IA pour l'analyse des documents médicaux : classification, simplification, extraction d'information

Natalia Grabar

CNRS, Univ. Lille, UMR 8163 - STL - Savoirs Textes Langage, F-59000 Lille, France
natalia.grabar@univ-lille.fr

15/05/2023, École d'été ESIA 2023, Berck-sur-Mer

Plan

- Introduction
- Approches et méthodes
- Thématiques
- Conclusion

Contexte

- Domaine médical :
  - un grand producteur de données
- Types de données :
  - structurées : formulaires, ressources terminologiques, tableaux...
  - non structurées : texte
    - résumés de sortie, comptes-rendus de consultation, d'imagerie ou d'hospitalisation, comptes-rendus d'imagerie, etc.
  - orales (dictée vocale, appels SAMU/hôpitaux, enregistrements de patients...)
  - imagerie
  - produits chimiques, pharmacie
  - séquences de gènes
  - ...
- Dimensions de données :
  - médicales : processus de soins, contexte clinique
  - de santé : pour le patient

Informatisation des hôpitaux

- Systématisation lors de la création et collecte
- Stockage
  - entrepôts de données : I2B2, Agfa Healthcare, Dedalus France...
- Unité : patient, visite
- Accumulation des données
- Enrichissement
  - sources de données
  - ressources terminologiques (normalisation, interopérabilité...)
  - données liées
  - ...
- Possibilité de leur exploitation pour la recherche
  - aspects légaux
  - aspects éthiques
  - ouverture vers le contexte interdisciplinaire

Centralité du patient

- HPST : hôpital, patient, santé, territoire
  - loi de 2009
- Prise de décisions
- Patient : propriétaire des données
- Consultation, compréhension

Besoins

- Professionnels de santé :
  - accéder rapidement à une information donnée
  - différents niveaux (document, paragraphe, phrase...)
  - besoins principaux dans le traitement de l'information :
    - trouver
    - extraire
    - catégoriser
    - structurer
    - lier...
- Patients :
  - trouver l'information correcte
  - comprendre l'information

    [Williams *et al.*, 1995, Patel *et al.*, 2002, Jucks & Bromme, 2007]
  - besoins principaux dans le traitement de l'information :
    - rechercher/trouver
    - comprendre

Objectif

- Présenter le TAL (Traitement Automatique des Langues)
  - approches
  - données
  - évaluation
- Proposer une vue de travaux de recherche autour des données médicales
  - exemples de tâches

TAL

les grands débuts

- Les années 1950 (guerre froide)
- Traduction automatique : automatisation de la traduction d'une langue vers une autre
- 1954 : Georgetown-IBM experiment
- Environ $20 millions investis en 10 ans
- Test :
  - *The spirit is willing, but the flesh is weak*
  - $\Rightarrow$ Russe $\Rightarrow$ Anglais
  - *The whisky is strong, but the meat is rotten*

TAL

dessous linguistiques

- Dictionnaire électronique
- Substitution de mots équivalents dans la langue cible transfert lexical
- Ordre syntaxique des mots
- Problématiques :
  - Ambiguïtés, polysémies, ...
  - Structures syntaxiques complexes
  - Relations sémantiques
  - Anaphores, ...

TAL

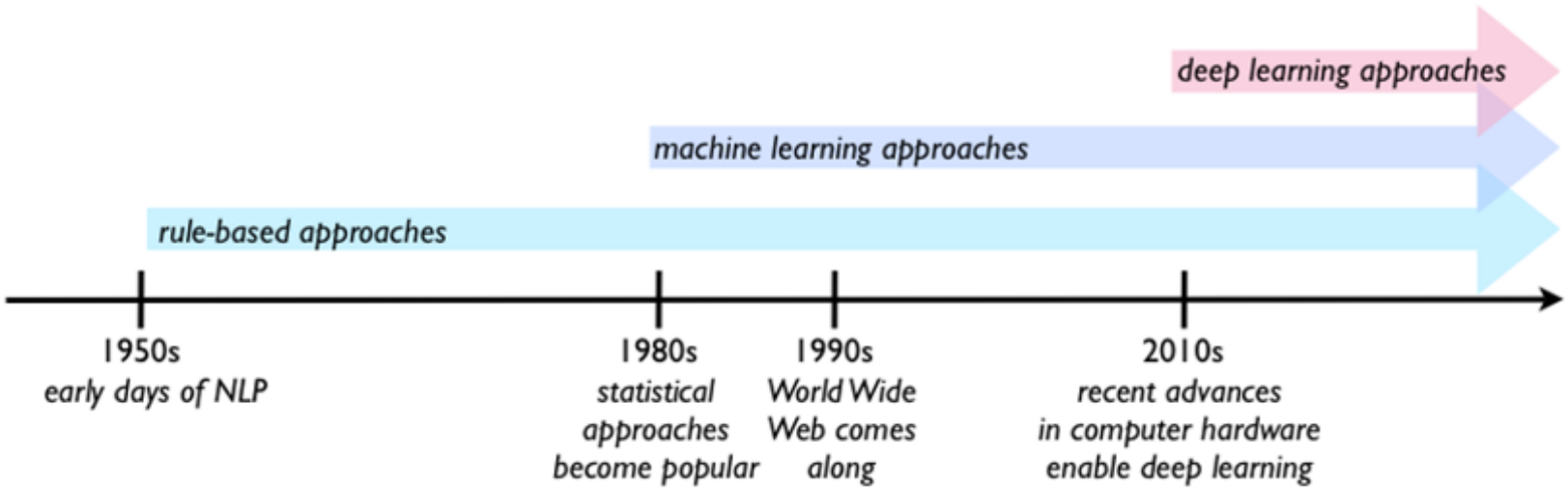
the "ALPAC report"

- En 1966, par the US National Academy of the Sciences
  Y. Bar-Hillel
  - "MT is hopeless"
  - La bonne qualité ou l'automatisation complète ne peuvent pas être atteintes
  - L'automatisation complète n'est pas souhaitable
    coûts éventuellement plus élevés qu'avec les traducteurs humains
  - Recommandation :
    - mettre plus d'effort dans la recherche en linguistique
    - qu'elle contribue ou non à la traduction automatique directement

$\Longrightarrow$ Début des travaux en TAL

TAL

types d'approches



- deux principales approches en TAL :
  - basées sur des règles
  - basées sur l'apprentissage automatique

TAL

approches basées sur des règles

- Approche intuitive
  - mots-clés, structures syntaxiques, terminologie, liens...
- Experts humains
- Création de règles
- Approche gérable manuellement
  - imbrication de règles
- Résultats interprétables :
  - précision élevée
  - rappel faible
- Avantages :
  - nouveau (domaine, question de recherche, langue...)
  - assez facile à mettre en place
  - utilisable par les non-experts
  - un bon niveau d'explicabilité

