

# TAL et RH

---

# Plan

---

- Le recrutement
  - Les métiers RH et les processus
  - Les outils, les pratiques
- Les outils TAL pour l'emploi
  - L'analyse de CV
  - L'analyse d'offres
  - Le matching
- Conclusion

# Où peut intervenir le TAL en RH ?

---

- Le recrutement / la recherche d'emploi
  - deux faces pas vraiment symétriques du même problème
- La « gestion des talents » (aka GPEC ou mobilité interne)
  - voir aussi le problème spécifique du « staffing »
- Les entretiens annuels et autres obligations légales

# Les métiers et les processus

---

# Les acteurs du recrutement

---

- Les acteurs traditionnels :
  - service public de l'emploi (Pôle Emploi),
  - entreprises,
  - cabinets de recrutement (Michael Page, Adecco, Manpower, ...),
  - associations (Apec)
- Les sites spécialisés en emploi
  - sites d'emploi (« jobboards ») généralistes (Monster, Meteojob, ...) ou sectoriels (dogfinance, fashionjob, agrojob, ...)
  - agrégateurs d'offres (Keljob, Jobijoba, Indeed),
  - réseaux sociaux professionnels (Linkedin, Viadeo, Xing, ...)
- Les nouveaux entrants (sites d'offres, de freelance, ...)
  - leboncoin.fr est le deuxième site d'emploi en France après Pole-Emploi, a développé une offre orientée emploi et va proposer une offre « cadres »
  - les applications mobiles (cf. Kudoz)

# Une prééminence du numérique, comme partout

---

- Tous les acteurs sont sur le Net qui est aujourd’hui leur canal principal (voire unique)
- Pourtant...
  - 18% des ménages français n’ont pas accès à Internet
  - 49% des français n’utilisent pas l’internet mobile
  - (source: Rapport du COE, mars 2015)
- Pour les chercheurs d’emploi, l’outil le plus important est :
  - les sites d’emploi (62%)
  - les candidatures spontanées (10%)
  - Pôle Emploi (8%)
  - le réseau personnel (7%)
  - (source: enquête RegionJobs, 2013)

# Classement des sites d'emploi

## Top 5 des sites d'Emploi / Carrière les plus visités en septembre 2016

| Rang | Brands / Channels                    | Visiteurs uniques<br>mensuels | Couverture<br>visiteurs uniques | Couverture<br>France Entière |
|------|--------------------------------------|-------------------------------|---------------------------------|------------------------------|
| 1    | Pole Emploi                          | <b>7 302 000</b>              | 15,5%                           | 11,7%                        |
| 2    | Indeed                               | <b>2 794 000</b>              | 5,9%                            | 4,5%                         |
| 3    | <i>Leboncoin.fr</i> Emploi - TS ACPM | <b>2 783 000</b>              | 5,9%                            | 4,5%                         |
| 4    | Meteojob - TS                        | <b>1 040 000</b>              | 2,2%                            | 1,7%                         |
| 5    | APEC                                 | <b>989 000</b>                | 2,1%                            | 1,6%                         |

*Les channels sont mentionnés en italique.*

Données Mediamétrie/Netratings Septembre 2016

# Modèle économique

---

- Une quasi-constante : la gratuité pour le candidat qui dépose son CV
  - Quelques exceptions (cf. Qapa qui permet de « promouvoir » son CV moyennant finance)
- Les jobboards sont rémunérés par les annonceurs, les aggrégateurs par la publicité et par les renvois de candidature vers les jobboards
  - Motivation : l'espoir en l'effet vertueux : plus il y a de CV, plus le site est attractif pour les annonceurs (resp. pour les visiteurs, donc pour une publicité plus rentable)
  - ... ce qui à son tour facilitera la vente des offres, donc le nombre d'offres en ligne
- Les réseaux sociaux professionnels vendent des « accès recruteurs » qui permettent de dépasser la limite du réseau personnel (amis, amis des amis)
  - En plus de la gratuité, les RSP poussent à la valorisation du candidat (vitrine)

# On recrute aussi via Facebook

---

- Possibilité de donner de plus en plus d'informations professionnelles dans le profil Facebook
- Promotion des « pages entreprise »
  - deviennent rapidement payantes, en particulier pour la création d'une offre
- Offres d'optimisation du recrutement sur Facebook (Work4, BranchOut, OhMyJob)
  - optimisation des publicités pour l'entreprise
  - mini-site recrutement dédié
  - cooptation/développement d'un réseau professionnel au sein de Facebook

