné : acte condamnable. condamnation [kodanasjö] n. f. (lat. condemnatio). Décision d'un tribunal imposant à l'un des plaideurs de s'incliner au moins partiellement devant les prétentions de son adversaire. Il Décision d'une juridiction prononcant une peine contre l'auteur d'un crime, d'un délit ou d'une contravention. (En cour d'assises, le jury juge la culpabilité de l'accusé, et la cour prononce la condamnation.) | La peine infligée : une condamnation à la réclusion criminelle. | Blâme, désapprobation : la condamnation des abus condamnatoire adj. Qui porte condamna-

condamné, e n. Personne qui a subi une condamnation. - Adj. Qui ne peut échapper à un sort prévu : molade condamné. condamner [k3dane] v. t. (lat. condem-

nare). Prononcer un jugement contre un plaideur ou un inculpé : condamner un criminel Astreindre, réduire à : condamner au silence, à l'immobilité. | Désapprouver, blàmer : condamner une opinion, un usage. Interdire : la loi condamne la bigamie. Déclarer perdu, incurable : les médecins l'oni condamné. | Battet, muter : condamner une

condensable adj. Qui peut être condensé. réduit à un moindre volume.

condensateur n. m. Phys. Appareil ser vant à emmagasiner une charge électrique la bouteille de Leyde est un condensateur électrique. | Lentille servant à éclairer un objet dont on yeut former une image.

condensation p. I. Action de condenser ou effet qui en résulte. | Liquéfaction d'un gaz. Soudure de plusieurs molécules chimiques, avec élimination d'eau.

condensé n. m. Résumé d'une œuvre litté-

condenser [kodose] v. J. Gat. condensare. rendre épais). Rendre plus dense, réduire à un moindre volume. | Liquéfier un gaz par refroidissement ou compression : le froid condense la vapeur d'eau. | Fig. Exprimer

208 PAct. Lorenzee

### Cascading CRF models

a, A 1. Türk alfabesinin ilk sırasında yer alan ve A adı verilen bu harf, ses bilimi bakı- → aba güreşi mından kalın ünlülerin düz ve geniş olanını abacı a. 1. Aba yapan veya satan kimse. 2. gösterir. 2. müz. Nota işaretlerini harflerle gösterme yönteminde la sesini bildirir. a'dan z'ye (kadar) baştan aşağı, tamamen, tamamıyla, bütünüyle: Evini a'dan z'ye değistirdi.

a ünl. (a:) Şaşma, hatırlama, sevinme, acıma, abacılık, -ğı a. Abacının yaptığı iş. üzülme, kızma vb. duyguların anlatımına abadi a. (a:ba:di:) Far. ābādi esk. Açık sagüç kazandıran söz: A, ne güzel! A, sen burada mıvdın?

a / e ünl. Dilek kipinin ikinci teklik ve çokluk şahıslarının çekiminden sonra gelerek anlamı pekiştiren ve güçlendiren bir söz: "Azıcık dursana oğlum, dedi," -A. Kabaklı, "Onu bana getiriversene!" -N. Hikmet. "Başka gazetelere baksanıza! Onlar da yazıyor." -N. F. Kısakürek. "Şimdi de başka çıkmazdayız desenize!" -N. Uygur.

ab a. (a:b) Far. āb esk. Su.

→ abıhayat, abıkevser, abuhava

aba (I) a. hlk. 1. Abla. 2. Anne.

aba (II) a. Ar. 'abā 1. Yünün dövülmesiyle abajurculuk, -ğu a. Abajurcunun yaptığı iş. yapılan kalın ve kaba kumaş. 2. Bu kumaştan yapılmış yakasız ve uzun üstlük. 3. sf. Bu kumaştan yapılan. 4. esk. Bu kumaştan yapılan ve dervişlerce giyilen hırka. 5. Kepenek (I). aba altında er yatar "bir insanın değeri giyimiyle kuşamıyla ölçülemez" anlamında kullanılan bir söz. aba altından sopa (veya değnek) göstermek birini imalı bir biçimde tehdit étmek. aba gibi kaba ve kalın (kumaş). aba vakti yaba, yaba vakti aba "gereksinimler vaktinden önce ve ucuz olduğu zaman karşılanmalıdır" anlamında abalı sf. Aba giymiş olan. kullanılan bir söz. abanın kadri yağmurda Abana öz. a. (aba'na) Kastamonu iline bağlı bilinir "bir şeyin gerçek değeri ona gereksinim duvulduõunda anlasılır" anlamında kullanılan bir söz. (bir yere) abayı sermek 1) istenilmediği hâlde teklifsizce yerleşmek; uzun süre yerleşip kalmak. abayı yakmak tkz. birine aşırı bir biçimde gönül vermek, tutulmak, âşık olmak: "Sen mi verdin ona gönül yoksa o mu yaktı sana daha önce

abayı?" -O. C. Kaygılı.