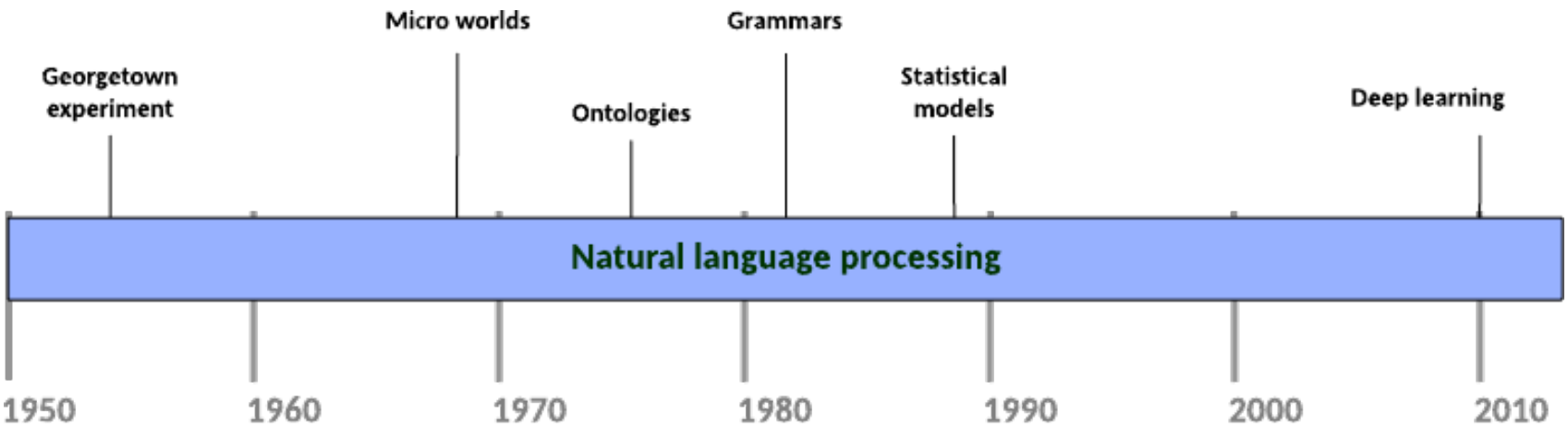
TAL

approches basées sur l'apprentissage automatique

- Création d'un corpus d'entraînement
  - annoté avec les catégories d'intérêt
  - de taille importante
- Annotation par des experts (au moins 2)
  - accord inter-annotateur
- Algorithmes d'apprentissage automatique
  - création de modèles
  - généralisation
  - transposition sur de nouvelles données
- Descripteurs :
  - internes, externes, contexte...
  - connaissances ajoutées

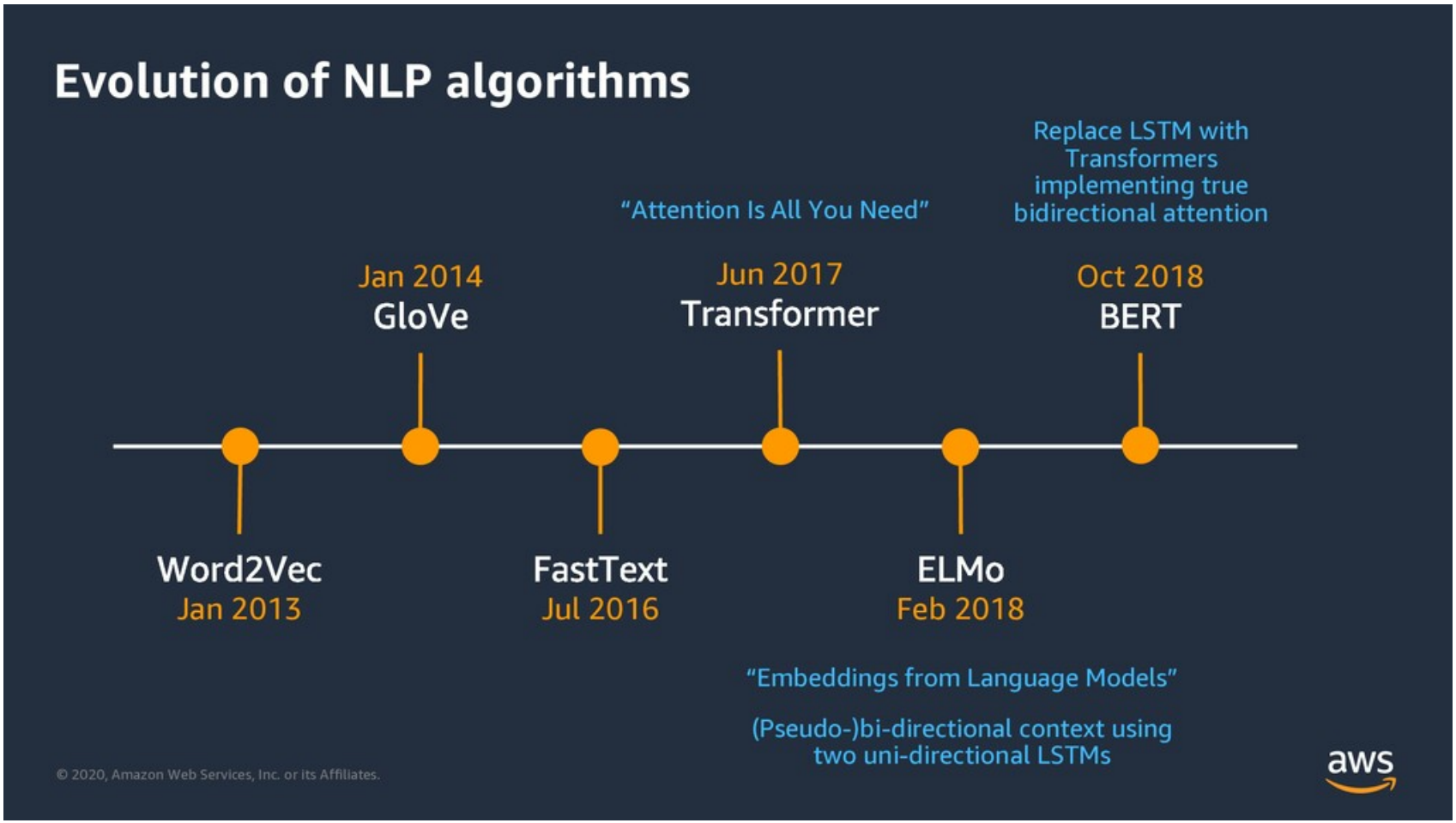
TAL

évolution des approches



TAL

évolution récente



TAL

évolution récente

BERTisation :

- fine-tuning de modèles existants

  [Flamholz *et al.*, 2022, Noh & Kavuluru, 2021, Bear Don't Walk Iv *et al.*, 2021, Michalopoulos *et al.*, 2021, Liu *et al.*, 2021a, Amir *et al.*, 2021]
- adaptation au domaine [Wang *et al.*, 2021, Liao *et al.*, 2021]
- transfer learning

  [Legrand *et al.*, 2021, Hussain *et al.*, 2021, Bear Don't Walk Iv *et al.*, 2021]
- self-training [Liao *et al.*, 2021]
- enrichissement avec des informations :
  - orthographe et lexique [Ding *et al.*, 2021]
  - classes syntaxico-sémantiques [Majewska *et al.*, 2021]
  - connaissances médicales [Roy & Pan, 2021]
- subword embeddings [Lauriola *et al.*, 2021, Kim *et al.*, 2021b]
- vector retrofitting [Ding *et al.*, 2021, Majewska *et al.*, 2021]

TAL

contribution des disciplines

Un domaine interdisciplinaire :

- linguistique
  - paramètres phonologiques
  - grammaire générative
  - syntaxe structurale
  - philosophie du langage
- informatique
  - algorithmique
  - génie logiciel
- mathématiques
  - logique
  - théorie des langages
  - probabilités