# Autres sources d'emploi : les plateformes

---

- Sites de « jobbing » (échanges de services entre particuliers)
  - Youpjob, jemepropose, etc.
- Les plateformes de services
  - Uber, mais aussi helping pour le ménage, etc.
  - Avec les problèmes légaux/éthiques que cela entraîne (cf. faillite de homejoy, plate-forme de ménage à domicile, suite à des plaintes et « class-actions » pour requalifier les « abonnés » en salariés)
    - Lyft, Instacart (courses), Postmates (livraison) ou Shyp (envoi de colis) sont aussi attaqués en justice.
- Les sites de freelance
  - depuis longtemps installés pour les traducteurs, p.ex.
- Les plateformes de crowdsourcing
  - pour la plupart illégales en France (cf. Amazon Mechanical Turk)

# Parenthèse sur la « marque employeur »

---

- La promotion de la « marque employeur », c'est le fait, pour une marque, de se présenter sous son meilleur jour afin d'attirer les meilleurs talents
  - D'où les pages Facebook, LinkedIn, etc.
  - Très souvent payantes et assimilées au SEO
- On rejoint la e-réputation, même si les efforts (et les budgets) mis en œuvre sont négligeables par rapport à ceux destinés à l'augmentation du chiffre d'affaires
  - ... ou d'empêcher qu'il baisse (cf. gestion des crises)

# Pour le recruteur

---

- Quatre grandes étapes dans le processus
  - recherche de candidats pertinents (« sourcing »)
  - traitement et analyse des candidatures
  - sélection et évaluation des candidats
  - intégration du candidat (et fidélisation)
- Le numérique est principalement sur les deux premières étapes
  - un peu sur les aspects sélection/intégration via les outils de test en ligne, les serious games, etc.
  - le tout passant essentiellement par l'écrit (malgré quelques timides avancées de la video - cf visiotalent)
- Outilage : ATS (Applicant Tracking System) ou... Excel

# Le « sourcing »

---

- Suite à diffusion d'annonce ou non (« chasse de tête »)
- S'effectue le plus souvent via les sites d'emploi en ligne
  - consultation des CV-thèques pour 72% des recruteurs
- Souvent complémenté par la recherche sur les réseaux sociaux (LinkedIn, Viadeo essentiellement - peu de facebook/twitter)
  - (quasi-systématiquement pour des profils de type cadre/bac+5)
  - pour affiner la perception
  - gérer l'obsolescence du CV : un CV déposé dans une CV-thèque est très rarement remis à jour, le profil social très fréquemment
  - pour entrer en contact (si les coordonnées ne sont pas sur le CV ou sont cachées par le job-board)
- Les RSP sont parfois aussi des outils de chasse pour les profils pénuriques

# L'analyse des candidatures

---

- Cette analyse passe souvent (dans les grosses structures) par un tiers, le chargé de recrutement (aka DRH, RRH, etc.)
  - ... qui ne connaît pas vraiment le métier pour lequel il recrute ni les besoins du recrutement
- C'est l'intermédiaire entre le donneur d'ordres (le manager qui veut recruter) et le candidat
- Il doit qualifier les candidatures souvent simplement au vu du CV et de la grille de compétences qu'il possède
  - mais dont il ne maîtrise généralement pas la signification
- Ses outils : les données que lui a fournies le donneur d'ordres (formations, compétences), les référentiels métier et sa propre expérience
  - tout cela est parfois bien peu...

# Les fonctions d'un ATS

---

- Collecter les candidatures
  - souvent automatisée via adresse mail dédiée
  - phase qui passe très fréquemment par une analyse automatique de CV
  - ... qui conditionne tout aussi fréquemment la suite du processus
  - (voir le post « How to beat an ATS » sur ResumeGenius)
- Constitution d'une CV-thèque (« vivier ») et recherche dans cette CV-thèque
- Envoi d'e-mail (pour RV ou pour décliner)
- Le tri des candidatures
  - plus ou moins outillé et très dépendant des technologies d'analyse/matching
- La rédaction et l'envoi des offres (via connecteurs vers les job-boards ou vers des plateformes spécialisées comme Multiposting)
- Quelques ATS : Taleo, CarrerBuilder, TalentSoft, Workday

# Les outils, les pratiques

---

# Le royaume des référentiels

---

- Idée de base : avoir une description exhaustive des métiers, des compétences requises, des formations qui fournissent ces compétences et des évolutions possibles
- Ainsi le processus de recrutement se standardise et peut être opéré par des gens dont c'est le métier
  - et peut être appuyé par des outils standard
- Enfin, valider l'adéquation d'un candidat à un poste revient à valider le nombre de critères qu'il remplit en mettant en regard la formalisation de son profil et celle de l'offre
- Le recrutement, n'est-ce donc finalement qu'un problème de classification ?