Abadan giyecek yapan veya satan kimse. 3. sf. mec. Asalak. 4. hlk. Bedavacı. abacı kebeci (ara yerde) sen neci? hlk. "seni ilgilendirmeyen işe niçin karışıyorsun?" anlamında kullanılan bir söz.

man renginde, ipekten yapılan, yarı mat, kalınca bir yazı kâğıdı türü.

aba güreşi a. sp. Aba giyilerek ve bele kuşak bağlanarak yapılan bir güreş türü.

abajur a. Fr. abat-jour 1. Işığı bir yere toplamak, doğrudan doğruya gözlere vurmasını önlemek için kullanılan, kâğıt, kumaş, maden veya renkli camdan yapılmış lamba siperi. 2. Genellikle üzeri siperli masa lambası veya ayaklı lamba: "Etajerin sağında bir telefon, ortasında bir heykel, solunda bir abajur." -N. F. Kısakürek.

abajurcu a. Abajur yapan veya satan kimse.

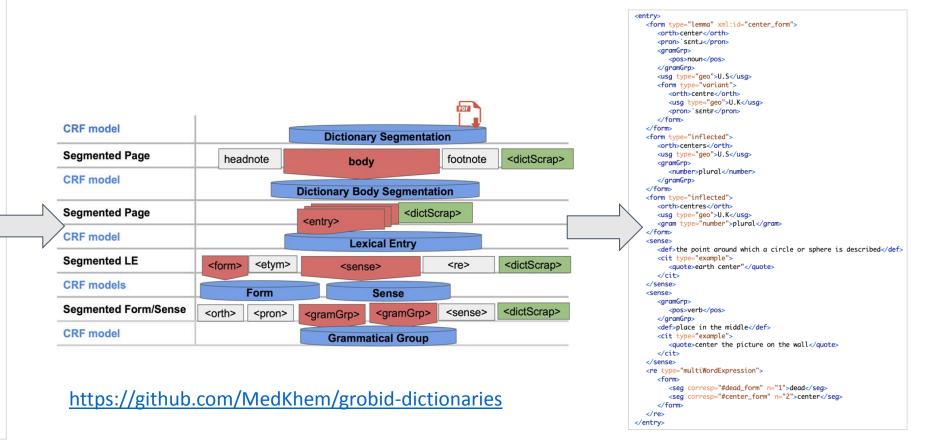
abajurlu sf. Abajuru olan: "Üstünde lacivert abajurlu, parlak bir madenden lamba." -P.

abajursuz sf. Abajuru olmayan: "Tavandaki abajursuz, çelik elektrik lambasını yakmış okuyordu." -S. F. Abasıyanık. abaküs a. Fr. abacus mat. 1. Savı boncuğu. 2.

mim. Sütun baslığının üstüne vatav olarak konan ve kenarlarından biraz dısarı tasan tas blok

abanabilme a. Abanabilmek işi.

abanabilmek (-e) 1. Abanma ihtimali veya imkânı bulunmak: "Sana, veni eserimde, takat getirilmez şekilde abanabilir miyim?" -N. F. Kısakürek. 2. Abanmaya gücü yet-



Mohamed Khemakhem, Luca Foppiano, Laurent Romary. Automatic Extraction of TEI Structures in Digitized Lexical Resources using Conditional Random Fields. electronic lexicography, eLex 2017, Sep

# Experiments and results - 1

amana

anuhu: a

anuhu: pr

anuhu naa:

anuhu ndal

atu: amarga

aya: amarga

about/a

a1 /e1/, A noun the first letter of the alphabet, followed by B O Do you mean 'dependant' spelt with an 'a' or 'dependent' with an 'e'? o from A to Z completely, all the way through

a2 /o, e1/, an /on, æn/ article 1. one o an enormous hole ○ a useful guidebook ○ She's bought a new car. O I want a cup of tea. O We had to wait an hour for the bus. (NOTE: an is used before words beginning with a, e, i, o, u and with h if the h is not pronounced: an apple or an hour. a is used before words beginning with all other letters and also before u where u is pronounced /ju:/: a useful guidebook) 2. for each or to each ○ Apples cost £1.50 a kilo. ○ The car was travelling at 50 kilometres an hour. ○ He earns £100 a day.

abandon /ə'bændən/ verb 1. to leave someone or something in an unkind way O The dog had been abandoned by its owner. 2. to give up or stop doing something o The company has decided to abandon the project. O We abandoned the idea of setting up a London office.