TAL

données : ressources

Langues de spécialité : spécificités à différents niveaux

- lexiques (ensemble de mots)
  - *veine, oeil, côte, vertèbre...*
- terminologies (termes, relations, définitions)
  - *veine porte*, *veine sous-clavière*
  - hiérarchique, partie de, synonymie, anonymie
  - localisé dans, traité par, traite...
- ontologies (+ contrôles)
- graphes de connaissances (triplets)
- syntaxe

TAL

données : terminologies

- *MeSH* (Medical Subject Headings) [NLM, 2001]
  - indexation et recherche d'information
- *CIM* (Classification Internationale des Maladies) [OMS, 1995]
  - encodage des (co)morbidités
- *CCAM* (Classification Commune des Actes Médicaux) [Trombert-Paviot *et al.*, 2003]
  - encodage des actes médicaux
- *MedDRA* (Medical Dictionary for Drug Regulatory Activities) [Brown *et al.*, 1999]
  - effets indésirables de médicaments
- *SNOMED International* [Côté *et al.*, 1993]
  - description des données médicales
- *SNOMED CT* [Wang *et al.*, 2002]
  - évolution vers une ontologie
- UMLS (Unified Medical Language System) [Lindberg *et al.*, 1993]
- ...

TAL

données : corpus

Matériel indispensable en TAL

- doit être proche des données traitées
- clinique, scientifique, réseaux sociaux...
- *MIMIC* (Medical Information Mart for Intensive Care) [Johnson *et al.*, 2016]
  - largement exploité
  - prédiction de ma mortalité [Anand *et al.*, 2018, Feng *et al.*, 2018] , identification de diagnostics et encodage [Perotte *et al.*, 2014, Li *et al.*, 2018] , étude de la temporalité [Che *et al.*, 2018] , identification de documents cliniques similaires [Gabriel *et al.*, 2018] ...
- sous-parties exploitées pour les compétitions de TAL

TAL

données : compétitions de TAL

- *i2b2* (Informatics for Integrating Biology and the Bedside) `https://www.i2b2.org/NLP/DataSets/Main.php`
  - financé par le NIH
  - promouvoir le développement et le test des outils de TAL pour l'anglais
  - améliorer les outils pour le traitement fin d'information cliniques
  - enrichi avec des annotations specifiques [Uzuner, 2008, Uzuner *et al.*, 2011, Sun *et al.*, 2013] :
    - dé-identification, statut de fumeur, informations sur les médicaments, relations sémantiques entre entités, temporalité
  - corpus et annotations : disponibles pour la recherche

TAL

données : compétitions de TAL

- *n2c2* (National NLP Clinical Challenges)
  `https://n2c2.dbmi.hms.harvard.edu/`
  - depuis 2018
  - tâches typiques du contexte clinique
  - inclusion de patients dans les essais cliniques
  - détection d'effets indésirables de médicaments
  - similarité textuelle
  - normalisation de concepts
  - extraction de l'histoire familiale

TAL

données : compétitions de TAL

- *CLEF-eHEALTH* challenges
  `https://sites.google.com/site/shareclefehealth/`
  - 2013 et 2014 : détection de maladies, normalisation d'abréviations
  - 2016 : structuration des notes d'infirmiers d'Australie
  - 2016 et 2017 : certificats de décès français de CépiDc (`http://www.cepidc.inserm.fr/`), extraction des causes de décès
- *eHealth-KD* 2019 challenge
  `https://knowledge-learning.github.io/ehealthkd-2019`
  - espagnol
  - identification et classification de mots-clés
  - détection de relations sémantiques entre les mots-clés

TAL

données : compétitions de TAL

- corpus CAS [Grabar *et al.*, 2020]
  - cas cliniques publiés
  - différents domaines cliniques
  - partiellement annoté
- compétitions DEFT 2019, 2020, 2021
  `https://deft.lisn.upsaclay.fr/`
  - âge, genre du patient
  - issu du traitement
  - maladies, médicaments et informations liées, procédures
  - signes et symptômes
  - ...

TAL

mesures d'évaluation

Dépendance des tâches :

- recherche et extraction d'information
  - précision : est-ce que ce qu'on trouve est correct ?
  - rappel : est-ce que ce qu'on trouve est complet ?
  - F-mesure : moyenne harmonique de P et R
- traduction automatique
  - BLEU (*bilingual evaluation understudy*) [Papineni *et al.*, 2002]
  - ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*) [Lin, 2004]
- simplification
  - SARI [Xu *et al.*, 2016]
  - index de lisibilité [Flesch, 1948, McLaughlin, 1969, Gunning, 1973, Kincaid *et al.*, 1975, Björnsson & Härd af Segerstad, 1979]
- génération de textes, résumé automatique...
- ...

TAL

exemples de tâches/applications

	Phonetique	Morphologie	Syntaxe	Semantique	Pragmatique
Ressources	prononciation syllabation prosodie	flexion derivation composition	lexiques syntaxiques	reseaux semantiques lexiques semantiques terminologies	regles desambiguisation
	lexique.org, ...	MorTAL, Celex, ...	LTAG, FTAG, LFG, ...	WordNet, DEC, ...	
Taches	Reconnaissance vocale Generation vocale (text speech)	Segmentation morphologique Analyse morphologique	Etiquetage morpho−syntaxique Analyse syntaxique Chunking	Extraction des unites de sens simples, complexes Detection de relations Decomposition en primitives Recherche de definitions	Structure de textes Anaphore Communication
Applications	Linguistique des corpus Generation de ressources TA (Traduction automatique) TAO	Reconnaissance de la parole	RI (Recherche d'information) EI (Extraction d'information) QR (Question/Reponses) Stylistique	Statistical NLP Dialogue homme−machine Correction orthographique	terminologies ontologies Generation sens−texte Resume automatique Generation automatique bulletins meteo, comptes−rendus, ...