# ROME (Répertoire Opérationnel des Métiers et Emplois)

---

- Référentiel-roi des référentiels pour l'emploi
- Conçu en 1989 par Pôle Emploi
- Recense 531 fiches-métier qui couvrent plus de 10.000 « appellations »
  - Une fiche fournit la définition du métier, les conditions d'accès, les activités, les compétences requises, etc.
- A été pendant longtemps l'unique point d'entrée pour la recherche sur [pole-emploi.fr](http://pole-emploi.fr)
  - et pour la recherche d'emplois par les conseillers
- Travail éditorial considérable
  - refondu en 1993, puis en 2009 (ROME-V3), amélioration continue depuis 2 ans
  - mobilise une équipe de plusieurs dizaines de personnes
  - mises à jour mineures plusieurs fois par an
- Récemment rendu disponible en open-data (via Etalab)
- Sert de pivot pour d'autres institutions (dont l'INSEE)

# Un extrait de fiche ROME

## M1805 - Études et développement informatique

### Appellations

- Analyste cogniticien / cogniticienne informatique
- Analyste concepteur / conceptrice informatique
- Analyste d'application
- Analyste de gestion informatique
- Analyste décisionnel - Business Intelligence
- Analyste d'étude informatique
- Analyste développeur / développeuse
- Analyste en cybersécurité
- Analyste fonctionnel / fonctionnelle informatique
- Analyste organique informatique
- Analyste réseau informatique
- Analyste responsable d'application informatique
- Analyste télématicien
- Analyste-programmeur / Analyste-programmeuse de micro-informatique
- Analyste-programmeur / Analyste-programmeuse d'étude informatique
- Analyste-programmeur / Analyste-programmeuse en informatique industrielle
- Analyste-programmeur / Analyste-programmeuse gestion informatique
- Analyste-programmeur / Analyste-programmeuse informatique
- Chef de projet TMA - Tierce Maintenance Applicative
- Concepteur / Conceptrice d'application informatique
- Concepteur / Conceptrice informatique
- Concepteur / Conceptrice logiciel informatique
- Développeur / Développeuse - jeux vidéo
- Développeur / Développeuse Big Data
- Développeur / Développeuse d'application
- Développeur / Développeuse de sécurité des systèmes d'information
- Développeur / Développeuse décisionnel - Business Intelligence
- Développeur / Développeuse full-stack
- Développeur / Développeuse informatique
- Développeur / Développeuse multimédia
- Développeur / Développeuse web
- Développeur / Développeuse web mobile
- Didacticien / Didacticienne informatique
- Homologateur / Homologatrice logiciel
- Informaticien / Informaticienne analyste
- Informaticien / Informaticienne d'application

# D'autres référentiels

---

- ESCO : initiative au niveau européen
  - tente de lier métiers, compétences, certifications, domaines d'activité
  - avec des passerelles entre ces éléments (« piliers »)
  - travail en cours dans le cadre de H2020
  - volonté de fluidifier le marché au niveau européen
- Référentiels INSEE
  - Métiers : FAP, PCS
  - Activités : NAF, NACE
- Référentiel de l'Apec
- Chaque branche de métier définit son référentiel
- (chacun constitue un problème intéressant pour appliquer des techniques de classification TAL)

# Le problème des référentiels

---

- Métiers trop pointus pour trouver leur correspondance dans un référentiel généraliste
  - p.ex. le « linguiste-informaticien » n'existe pas dans le ROME
  - correspondent souvent à des métiers pénuriques (cf. « data-scientist », « ingénieur data », etc.)
- Ou à l'inverse des visions trop larges des métiers qui rendent inutilisable la classification elle-même
  - cf. le ROME M1806 qui regroupe les fonctions de consultant en informatique
  - Ou le M1803 qui comprend toutes les fonctions de direction des services informatiques (du chef de projet au DSI)
    - dont le « Chef de projet en linguistique informatique » !