abbreviation /əˌbri:vi'eɪʃ(ə)n/ noun a short form of a word

ability /ə'bılıti/ noun 1. a natural tendency to do something well ○ I admire his ability to stay calm in difficult situations. O We can develop their natural abilities. (NOTE: The plural is abilities.) ☐ I'll do it to the best of my ability I'll Til do it to the best of my adzi cuvui: o (disyuntiwa); por ventura; quizás do it as well as I can 2. the fact of being adzi cuvui: o (disyuntiwa); por ventura; quizás clever O suitable for different levels ability (NOTE: no plural)

able /'e1b(a)l/ adjective 1. ato be able do something to be capable of som thing or have the chance to do son thing or have the chance to find thing o They weren't able to find thing o They weren't be to come to thing O They weren ay house. O Will you be able to come to ama: así meeting? 2. good at doing something good at doing many things O She

a a: a (del que halla a otro en maleficio)

a dzuchica caa cuvui: poco más o menos

thing a dzuchica añandaa: poco más o menos

o s a dzuchica coo cuvui: poco más o menos

a huihica añandaa: poco más o menos

a na ndehe cuvui ndatu nicay: o bienaventurado,

a yoo ee ñahando: por ventura alguno de voso-

aa: ya, acordándoseme lo que se me había olvi-

aa ndica huii: ya (acordándoseme lo que se me

adzi yoo ee ñahando: por ventura alguno de vo-

aa dzuhua huii: así, así (sonriéndose)

adzi: o (disyuntiva); por ventura; quizás

a hua dzevui dzavua: pues no

do so a a a: a (interiectio admirantis)

hov a hua dzevui: o no

ye a huitnani: ahora poco

a ñaha: o no; pues no

a sa dzevui: pues no

aa: de manera que

adzi q cuvui: por ventura

adzi yoo: por ventura alguno

a yoo: por ventura alguno

### DICTIONNAIRE

### LANGUE FRANÇAISE.

A Rarde.

A Ray A Ray (m) n.4, pl. mebaka abaka akan (m) n.4, pl. mebaka du nkan (m) n.4, pl. mebaka du nkan (m) n.4, pl. mebaka (m) n.4, pl. meba

PREMIÈRE FANG.FRA

on dit abañ e

de porter (vb bab

ABABGA (b) n.

ABACHA (h) n.4,

ABAGHE

ABAGHE

(bb baghé b). Abatta

Abaghé chen, coup de

le travail d'abattage des al

l'abattage des al

ABAGHI

ABAGH

ABAGHLE (bm) n.4, pl. (vb baghle b). Action de garden abaghle dam, tu as v

a mon, père de enfant garçon (on Aler-a-mor,

chemin glis.

(bmh), nsilga nyôl (bh). ABAISSER baisser, faire descendre, silé (b). S'abaisser, zoghbe (b), zombe

donner quelqu'un, ndeghé môr (bb).

~ ABATTAGE des arbres, abaghé bili (bmh). Abattage du sous-bois à la matchette pour faire une plantation, meli (h), meli me tsi (hh). Abattage du sousbois recommencé, parce qu'on n'avait pas abattu les arbres à temps, ékwalé (h). Abattage des fruits en tapant avec un bâton, mfarga (m). Abattage collectif

ABATTRE les fruits, far (h).

narique), ékyel (b), zor (m). Gros ab-cès, anthrax, phlegmon, bome (b). Ab-cès à l'aîne, adénite, ntsima (h), fikar (h), koko (m), ékokol (m). Abcès de l'aisselle, lañ-bone (hh), lè (h). Abcès au doigt, panaris, ôbîr (h). Abcès au sein d'une femme qui allaite, mbembè

Petite abeille de forêt, abê (m); son miel, abê (m). Abeille encore plus petite, fait, éfôn (h). Ruche abôn e fôfôn (mb).

DEUXIÈME PARTIE FRANÇAIS - FANG

A

ABAISSEMENT asilé (bm), nsilga (b). Humiliation personnelle, asilé nyôl

(b). Baisser la tête, zoghé nlô (bh).

ABANDONNER délaisser, zamé (h) bore (b), sie (b). Abandonner les charges pour se sauver, sô byôm (bh). Aban-

ABANGA affluent de l'Ogooué, Aba-

de petits arbres, ñkume (h).