Ressources terminologiques

Base pour plusieurs thématiques et applications

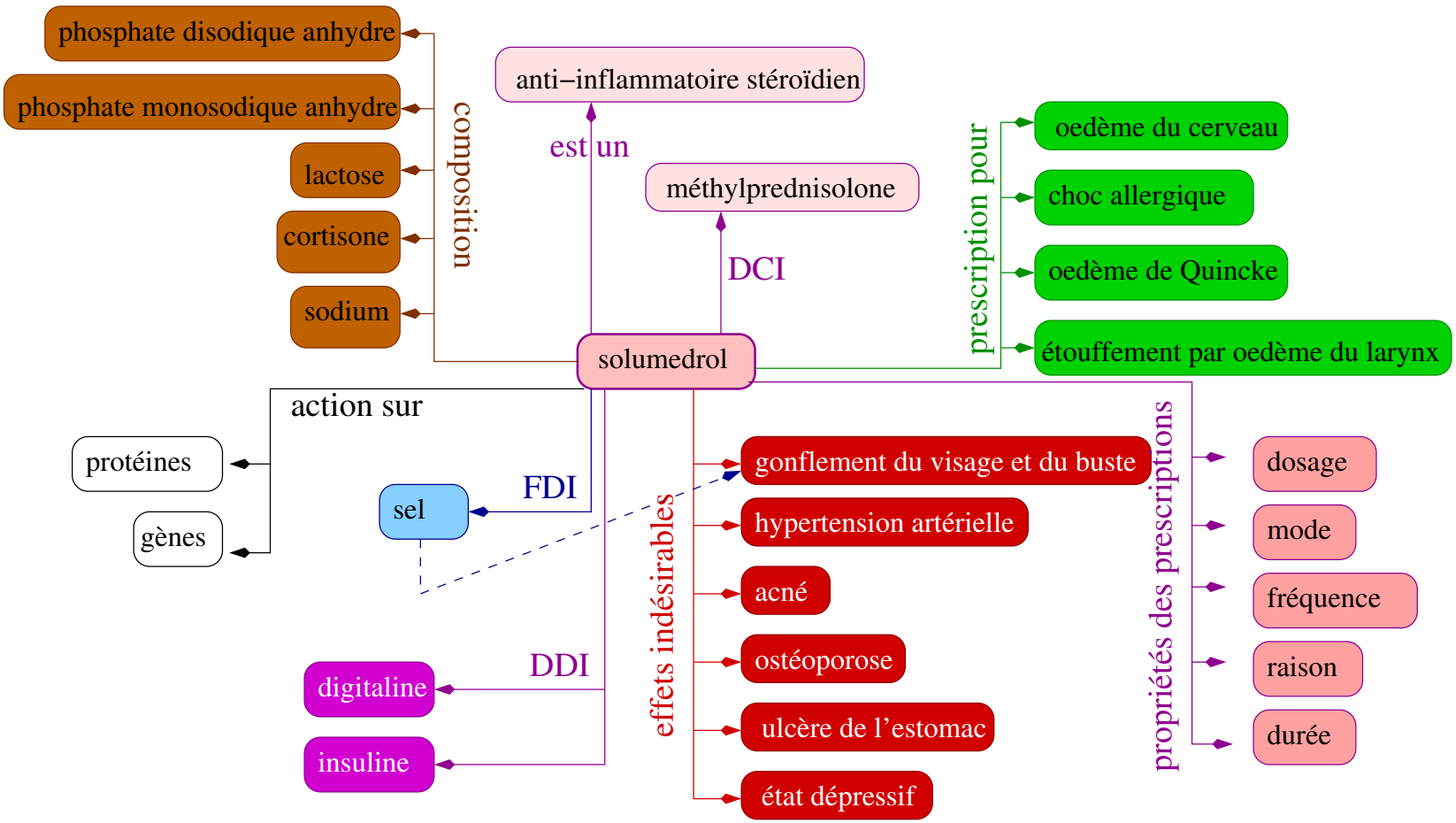
- extraction de termes
- structuration de termes
- fusion de ressources terminologiques [Lindberg *et al.*, 1993]
- mise à jour, adaptation, enrichissement...
- données liées : apprentissage supervisé distant
  [Riedel *et al.*, 2010, Surdeanu *et al.*, 2012, Mintz *et al.*, 2009]
- transfert
  - domaines [Blitzer *et al.*, 2006, Jiang & Zhai, 2007, Plank & Moschitti, 2013, Legrand *et al.*, 2021]
  - langues [Hamon & Grabar, 2016]

Applications :

- codage PMSI de l'activité des hôpitaux
  [Ruch *et al.*, 2008, Zhou *et al.*, 2021, Liu *et al.*, 2021c, Dong *et al.*, 2021]
- interopérabilité sémantique [Degoulet *et al.*, 1997, Walker *et al.*, 2005, DGME, 2009, Stroetmann *et al.*, 2009, Wajsbürt *et al.*, 2021]

Extraction d'information

pharmacovigilance



Extraction d'information

recrutement de patients pour les essais cliniques

- Tâche nécessaire mais très lourde

  [Campillo-Gimenez *et al.*, 2015, Fletcher *et al.*, 2012]
- Retards dans les recrutements [Center Watch, 2013]
- Systèmes automatiques

  [Cuggia *et al.*, 2011, Embi *et al.*, 2005, Olasov & Sim, 2006, Ross *et al.*, 2010, Tu *et al.*, 2011, Pressler *et al.*, 2012, Shivade *et al.*, 2014]
- Recherche clinique
  - compétition N2C2 [Oleynik *et al.*, 2019, Percha *et al.*, 2021, Du *et al.*, 2021, Liu *et al.*, 2021b]
- Différents types de critères (inclusion, exclusion) selon les objectifs des essais cliniques

Extraction d'information

revues systématiques d'articles

- Faciliter l'accès aux derniers résultats de la recherche
  - pour les professionnels de santé
- Différents types de revues d'articles
  - revues systématiques
  - revues narratives
  - revues rapides
  - revues de la portée (*scoping review*)
  - pronostique, diagnostique...
- Collaboration Cochrane [Collaboration, 2009]
- Travaux intenses en TAL [Kreimeyer *et al.*, 2017]
- Orientation récente sur les patients
  - *Cochrane plain language summary*

Maladies neurodégénératives

Détection de patients atteints (AZH) :

- données orales (entretiens, interviews)
- transcription
- analyse du discours :
  - interaction verbale [Boyé *et al.*, 2014]
  - répétitions, disfluences [Duong *et al.*, 2003, Berrewaerts *et al.*, 2003, Ska & Duong, 2005, Lee, 2011, Boyé *et al.*, 2014]
  - syntaxe [Kynette & Kemper, 1986, Kemper *et al.*, 1990, Nef & Hupet, 1992]
  - lexique [Kynette & Kemper, 1986, Boyé *et al.*, 2014]
  - sémantique [Croisile *et al.*, 1996, Rousseau *et al.*, 2009]
  - pragmatique [Ska & Duong, 2005, Lee, 2011, Gaspers *et al.*, 2012]

$\Rightarrow$ détection à des stades précoces de la maladie

Santé mentale

- Maladies neurologiques et psychiatriques
  - COVID-19 et phases de confinement
- Thématiques récentes :
  - amélioration du diagnostic [Shiner *et al.*, 2021]
  - évolution temporelle des phases de la maladies [Viani *et al.*, 2021]
  - idées suicidaires, auto-mutilation [Rozova *et al.*, 2022, Cliffe *et al.*, 2021]
  - mésusage de substances [Cameron *et al.*, 2013, Kalyanam *et al.*, 2017, Bigeard *et al.*, 2019, Cox *et al.*, 2021]
  - assistance en ligne [Leung *et al.*, 2021, Hassan *et al.*, 2021]
- Provenance des données :
  - réseaux sociaux
  - dossiers cliniques