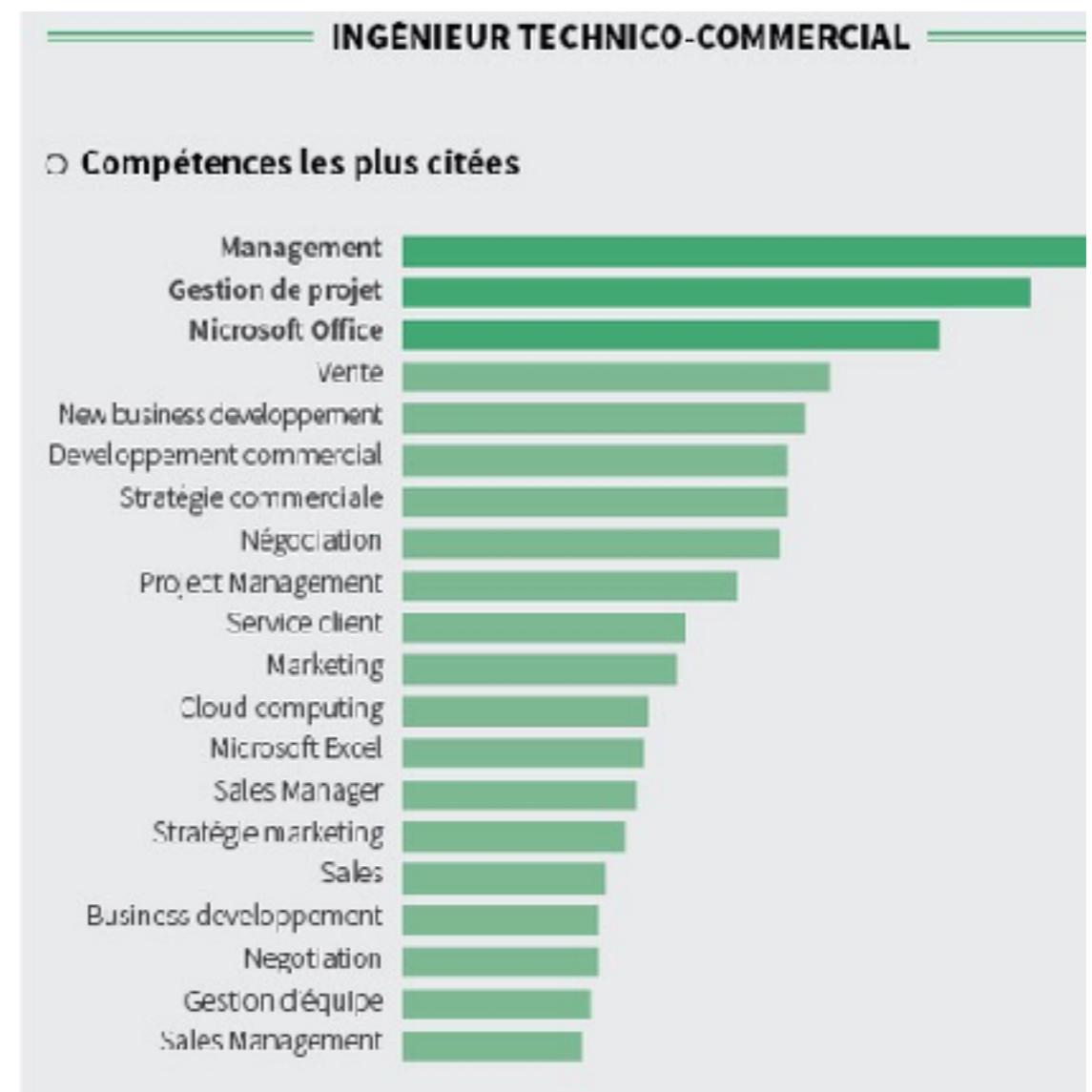
# Une alternative : les folksonomies ?