ABCES tumeur, éküküa (h) (nom gé-

ABEILLE nom générique, abè (m). mvem (b); son miel, ôbwé (h). Abeille de grandeur normale, fôfôñ (b) et bôe (b); son miel, yvi (h). Abeille plus grande que mvem, plus petite que fôfôn, memvame (h). Cire d'abeille, éko (m). Butinement des abeilles, éleélé (h). Essaim d'abeilles en place sur un tronc d'arbre, duma fôfôñ (bmb). Essaim d'abeilles en marche et le bruit qu'il

Butin des abeilles (miel), ngôr (h). Aiguillon d'abeille, nson ô fôfôn (bb).

ABÎME géhenne, tôtôla (b)

ABÎMER être abîmé, ndama (h), bîra (h). Très abîmé, ne-meghlé (b), ne-mekh b), nnžugha (bm).

ABOIEMENT du chien, abômé mvu (hh), mbômga ô mvu (hh). Aboiement de l'antilope ñkokh, mbômga ñkokh (hh). ABOKWE affluent du Como, Aboghe

ABONDANCE afokh (b) (vivres, richesses, hommes). Etre dans l'abondance,

ABOYER bôm (h). Le chien aboie,

mvu za bôm. ABRÉGER raccourcir, seghé (b).

ABRI hutte, cabane, abana (b), ébem (m), mba (h).

ABSENTER (s'), être parti, saña (b). ABSTENIR (s') de choses diverses, kî (b) (voir ce mot). Celui qui s'abstient, nki (b). S'abstenir de rapports sexuels entre époux, kia (b).

ACAJOU rouge du Gabon, mbegha (h), nžamengila (mb), ôven mevu (bb),

ACCABLER les gens de travail, tughle bôr (hb). Celui qu'on accable de travail, étugha (h).

ACCAPARER drainer, prendre tout,

ACCEPTATION ayebe (h).

ACCEPTER dire oui, yebe (h), bame

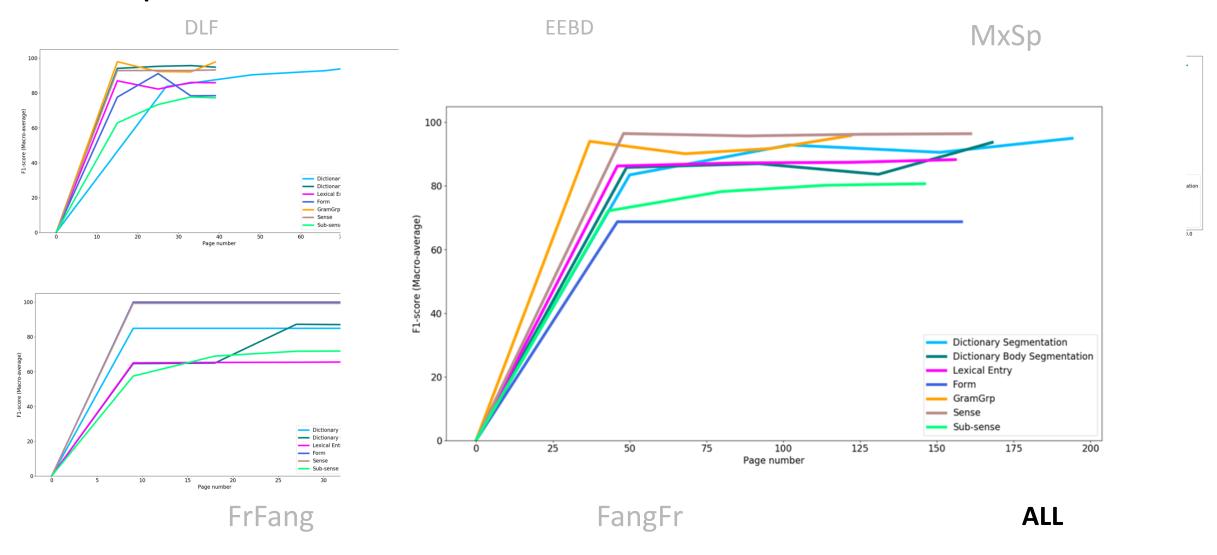
ACCOMPAGNER lera (h), lere môr (hb). Action d'accompagner ou de porter un paquet un petit bout de chemin,

ACCOMPLIR une promesse, tsvie magha (hm). La promesse s'accomplit, magha ma tsvie.

Easier English Basic Dictionary (2009)

Dictionnaire de la langue Française (1873)

# Experiments and results - 2



# Further steps in language resources management

### Training

- Basic digital literacy
  - Even for very simple issues related to online presence (authorship, affiliation)
- Concrete knowledge of existing standards for language resources
  - Preventing the development of (any) new formats, unless strongly supported by evidence

### Active contributions

- Participation to standardisation activities
- Building up standards-based infrastructures

### General policies

- A strong relation to open science principles (the I in FAIR, which leads to the R)
- Hence a necessary stance for any research project

# Merci pour votre attention