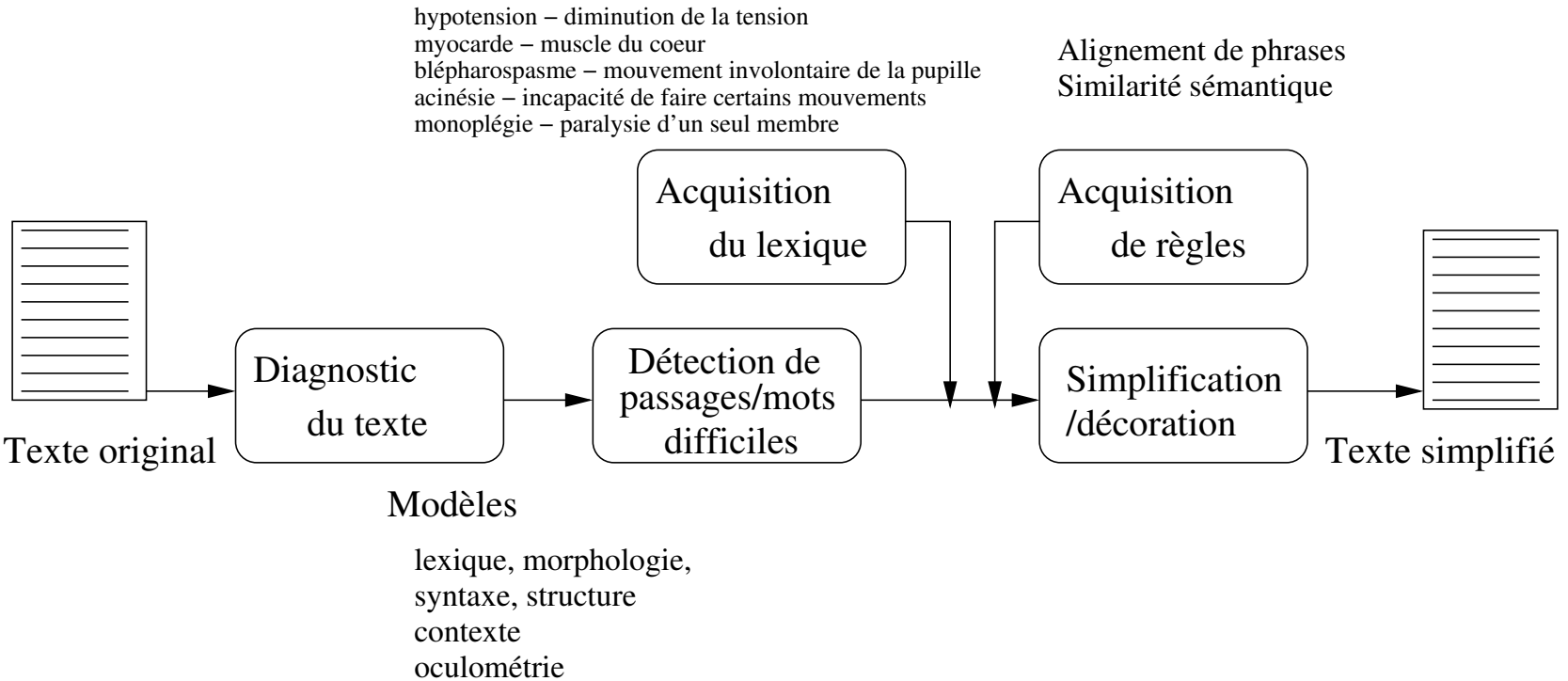
Exploration des réseaux sociaux

- Pourquoi utiliser les données des réseaux sociaux ?
- Thématiques :
  - santé mentale, dépression [Rozova *et al.*, 2022, Cliffe *et al.*, 2021]
  - mésusage de médicaments [Cameron *et al.*, 2013, Kalyanam *et al.*, 2017, Bigeard *et al.*, 2019, Cox *et al.*, 2021]
  - jugement sur les traitements, les vaccins, les procédures [Osadchiy *et al.*, 2020, Aljedaani *et al.*, 2022]
  - nouvelles pratiques
  - émotions et vécu des malades
  - effets indésirables de médicaments (Mediator)

Caractériser l'urgence

- Appels du SAMU :
  - des milliers d'heures d'enregistrements
  - signal audio
- Aider à réguler les appels :
  - gravité et urgence [Kim *et al.*, 2021a]
- Indicateurs :
  - lexique
  - intonation, prosodie
  - qualité de la voix
  - ...
- Difficultés :
  - différencier les rôles
  - transcription
  - bruitage, accents...
  - annotation

Simplification



- *une hypotension artérielle peut être observée en cas d'administration intraveineuse trop rapide, inférieure à 60 minutes (voir rubrique 4.2)*
- *une diminution de la tension artérielle peut être observée en cas d'administration intraveineuse trop rapide, inférieure à 60 minutes*

[Koptient & Grabar, 2020, Cardon & Grabar, 2020a, Cardon & Grabar, 2020b]

Difficultés

- Disponibilité et accessibilité des données :
  - accès réglémenté et difficile
  - corpus de cas cliniques [Grabar *et al.*, 2020]
- Corpus et données annotés :
  - documents cliniques
  - fiabilité des annotations
  - expertise des annotateurs
  - accord, consensus

Difficultés

- Données rares :
  - maladies rares (1/2 000, 1/10 000, 1/100 000...)
- Données incomplètes :
  - histoire de la maladie, analyses de laboratoire, prescriptions...
- Données incertaines [Simianu *et al.*, 2016, Reiner, 2018] :
  - incertitude scientifique
  - incertitude diagnostique
- Données distribuées :
  - entre différentes sources ou différents documents
- Fiabilité :
  - sources
  - modèles, résultats
- Explicabilité :
  - pour la prise de décisions

Considérations éthiques

- Disponibilité et accessibilité des données
- Informations identifiantes
- Fiabilité des modèles et résultats
- Explicabilité des résultats
- Stabilité des résultats avec différentes approches

Conclusion

- Données de différents types
- Besoin de collaborations pour les traiter et valoriser
- Besoin de gros volumes de données
  - approches modernes


Aljedaani, W., Abuhaimed, I., Rustam, F., Mkaouer, M., Ouni, A. & Jenhani, I. (2022).
Automatically detecting and understanding the perception of COVID-19 vaccination : a middle east case study.
Soc Netw Anal Min, **12**(1), 1–12.


Amir, S., van de Meent, J.-W. & Wallace, B. C. (2021).
On the impact of random seeds on the fairness of clinical classifiers.
In *Proc of the 2s021 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies*, pp. 3808–3823 : Association for Computational Linguistics.


Anand, R., Stey, P., Jain, S., Biron, D., Bhatt, H., Monteiro, K., Feller, E., ML, R., IN, S. & ES, C. (2018).
Predicting mortality in diabetic ICU patients using machine learning and severity indices.
In *AMIA Jt Summits Transl Sci Proc*, pp. 310–319.


Bear Don't Walk Iv, O. J., Sun, T., Perotte, A. & Elhadad, N. (2021).
Clinically relevant pretraining is all you need.
J Am Med Inform Assoc, **28**(9), 1970–1976.


Berrewaerts, J., Hupet, M. & Feyereisen, P. (2003).
Langage et démence : examen des capacités pragmatiques dans la maladie d'Alzheimer.
Revue de Neuropsychologie, **13**(2), 165–207.


Bigeard, E., Thiessard, F. & Grabar, N. (2019).
Detecting drug non-compliance in internet fora using information retrieval and machine learning approaches.
In *MEDINFO 2019*, pp. 1–6.


Björnsson, H. & Härd af Segerstad, B. (1979).
Lix pä franska och tio andra spräk.
Stockholm : Pedagogiskt centrum, Stockholms skolförvaltning.


Blitzer, J., McDonald, R. & Pereira, F. (2006).
Domain adaptation with structural correspondence learning.

In *Proc of the 2006 Conf on Empirical Methods in Natural Language Processing*, pp. 120–128 : Association for Computational Linguistics.

Boyé, M., Tran, T. & Grabar, N. (2014).

Nlp-oriented contrastive study of linguistic productions of alzheimer and control people.
In *POLTAL*, pp. 412–424 : Springer, Advances in Natural Language Processing, LNCS 8686.

Brown, E., Wood, L. & Wood, S. (1999).

The medical dictionary for regulatory activities (MedDRA).
Drug Saf., **20**(2), 109–117.

Cameron, D., Smith, G. A., Daniulaityte, R., Sheth, A. P., Dave, D., Chen, L., Anand, G., Carlson, R., Watkins, K. Z. & Falck, R. (2013).