---

- Approche utilisée par les réseaux sociaux professionnels
  - évite d'avoir à se positionner sur un référentiel plutôt qu'un autre
- On laisse les utilisateurs proposer des compétences, des métiers, des « tags » en général pour décrire leurs relations professionnelles
  - ceux qui ont le plus de clics deviennent plus souvent proposés et le système est censé converger
- Mais rien ne garantit cette convergence

# Un exemple de perte de sens

- Extrait d'une étude réalisée par LinkedIn pour Le Monde Campus sur les métiers les plus demandés : les 20 métiers les plus demandés sont dans le monde du logiciel
- Liste en illustration les compétences extraites des tags de LinkedIn
- Beaucoup de compétences équivalentes et pourtant distinguées
  - « Gestion de projet » vs. « Project Management »
  - « Sales » vs. « Vente »
  - « new business development », « développement commercial », « business development », « stratégie commerciale »
- Conclusion : on retombe dans le même travers que les classifications « top-down » : si vous vous inscrivez dans la mauvaise case, vous n'êtes pas visible du recruteur



# L'analyse de CV

---

Les outils TAL pour l'emploi

# Cas d'usage

---

- Les analyseurs de CV sont utilisés dans les ATS et les sites de recrutement
  - présentent un formulaire pré-rempli à l'utilisateur, parfois restreint aux données importantes, c'est-à-dire l'état-civil (!)
  - principe : simplifier la vie au candidat/recruteur, mais aussi capter les « meilleurs » candidats
- La plupart se présentent sous la forme de Web Services
- Quelques analyseurs : LeaCV, textkernel, RChili, Daxtra, Sovren, CVReader
  - CVReader permet un test en ligne

# Exemple d'intégration (1/4)

The screenshot shows the Manpower website's 'Mon Manpower' section. At the top, there is a navigation bar with links for ACCUEIL, CANDIDATS, ENTREPRISES, LE GROUPE, and MON MANPOWER. There is also a search bar labeled 'Rechercher' with an 'OK' button, and links for 'Créer un compte' and 'Se connecter'.

The main content area has a breadcrumb navigation 'Accueil > Mon Manpower'. The title 'Espace Mon Manpower' is displayed prominently. Below it, the heading 'CRÉER mon compte mon Manpower' is shown.

The registration form is divided into four steps:

- 1 Votre profil**: A dropdown menu labeled 'Vous êtes :' with the option 'Candidat' selected.
- 2 Votre e-mail**: Fields for 'Votre e-mail:' containing 'jacdy@adinet.com' and 'Confirmez votre e-mail:' also containing 'jacdy@adinet.com'.
- 3 Votre mot de passe**: Fields for 'Votre mot de passe:' (containing '\*\*\*\*\*') and 'Confirmez votre mot de passe:' (containing '\*\*\*\*\*'). A note indicates '6 à 8 caractères sans espace'.
- 4 Votre CV**: A file upload field labeled 'Téléchargez votre CV :' with 'Parcourir...' and 'CV-JackDy-2014' selected. Notes say 'Formats acceptés : docx, doc, rtf, pdf' and '500 ko max.'. A link 'Vous n'avez pas de CV à télécharger' is provided. A checkbox 'J'accepte les conditions générales d'utilisation' is checked.

A large green 'Créer' button is located at the bottom right of the form.

# Exemple d'intégration (2/4)

Accueil > Mon Manpower > Espace candidats > Compléter mon profil - Étape 1

Profil complété à : 0 %

Compléter mon profil

Mes services

Gérer mes alertes e-mail

Gérer mes candidatures

Trouver une agence

Où ? (ex : Paris, Ile-de-France)

Tous les métiers ▾

Rechercher

## Compléter mon profil

Nous vous remercions d'avoir créé votre compte. Un e-mail vous indiquant vos identifiants vous a été adressé.

Afin que nos agences puissent évaluer au mieux votre candidature, nous vous invitons à compléter rapidement votre profil.

1. Mes coordonnées    2. Mes postes recherchés    3. Mes formations et compétences    4. Mes expériences    5. Mes CV

**MES COORDONNÉES**

Monsieur     Madame

Jack    Dy    09/10/1993

**Mon adresse**

N°    Type de voie    271 AV DAUMESNIL

Complément d'adresse    75012    PARIS

**Mes numéros de téléphone**  
Indiquez au moins un numéro de téléphone

0610566214    Tél. 2    Tél. 3

jacd@adinet.com.uy

**Ma disponibilité \***

Immédiate  
 À la date du jj/mm/aaaa  
 Avec un préavis de Choisissez ▾

Je bénéficie d'une reconnaissance Travailleur Handicapé [Consulter les engagements de Manpower](#)

# Exemple d'intégration (3/4)

Profil **complété à : 50 %**



**Compléter mon profil**

Mes services

Gérer mes alertes e-mail

Gérer mes candidatures

Trouver une agence

Où ? (ex : Paris, Ile-de-France)

Tous les métiers ▾

**Rechercher**

Compléter **mon profil**

Afin que nos agences puissent évaluer au mieux votre candidature, nous vous invitons à compléter rapidement votre profil.

