

DOING: Données intelligentes



18 décembre 2020

DOING en différentes couleurs

@DIAMS

- Dynamique locale
- Base pour la formation des sous-groupes de travail

@MADICS

- avec Anne-Lyse (LLL) et Genoveva (LIRIS)
- Perspective nationale
- Ouverture, plus d'opportunités

DOING : un atelier MADICS

- Planification : journée d'étude le 8 avril à Paris (annulée)
- Deux webinars (1 keynote + 2 exposés courts)
 - ▶ le 5 juin (26 participants) et
 - ▶ le 17 juin (36 participants)
- 2nd Symposium GDR CNRS MADICS - Webinar DOING : 17 Juillet (40 participants) avec 2 keynotes
- Participation à d'autres ateliers (Symposium) : RoD et HN
- **Workshop DOING@ADBIS-TPDL-EDA** : Carmem (UFPR) et Mirian (UO)
25 août 2020, environ 35 participants ; 17 soumissions (8 long + 1 court)
- Mise en place des sous-groupes de travail ; co-encadrement de deux stages
(**financement ICVL et LIFO**)
- Développement des **prototypes** au LIG (Genoveva) : des workflows pour l'exploration des données textuelles
- Réunion avec HN-MADICS
- Contacts avec des spécialistes : l'UFJF (Brésil)

DOING 2021 (1)

La bonne nouvelle

Le comité de direction du GDR MaDICS a le plaisir de vous annoncer que votre proposition a été retenue.

L'Action DOING est ainsi créée à partir du 1er janvier 2021

Les bons commentaires

- Le comité a apprécié le **fond scientifique de votre proposition qui tout en étant particulièrement ambitieux dégage des objectifs scientifiques précis sur l'extraction d'information de données textuelles pour la construction et l'interrogation de bases de connaissance**
- Le projet, croisant des domaines variés de l'informatique, nous paraît **fédérateur** avec des objectifs d'animation en phase avec celui-ci

Les recommandations

Nous vous recommandons d'ajouter une implication **d'experts métier** sur les jeux de données identifiés par DOING (santé, environnement) ce qui conduira à accroître l'interdisciplinarité.

DOING : proposition d'action

Des petits changements : les nuances... importantes

- L'action DOING se concentre sur la **transformation des données en information puis en connaissance**
- Collaboration TAL, BD et IA
 - ▶ **extraire des informations** dans les données **textuelles**
 - ▶ peupler des bases de connaissances (base de donnée **graphe**)
 - ▶ **manipulation et maintenance intelligente** de la base
 - ▶ requêtes *data science* : vers des mécanismes d'analyse efficaces, flexibles, sûrs, adaptés à l'utilisateur et qui respectent des contraintes de qualité et de vie privée
 - ▶ intérêt particulier : *graphes de propriété, Cypher, la sémantique du monde fermé, algorithmes d'analyse des graphes*
 - ▶ Un contour : **extraire et stocker les données les plus intéressantes et importantes pour les consultations et les analyses des experts**
↪ travail en amont pour établir l'intérêt des experts !

Santé

- Corpus de cas cliniques développé et annoté pour DEFT 2019/2020 et accessible après acceptation de la licence ;
- Corpus de cas cliniques en cours de développement (fin du projet en juillet 2021) multilingue (français, anglais, italien, espagnol, basque) dans le cadre du projet ELG (European Language Grid) E3C () ; licence CC-BY-NC-SA ;
- QUAERO FrenchMedical corpus qui contient entre autres des notices de médicament annotées ;
- COVID-19 Open Research Dataset (non annoté)
- S2ORC : The Semantic Scholar Open Research Corpus (non annoté)

Environnement

- S2ORC : The Semantic Scholar Open Research Corpus (non annoté)
- PANACEA Environment French monolingual corpus (non annoté) :
- ARD-CVL-JUNON (requêtes *data science*)

Actions prévues

Le projet...

- Ateliers de travail (nos réunions @DIAMS et @MADICS)
- Mise en place des sous-groupes de travail : co-encadrements divers
- Organiser des tutoriaux
- 2nd International Workshop DOING@ADBIS : Tartu, Estonia 2021 (24 août)
- Préparation d'un article de synthèse
- Travail avec médecins UFJF (préparation/mission)
- Journées communes avec autres ateliers

Les opportunités ...

- Des invités (PAMDA : projet Studium, proposition S. Limet))
- Stages
- Thèses
- Les publications
- Interaction avec JUNON
- Élaboration des projets