PREDOSE : A semantic web platform for drug abuse epidemiology using social media.
Journal of Biomedical Informatics, **46**, 985–997.

Campillo-Gimenez, B., Buscail, C., Zekri, O., Laguerre, B., Le Prisé, E., De Crevoisier, R. & Cuggia, M. (2015).

Improving the pre-screening of eligible patients in order to increase enrollment in cancer clinical trials.
Trials, **16**(1), 1–15.

Cardon, R. & Grabar, N. (2020a).

Construction d'un corpus parallèle à partir de corpus comparables pour la simplification de textes médicaux en français.
volume 61, pp. 15–39.

Cardon, R. & Grabar, N. (2020b).

French biomedical text simplification : When small and precise helps.
In *COLING 2020*, pp. 1–8.

Center Watch (2013).

State of the Clinical Trials Industry : A Sourcebook of Charts and Statistics.
Technical report, Center Watch.

Che, Z., Purushotham, S., Cho, K., Sontag, D. & Liu, Y. (2018).

Recurrent neural networks for multivariate time series with missing values.
Sci Rep, **8**(1), 6085.

Cliffe, C., Seyedsalehi, A., Vardavoulia, K., Bittar, A., Velupillai, S., Shetty, H., Schmidt, U. & Dutta, R. (2021).
Using natural language processing to extract self-harm and suicidality data from a clinical sample of patients with eating disorders : a retrospective cohort study.
BMJ Open, **11**(12), e053808.

Collaboration, C. (2009).
Cochrane : systematic review of biomedical literature.
Cochrane Collaboration.
www.cochrane.org.

Côté, R. A., Rothwell, D. J., Palotay, J. L., Beckett, R. S. & Brochu, L. (1993).
The Systematised Nomenclature of Human and Veterinary Medicine : SNOMED International.
Northfield : College of American Pathologists.

Cox, D. J., Garcia-Romeu, A. & Johnson, M. W. (2021).
Predicting changes in substance use following psychedelic experiences : natural language processing of psychedelic session narratives.
Am J Drug Alcohol Abuse, **47**(4), 444–454.

Croisile, B., Ska, B., Brabant, M., Duchene, A., Lepage, Y., Aimard, G. & Trillet, M. (1996).
Comparative study of oral and written picture description in patients with Alzheimer's disease.
Brain and language, **53**, 1–19.

Cuggia, M., Besana, P. & Glasspool, D. (2011).
Comparing semi-automatic systems for recruitment of patients to clinical trials.
Int J of Medical Informatics, **80**(6), 371–88.

Degoulet, P., Fieschi, M. & Attali, C. (1997).
Les enjeux de l'interopérabilité sémantique dans les systèmes d'information de santé.
Informatique et Santé, **9**, 203–12.

DGME (2009).
Référentiel général de l'interopérabilité. RGI.
Technical report, Direction générale de la modernisation de l'état.

Ding, X., Mower, J., Subramanian, D. & Cohen, T. (2021).
Augmenting aer2vec : Enriching distributed representations of adverse event report data with orthographic and lexical information.
Journal of Biomedical Informatics, **119**, 103833.

Dong, H., Suárez-Paniagua, V., Whiteley, W. & Wu, H. (2021).
Explainable automated coding of clinical notes using hierarchical label-wise attention networks and label embedding initialisation.
J Biomed Inform, **116**, 103728.

Du, J., Wang, Q., Wang, J., Ramesh, P., Xiang, Y., Jiang, X. & Tao, C. (2021).
COVID-19 trial graph : a linked graph for covid-19 clinical trials.
J Am Med Inform Assoc, **28**(9), 1964–1969.

Duong, A., Tardif, A. & Ska, B. (2003).
Discourse about discourse : What is it and how does it progress in Alzheimer's disease ?
Brain and cognition, **53**, 177–180.

Embi, P., Jain, A., Clark, J. & Harris, C. (2005).
Development of an electronic health record-based clinical trial alert system to enhance recruitment at the point of care.
In *Ann Symp Am Med Inform Assoc (AMIA)*, pp. 231–35.

Feng, M., McSparron, J., Kien, D., Stone, D., Roberts, D., Schwartzstein, R., Vieillard-Baron, A. & Celi, L. (2018).
Transthoracic echocardiography and mortality in sepsis : analysis of the MIMIC-III database.
Intensive Care Med, **44**(6), 884–892.

Flamholz, Z. N., Crane-Droesch, A., Ungar, L. H. & Weissman, G. E. (2022).

Word embeddings trained on published case reports are lightweight, effective for clinical tasks, and free of protected health information.
J Biomed Inform, **125**, 103971.

Flesch, R. (1948).
A new readability yardstick.
Journ Appl Psychol, **23**, 221–233.

Fletcher, B., Gheorghe, A., Moore, D., Wilson, S. & Damery, S. (2012).
Improving the recruitment activity of clinicians in randomised controlled trials : A systematic review.
BMJ Open, **2**(1), 1–14.

Gabriel, R., Kuo, T., McAuley, J. & Hsu, C. (2018).
Identifying and characterizing highly similar notes in big clinical note datasets.
J Biomed Inform, **82**, 63–69.

Gaspers, J., Thiele, K., Cimiano, P., Foltz, A., Stenneken, P. & Tscherepanow, M. (2012).
An evaluation of measures to dissociate language and communication disorders from healthy controls using machine learning techniques.
In *IHI 2012*, pp. 209–218.

Grabar, N., Dalloux, C. & Claveau, V. (2020).
CAS : corpus of clinical cases in French.
Journal of BioMedical Semantics, **11**(1), 1–7.

Gunning, R. (1973).
The art of clear writing.
New York, NY : McGraw Hill.

Hamon, T. & Grabar, N. (2016).
Adaptation of cross-lingual transfer methods for the building of medical terminology in ukrainian.
In *Computational Linguistics and Intelligent Text Processing*, pp. 1–12.

Hassan, A., Ali, M. D. I., Ahammed, R., Bourouis, S. & Khan, M. M. (2021).
Development of NLP-integrated intelligent web system for e-mental health.

Comput Math Methods Med, **1**, 1546343.

Hussain, M., Satti, F. A., Hussain, J., Ali, T., Ali, S. I., Bilal, H. S. M., Park, G. H., Lee, S. & Chung, T. (2021).
A practical approach towards causality mining in clinical text using active transfer learning.
J Biomed Inform, **123**, 103932.

Jiang, J. & Zhai, C. (2007).
Instance weighting for domain adaptation in NLP.
In *Proc of the 45th Ann Meeting of the Assoc of Comp Linguistics*, pp. 264–271 : Association for Computational Linguistics.

Johnson, A. E., Pollard, T. J., Shen, L., wei H. Lehman, L., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L. A. & Mark, R. G. (2016).
MIMIC-III, a freely accessible critical care database.
Scientific Data, **3**(160035), 1–9.

Jucks, R. & Bromme, R. (2007).
Choice of words in doctor-patient communication : an analysis of health-related internet sites.
Health Commun, **21**(3), 267–77.

Kalyanam, J., Katsuki, T., Lanckriet, G. R. G. & Mackey, T. K. (2017).
Exploring trends of nonmedical use of prescription drugs and polydrug abuse in the twittersphere using unsupervised machine learning.
Addictive Behaviors, **65**, 289–295.