1. Mes coordonnées    2. Mes postes recherchés    3. Mes formations et compétences    4. Mes expériences    5. Mes CV

Niveau d'études atteint

BAC

Formations - diplômes

2011  
: Études musicales au conservatoire de Créteil (94) en basse continue et en piano jazz  
: Bac TMD

il vous reste 1887 caractères (2000 car. maximum)

Mes connaissances linguistiques

Saisissez au moins 3 lettres de la langue et choisissez dans la liste puis sélectionnez votre niveau. (1 à 10 choix possibles).

| Mes langues : 4 sélectionnée(s) | Niveau         |                           |
|---------------------------------|----------------|---------------------------|
| ANGLAIS                         | BONNE MAITRISE | <a href="#">Supprimer</a> |
| RUSSE                           | NOTIONS        | <a href="#">Supprimer</a> |
| ESPAGNOL                        | COURANT        | <a href="#">Supprimer</a> |
| ALLEMAND                        | NOTIONS        | <a href="#">Supprimer</a> |

Mes connaissances linguistiques

Expérience

# Exemple d'intégration (4/4)

The screenshot shows a user interface for profile completion. On the left, there's a sidebar with navigation links: 'Profil complété à : 56 %' (Profile completed at 56%), 'Compléter mon profil' (Complete my profile), 'Mes services' (My services) with 'Gérer mes alertes e-mail' (Manage my email alerts) and 'Gérer mes candidatures' (Manage my applications), and 'Trouver une agence' (Find an agency) with a search bar ('Où ? (ex : Paris, Ile-de-France)', 'Tous les métiers' dropdown, and 'Rechercher' button).

The main content area is titled 'Compléter mon profil' (Complete my profile). It includes a message: 'Afin que nos agences puissent évaluer au mieux votre candidature, nous vous invitons à compléter rapidement votre profil.' (In order for our agencies to evaluate your application as best as possible, we invite you to quickly complete your profile.)

There are five tabs for completing the profile: '1. Mes coordonnées' (My coordinates), '2. Mes postes recherchés' (Jobs I am looking for), '3. Mes formations et compétences' (My training and skills), '4. Mes expériences' (My experiences), and '5. Mes CV' (My CV). The '4. Mes expériences' tab is active.

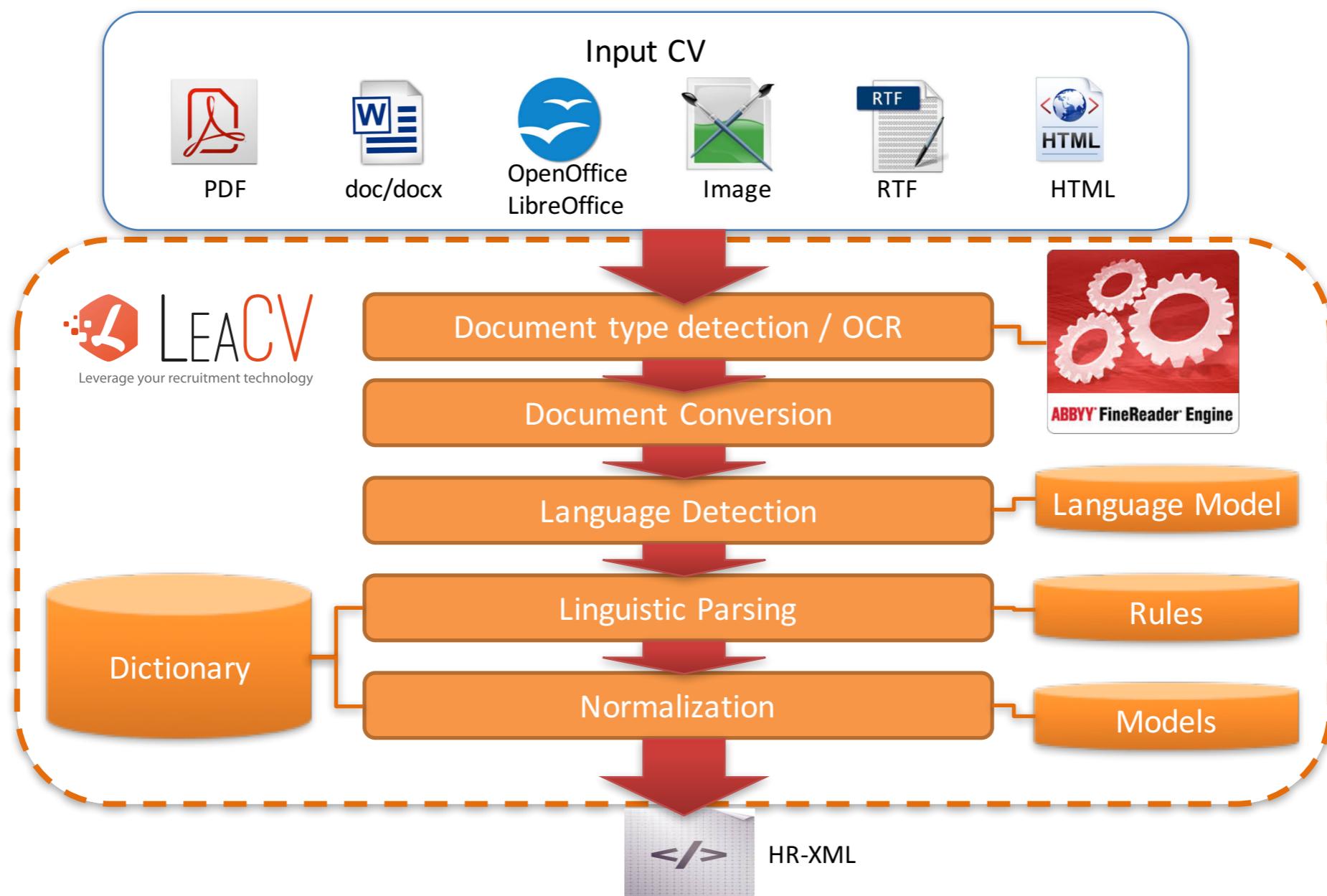
Under 'Mon niveau d'expérience' (My level of experience), a dropdown menu shows '1 an'. Below this is a section for 'Mes expériences professionnelles' (My professional experiences) with a table:

