CONFÉRENCES PROFESSIONNELLES

- Onate Tejeiro
  - APP
    - ROLE
      - Intervenant
        - Journalisme-information
          - Informatique
          - Pas spécialiste en TAL
      - Entreprise
        - n agence de presse
        - Journalisme
          - reporters
          - éditeurs de texte
          - photographes pro
          - Service photo de l'AFP
      - De nombreux prix et reconnaissances
      - Actualités
        - Infographies
          - Multilingue
            - anglais
            - français
            - arabe
            - espagnol
            - portugais
            - allemand
        - Vidéographies
          - Types
            - textes
            - images
            - photos
            - vidéos
          - modélisations 3D
          - Compréhension approfondie
            - commentaire audio
            - script
          - Domaines
            - science
            - environnement
            - technologie
            - espace
    - FACT CHECKING
      - Depuis 2017
      - Identifie ce qui est faux, infondé, très trompeur
      - Cellule d'investigation numérique
      - Ex : Russes qui propagent leur d'infos fausses
      - Site dédié : https://factual.afp.com/
      - ?
    - Multilingue
      - 24 langues
      - 140 journalistes
    - MEDIALAB
      - Scribe
        - Extraction & résolution des entités nommées
          - basé sur des ontologies
        - EDyLan
          - Environnement de ressources textuelles
        - Transcriber
          - Transcription de la parole
        - Chronoline
          - Information temporelle
        - ARIEL
          - Acquisition de schémas pour la reconnaissance et l'annotation d'événements liés
        - APIMR
          - Moteur de recherche
            - Entités nommées
          - API
          - Visualisations multimedias
          - WHAT, WHO, WHEN, WHERE
          - HTML5 videos
            - Videos interactives
        - IrvIS, WeVerify, vera.ai
          - Vérification de l'info
            - Fake news
    - LLMs
  - Alexandre Courtiaux
    - PARCOURS
      - Linguistique
        - M1 TAL à INALCO
        - M2 TAL PRO à Nanterre
    - CRÉEST AGRICOLE
      - Chef de projet info-linguiste
      - Son rôle
        - Gestion des projets
          - Méthode AGILE
        - Business analyst
        - Prototématrices & solutions du TAL et IA
        - Gestion de projet
        - Métier + technique
          - Start-up
          - Polyvalence professionnelle
        - Grande entreprise
          - Bug du compétition
          - Grosse demande de compétences technique
        - Info-linguiste
          - Chef de projet
          - Assistant de projet
  - Hélène Flamen
    - PARCOURS
      - Linguistique
        - Licence SOL
        - Master SOL
        - Doctorat
          - Thèse
            - Laboratoire Ligérien de Linguistique, Orléans
            - Étude de la perception d'une ville
              - Représentations, analyse et visualisation
              - Corpus : ESLO
    - SNCF
      - ?
      - Plateau IA
        - Vision Artificielle
        - Prédiction
        - Cognitique
        - IA de confiance
        - TALN
      - Rôle
        - Chef de projet : TAL & IA Générative
        - Gestion de projets
      - TAL
        - Analyse émotionnelle
        - Agent conversationnel
        - Analyse de REX
          - Exploration doc
            - PyQtBeef
              - 2020-2021
                - Fusition 16 ffé dans la documentation
                - Extraction & structuration des données
                - Parsing de doc.
                - Extraction d'infos
                - Suivi de mise à jour de contenus
                - Identification d'infos contradictoires
      - Nouvelles perspectives
        - Hub France IA
        - Axes non-technicaux
          - Développement & adoption d'une IA responsable
        - Chat GPT
          - 5 millards de parametres
          - Fiabilité des réponses ?
          - Gouvernance des données ?
          - Coûts financiers et environnementaux ?
        - LLM
          - Usages ?
            - Capacité de ...
              - rédaction
              - classification
              - traduction
              - synthèse
    - ENACT
      - Cluster IA en Grand Est
        - Catalyseur pour l'ingénierie & la découverte scientifique
          - TAL
          - LLM multimodaux
          - Résolution de problèmes complexes
        - Accès facilité aux compétences
        - Veille technologique
        - Développement de projets
        - Réseau
  - Géraldine Damnati
    - ORANGE
      - R&D
        - Innovation
        - Data & AI
        - Intelligence artificielle
      - Équipe Deskit
      - TAL
        - Apprentie : base compétences
          - Reprendre une compétence à l'aide de modèles pré-entraînés mais contraints de zero-shot
            - Grammaires
            - Langages formels
        - Apprentie : numériques
          - Reprendre la performance linguistique à l'aide de modèles pré-entraînés mais moins contraints de zero-shot
            - Observation de corpus
            - Apprentissage automatique
            - Modèles préalablés
            - Deep-learning
    - TRANSLIB
      - Linguistique
        - SIMBOW
          - Similarité sémantique
        - BoW
        - Présente sémantique entre mots/ entre phrases
      - Analyse syntaxique et dépendance
        - Analyse de relaxes clients
      - Analyse multilingue d'opinion
        - MRQA
          - Machine Reading Question Answering
            - Tâche de compréhension de lecture
              - SEXT
      - Analyse sémantique
        - Modélisation des mots & leurs interactions dans une phrase
          - Frameïdel
            - BI-LSTM
              - Étiquetage de séquence
            - Abstract Meaning Representation
  - Bénédicte Goujon
    - PARCOURS
      - SCIENTIFIQUE
        - M1 MASS
          - maths, info, éco
        - M2 : DEA TALN
        - Thèse CIFRE : 2000
    - THALES
      - Secteurs d'activité
        - Identité & sécurité numérique
        - Transport terrestre
        - Espace
        - Aéronautique
        - Défense et sécurité
      - Missions
        - traitement & prise de décision du data
        - transmission & stockage du data
        - détection & collecte
      - TRT
        - Phase 1
          - 2006 - 2011
            - Apprentie symbolique
              - règles écrits manuellement
                - Unitex
                  - Limbes
                    - Cadi dével. : construction connaissances
                    - Difficile de couvrir tous les cas
            - Apprentie statistique
              - Corpus annotés volumineux
            - TVS
              - 2003-2004
              - Plateforme commerciale vocaleo
                - Reconnaissance vocale
              - Écritures grammaires
                - génériques
                - spécifiques
            - Outilex
              - ANR INTL
              - 2004-2006
              - Plateforme-style Unitex
                - Gestion XML
                - Dictionnaires multilingues
                - ?
                - Filtrage thématique de doc audio
            - Annotation automatique de l'incertain dans des textes
              - 2007-2010
              - Extraction auto d'événements + aspects
                - temps
                - incertitude
                - polarité
                - source
            - Outil
              - SemFlux
                - 2005-2011
                - java
                - Extraction de relations binaires entre entités nommées
        - Phase 2
          - 2016 - présent
            - Extraction d'infos
              - Recrutés
                - Entités & événements spécifiques
                - Arguments relient événements & entités
                - Relations entre entités
            - STRASS
              - 2020
              - Extraction de relations entre entités
                - python
                - spaCy
              - Aide à acquisition de connaissances linguistiques
            - INCEPTUM
              - Annotation manuelle
              - Ontologie
    - Chat GPT
      - ?
      - Support pour des tâches
        - Génération de code
        - Génération de textes bien formés
      - ?
      - Pas de connaissances valables et structurées
        - Peut affirmer des contre vérités
        - Aide à la gestion de connaissances