Kemper, S., Rash, S., Kynette, D. & Norman, S. (1990).
Telling stories : The structure of adults' narratives.
European Journal of Cognitive Psychology, **2**, 205–228.

Kim, D., Oh, J., Im, H., Yoon, M., Park, J. & Lee, J. (2021a).
Automatic classification of the korean triage acuity scale in simulated emergency rooms using speech recognition and natural language processing : a proof of concept study.
J Korean Med Sci, **36**(27), e175.

Kim, T., Han, S. W., Kang, M., Lee, S. H., Kim, J.-H., Joo, H. J. & Sohn, J. W. (2021b).

Similarity-based unsupervised spelling correction using biowordvec : Development and usability study of bacterial culture and antimicrobial susceptibility reports.

JMIR Med Inform, **9**(2), e25530.

Kincaid, J., Fishburne, R., Rogers, R. & Chissom, B. (1975).

Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel.

Technical report, Naval Technical Training, U. S. Naval Air Station, Memphis, TN.

Koptient, A. & Grabar, N. (2020).

Rated lexicon for the simplification of medical texts.

In *Proc of HEALTHINFO 2020*, pp. 1–6.

Kreimeyer, K., Foster, M., Pandey, A., Arya, N., Halford, G., Jones, S., Forshee, R., Walderhaug, M. & Botsis, T. (2017).

Natural language processing systems for capturing and standardizing unstructured clinical information : A systematic review.

J Biomed Inform, (73), 14–29.

Kynette, D. & Kemper, S. (1986).

Aging and the loss of grammatical forms : A cross-sectional study of language performance.

Language and Communication, **6**, 65–72.

Lauriola, I., Aiolli, F., Lavelli, A. & Rinaldi, F. (2021).

Learning adaptive representations for entity recognition in the biomedical domain.

J Biomed Semant, **12**(10), 1–16.

Lee, H. (2011).

Vieillissement normal et maladie d'Alzheimer : analyse comparative de la narration semi-dirigée au niveau lexical.

In *Méthodes et analyses comparatives en sciences du langage*.

Legrand, J., Toussaint, Y., Raïssi, C. & Coulet, A. (2021).

Syntax-based transfer learning for the task of biomedical relation extraction.
J Biomed Semant, **12**(16).

Leung, Y. W., Wouterloot, E., Adikari, A., Hirst, G., de Silva, D., Wong, J., Bender, J. L., Gancarz, M., Gratzer, D., Alahakoon, D. & Esplen, M. J. (2021).
Natural language processing-based virtual cofacilitator for online cancer support groups : Protocol for an algorithm development and validation study.
JMIR Res Protoc, **10**(1), e21453.

Li, M., Fei, Z., Zeng, M., Wu, F., Li, Y., Pan, Y. & Wang, J. (2018).
Automated ICD-9 coding via a deep learning approach.
In *IEEE/ACM Trans Comput Biol Bioinform*.

Liao, S., Kiros, J., Chen, J., Zhang, Z. & Chen, T. (2021).
Improving domain adaptation in de-identification of electronic health records through self-training.
J Am Med Inform Assoc, **28**(10), 2093–2100.

Lin, C.-Y. (2004).
ROUGE : a package for automatic evaluation of summaries.
In ELRA, Ed., *In Proceedings of the Workshop on Text Summarization Branches Out (WAS 2004)*, pp. 1–12, Barcelona, Spain.

Lindberg, D., Humphreys, B. & McCray, A. (1993).
The Unified Medical Language System.
Methods Inf Med, **32**(4), 281–291.

Liu, F., Shareghi, E., Meng, Z., Basaldella, M. & Collier, N. (2021a).
Self-alignment pretraining for biomedical entity representations.
In *Proc of the 2021 Conf of the North American Chapter of the Ass for Comp Linguistics : Human Language Technologies*, pp. 4228–4238 : Association for Computational Linguistics.

Liu, H., Chi, Y., Butler, A., Sun, Y. & Weng, C. (2021b).
A knowledge base of clinical trial eligibility criteria.
J Biomed Inform, **117**, 103771.

Liu, Y., Cheng, H., Klopfer, R., Gormley, M. R. & Schaaf, T. (2021c).
Effective convolutional attention network for multi-label clinical document classification.
In *Proc of the 2021 Conf on Empirical Methods in Natural Language Processing*, pp. 5941–5953, Online and Punta Cana, Dominican Republic : Association for Computational Linguistics.

Majewska, O., Collins, C., Baker, S., Björne, J., Brown, S. W., Korhonen, A. & Palmer, M. (2021).
BioVerbNet : a large semantic-syntactic classification of verbs in biomedicine.
J Biomed Semantics, **12**(1), 12.

McLaughlin, G. H. (1969).
SMOG grading – a new readability formula.
Journal of reading, **12**(8), 639–646.

Michalopoulos, G., Wang, Y., Kaka, H., Chen, H. & Wong, A. (2021).
UmlsBERT : Clinical domain knowledge augmentation of contextual embeddings using the Unified Medical Language System Metathesaurus.
In *Proc of the 2021 Conf of the North American Chapter of the Ass for Comp Linguistics : Human Language Technologies*, pp. 1744–1753 : Association for Computational Linguistics.

Mintz, M., Bills, S., Snow, R. & Jurafsky, D. (2009).
Distant supervision for relation extraction without labeled data.
In *Proc of the Joint Conf of the 47th Ann Meeting of the ACL and the 4th Int Joint Conf on Natural Language Processing of the AFNLP*, pp. 1003–1011, Suntec, Singapore : Association for Computational Linguistics.

Nef, F. & Hupet, M. (1992).
Les manifestations du vieillissement normal dans le langage spontané oral et écrit.
L'année psychologique, **9**(3), 393–419.

NLM (2001).
Medical Subject Headings.
National Library of Medicine, Bethesda, Maryland.
www.nlm.nih.gov/mesh/meshhome.html.

Noh, J. & Kavuluru, R. (2021).
Improved biomedical word embeddings in the transformer era.
J Biomed Inform, **120**, 103867.

Olasov, B. & Sim, I. (2006).
Ruleed, a web-based semantic network interface for constructing and revising computable eligibility rules.
In *Ann Symp Am Med Inform Assoc (AMIA)*, pp. 1051.

Oleynik, M., Kugic, A., Kasáč, Z. & Kreuzthaler, M. (2019).
Evaluating shallow and deep learning strategies for the 2018 n2c2 shared task on clinical text classification.
J Am Med Inform Assoc, **26**(11), 1247–1254.

OMS (1995).
Classification statistique internationale des maladies et des problèmes de santé connexes — Dixième révision.
Organisation mondiale de la Santé, Genève.

Osadchiy, V., Jiang, T., Mills, J. & Eleswarapu, S. (2020).
Low testosterone on social media : Application of natural language processing to understand patients' perceptions of hypogonadism and its treatment.
J Med Internet Res, (22), 1–15.

Papineni, K., Roukos, S., Ward, T., Henderson, J. & Reeder, F. (2002).
BLEU : a method for automatic evaluation of machine translation.
In *Annual Meeting of the Association for Computational Linguistics*, pp. 311–318.

Patel, V., Branch, T. & Arocha, J. (2002).
Errors in interpreting quantities as procedures : The case of pharmaceutical labels.
Int Journ Med Inform, **65**(3), 193–211.

Percha, B., Pisapati, K., Gao, C. & Schmidt, H. (2021).
Natural language inference for curation of structured clinical registries from unstructured text.
J Am Med Inform Assoc, **29**(1), 97–108.