| Mes expériences       | Entreprise - Lieu | Contrat | Début   | Fin     | Modifier                 | Supprimer                 |
|-----------------------|-------------------|---------|---------|---------|--------------------------|---------------------------|
| Serveur au Bar        | Club Jazz -       |         | 08/2012 | 08/2013 | <a href="#">Modifier</a> | <a href="#">Supprimer</a> |
| Serveur au Restaurant | Boeuf Sous -      |         | 08/2011 | 08/2011 | <a href="#">Modifier</a> | <a href="#">Supprimer</a> |

A 'Nouvelle expérience' (New experience) button is located below the table.

At the bottom, there's a section for 'Mes expériences extra-professionnelles' (My extra-professional experiences) with a placeholder text: 'Précisez vos expériences associatives, d'animation, d'organisation d'événements ...' (Specify your associative, animating, event organization experiences ...).

# Une chaîne d'analyse de CV



Avant



## EXPERIENCES

- Juillet-Août 2014 : Réceptionniste trilingue au Camping Paris Est (Groupe Homair Vacances), Champigny (94).
- Août 2012 et 2013 : Serveur au Bar Club du festival JAZZ IN MARCIAC (32).
- Août 2011 : Bénévolat au Restaurant Le Boeuf Sous La Toile du festival JAZZ IN MARCIAC (32). Tâches diverses en cuisine et salle.
- Juillet 2011: Stage régie son au festival LES SUDS à Arles (13).
- 2013-2014: Groupes musicaux NO MAD et WAK: participation à des festivals et concerts dans des salles culturelles.
- 2006-2012: Groupe musical The Chat Spleen : participation à des festivals et concerts dans des salles culturelles ainsi qu'au Gibus et au New Morning (75) et au festival Jazz In Marciac (32).

## ETUDES

- 2014 : Prépa DEM Jazz, EDIM (Ecole de musiques actuelles), Cachan (94).
- 2012-2013 : Etudes de Sciences pour l'ingénieur (parcours acoustique), Université du Maine (72).
- 2011 : Obtention du Bac TMD avec mention Assez bien.
- 2009-2011: Etudes musicales au conservatoire de Créteil (94) en basse continue et en piano jazz.
- 2004-2009 : Etudes musicales à l'école de musique de Joinville-le-Pont (94) en clavescin.

## LANGUES

- Bilingue Français-Espagnol.
- Anglais : Très fluide oral et écrit.
- Italien et Allemand : Bonne compréhension.
- Russe : Notions.