Perotte, A., Pivovarov, R., Natarajan, K., Weiskopf, N., Wood, F. & Elhadad, N. (2014).
Diagnosis code assignment : models and evaluation metrics.
J Am Med Inform Assoc, **21**, 231–237.


Plank, B. & Moschitti, A. (2013).
Embedding semantic similarity in tree kernels for domain adaptation of relation extraction.
In *Proc of the 51st Ann Meeting of the Assoc for Comp Linguistics*, pp. 1498–1507.


Pressler, T., Yen, P., Ding, J., Liu, J., Embi, P. & Payne, P. (2012).
Computational challenges and human factors influencing the design and use of clinical research participant eligibility pre-screening tools.
BMC Med Inform Dec Mak, **12**, 47.


Reiner, B. (2018).
Quantitative analysis of uncertainty in medical reporting : Creating a standardized and objective methodology.
J Digit Imaging, **31**(2), 145–149.


Riedel, S., Yao, L. & McCallum, A. (2010).
Modeling relations and their mentions without labeled text.
In *Proc of ECML PKDD*, pp. 148–163.


Ross, J., Tu, S., Carini, S. & Sim, I. (2010).
Analysis of eligibility criteria complexity in clinical trials.
In *Summit on Translational Bioinformatics*, pp. 46–50.


Rousseau, T., De Saint-André, A. & Gatignol, P. (2009).
Évaluation pragmatique de la communication des personnes âgées saines.
Neurologie Psychiatrie Gériatrie, **9**, 271–280.


Roy, A. & Pan, S. (2021).
Incorporating medical knowledge in BERT for clinical relation extraction.
In *Proc of the 2021 Conf on Empirical Methods in Natural Language Processing*, pp. 5357–5366, Online and Punta Cana, Dominican Republic : Association for Computational Linguistics.

Rozova, V., Witt, K., Robinson, J., Li, Y. & Verspoor, K. (2022).
Detection of self-harm and suicidal ideation in emergency department triage notes.
J of the Am Med Inform Ass, **29**(3), 472–480.

Ruch, P., Gobeil, J., Tbahriti, I. & Geissbühler, A. (2008).
From episodes of care to diagnosis codes : automatic text categorization for medico-economic ecoding.
In *Ann Symp Am Med Inform Assoc (AMIA)*, pp. 636–40.

Shiner, B., Levis, M., Dufort, V. M., Patterson, O. V., Watts, B. V., DuVall, S. L., Russ, C. J. & Maguen, S. (2021).
Improvements to PTSD quality metrics with natural language processing.
J Eval Clin Pract.

Shivade, C., Raghavan, P., Fosler-Lussier, E., Embi, P., Elhadad, N., Johnson, S. & Lai, A. (2014).
A review of approaches to identifying patient phenotype cohorts using electronic health records.
J Am Med Inform Assoc, **21**(2), 221–30.

Simianu, V. V., Grounds, M. A., Joslyn, S. L., LeClerc, J. E., Ehlers, A. P., Agrawal, N., Alfonso-Cristancho, R., Flaxman, A. D., & Flum, D. R. (2016).
Understanding clinical and non-clinical decisions under uncertainty : a scenario-based survey.
BMC Medical Informatics and Decision Making, **16**(153), 1–9.

Ska, B. & Duong, A. (2005).
Communication, discours et démence.
Psychologie et NeuroPsychiatrie du Vieillissement, **3**(2), 125–133.

Stroetmann, V. N., Kalra, D., Lewalle, P., Rector, A., Rodrigues, J.-M., Stroetmann, K. A., Surjan, G., Ustun, B., Virtanen, M. & Zanstra, P. E. (2009).
Semantic interoperability for better health and safer healthcare.
Technical report, European commission, Information society and media.

Sun, W., Rumshisky, A. & Uzuner, Ö. (2013).
Evaluating temporal relations in clinical text : 2012 I2B2 challenge.

JAMIA, **20**(5), 806–813.

Surdeanu, M., Tibshirani, J., Nallapati, R. & Manning, C. D. (2012).
Multi-instance multi-label learning for relation extraction.
In *Proc of the 2012 Joint Conf on Empirical Methods in Natural Language Processing and Comp Natural Language Learning*, pp. 455–465, Jeju Island, Korea : Association for Computational Linguistics.

Trombert-Paviot, B., Rector, A., Baud, R., Zanstra, P., Martin, C., van der Haring, E., Clavel, L. & Rodrigues, J. (2003).
The development of ccam : the new French coding system of clinical procedures.
HIM J, **31**(1), 1–11.

Tu, S., Peleg, M., Carini, S., Bobak, M., Ross, J., Rubin, D. & Sim, I. (2011).
A practical method for transforming free-text eligibility criteria into computable criteria.
J Biomed Inform, **44**(2), 239–50.

Uzuner, O. (2008).
Second I2B2 workshop on natural language processing challenges for clinical records.
In *Ann Symp Am Med Inform Assoc (AMIA)*, pp. 1252–3.

Uzuner, O., South, B. R., Shen, S. & DuVall, S. L. (2011).
2010 I2B2/VA challenge on concepts, assertions, and relations in clinical text.
J Am Med Inform Assoc, **18**(5), 552–556.

Viani, N., Botelle, R., Kerwin, J., Yin, L., Patel, R., Stewart, R. & Velupillai, S. (2021).
A natural language processing approach for identifying temporal disease onset information from mental healthcare text.
Sci Rep, **11**(1), 757.

Wajsbürt, P., Sarfati, A. & Tannier, X. (2021).
Medical concept normalization in French using multilingual terminologies and contextual embeddings.
J Biomed Inform, **114**, 103684.

Walker, J., Pan, E., Johnston, D., Adler-Milstein, J., Bates, D. W. & Middleton, B. (2005).

The Value of Healthcare Information Exchange and Interoperability.
Technical report, US Health Affairs.

Wang, A., Sable, J. & Spackman, K. (2002).
The snomed clinical terms development process : refinement and analysis of content.
In *Ann Symp Am Med Inform Assoc (AMIA)*, pp. 845–9.

Wang, J., Abu-El-Rub, N., Gray, J., Pham, H. A., Zhou, Y., Manion, F. J., Liu, M., Song, X., Xu, H., Rouhizadeh, M. & Zhang, Y. (2021).
COVID-19 SignSym : a fast adaptation of a general clinical NLP tool to identify and normalize COVID-19 signs and symptoms to omop common data model.
J Am Med Inform Assoc, **28**(6), 1275–1283.

Williams, M., Parker, R., Baker, D., Parikh, N., Pitkin, K., Coates, W. & Nurss, J. (1995).
Inadequate functional health literacy among patients at two public hospitals.
JAMA, **274**(21), 1677–1682.

Xu, W., Napoles, C., Pavlick, E., Chen, Q. & Callison-Burch, C. (2016).
Optimizing statistical machine translation for text simplification.
Transactions of the Association for Computational Linguistics, **4**, 401–415.

Zhou, T., Cao, P., Chen, Y., Liu, K., Zhao, J., Niu, K., Chong, W. & Liu, S. (2021).
Automatic ICD coding via interactive shared representation networks with self-distillation mechanism.
In ACL, Ed., *Proc of the 59th Ann Meeting of the Assoc for Comp Linguistics and the 11th Inter Joint Conf on Natural Language Processing*, pp. 5948–5957.