## COMPETENCES INFORMATIQUES

- Logiciels : Word, Excel, Cubase 7, Magix Music Maker, Octave.

## DIVERS

- Association des Orchestres Universitaires du Maine (AOUM).
- Voyages : Séjours en Amérique du Sud (Uruguay et Brésil).

# Après

```
<EmployerOrg>
  <EmployerOrgName>Camping Paris Est</EmployerOrgName>
  <PositionHistory positionType="contract">
    <Title>Réceptionniste</Title>
    <OrgName>
      <OrganizationName>Camping Paris Est</OrganizationName>
    </OrgName>
    <OrgInfo>
      <LocationSummary>
        <Municipality>Champigny</Municipality>
      </LocationSummary>
    </OrgInfo>
    <Description>Juillet-Août 2014 : Réceptionniste trilingue au Camping Paris Est (Groupe Homair Vacances), Champigny (94).</Description>
    <StartDate>
      <YearMonth>2014-07</YearMonth>
    </StartDate>
    <EndDate>
      <YearMonth>2014-08</YearMonth>
    </EndDate>
    <UserArea>
      <lea:Category>
      </lea:Category>
      <lea:WorkingTime>unknown</lea:WorkingTime>
    </UserArea>
  </PositionHistory>
</EmployerOrg>
<EmployerOrg>
  <EmployerOrgName>Club du festival JAZZ IN MARCIAC</EmployerOrgName>
  <PositionHistory positionType="contract">
    <Title>Serviteur au Bar</Title>
    <OrgName>
      <OrganizationName>Club du festival JAZZ IN MARCIAC</OrganizationName>
    </OrgName>
    <OrgInfo>
      <LocationSummary>
        <Municipality>MARCIAC</Municipality>
      </LocationSummary>
    </OrgInfo>
```

# Variété des formats

---

- Comme (pratiquement) toujours dans une industrialisation du TAL, le premier problème est celui du format d'entrée des documents
  - PDF (dans toutes ses versions, croisées avec les différents générateurs) - dans la majorité des cas
    - associé au mythe du « format non-modifiable »
  - Image : numérisation de CV papiers (cf. salons)
  - Office (Word, mais aussi Powerpoint) dans ses différentes déclinaisons (doc vs. docx, etc.)
  - Open Office
  - HTML (un peu)

# Exemple de problème « presque TAL » : les tableaux

---

- L'utilisation des tableaux (au sens large) est une forme très usitée de mise en page des CV
- Mais tous les tableaux n'ont pas la même signification
- Un outil de conversion en texte devrait reconnaître la sémantique des colonnes du tableau pour fournir des données « propres » à l'analyseur linguistique
  - sinon : « garbage in, garbage out »
- Mais on est en amont du traitement, donc avant interprétation, avant même détection de la langue !

# Un exemple réel de colonnes à interpréter

---

## EXPERIENCES

**Linguiste informaticienne** - Stage- Laboratoire d'informatique Etis-Cergy-Pontoise- Avril-Juillet2015

- analyse sémantique et syntaxique d'un corpus de tweets de politiques français extrait de Twitter
- mise en place d'un outil de détection de la négation à base de règles linguistiques.

**Traduction et interprétariat**- Missions- (arabe/français, français/arabe)- Vidéo Adapt, Label Vidéo, Amnesty International, Passerelles- Paris - 2011/2014

- Traduction de documents, sous-titrage de films, et traduction consécutive lors de conférences de presse

**Rédaction de la méthode d'initiation à langue arabe pour les francophones**

« *Démarrez l'arabe - spécial vacances !* » -Larousse, Paris - Novembre 2012/Mars 2013

- Rédaction en français de textes sur la culture et la langue arabes
- Conception d'activités d'apprentissage ludiques pour faciliter l'apprentissage.

## COMPETENCES

Langages de Programmation :  
Python, Perl, PHP, HTML &CSS

Outils de Traitement Automatique :  
Unitex, Treetagger,  
Expressions régulières

Systèmes d'exploitation :  
Windows 7, Ubuntu

# D'autres colonnes, sémantiques variables

## EXPERIENCE PROFESSIONNELLE

Aout 2015 :

Stagiaire chez la société SAM-BI - Ariana, Tunisie

**Sujet :** Développement d'une application mobile la communication entre les dentistes et les laboratoires en Allemagne.

**Technologie :** Android, MySQL, SCRUM.

Juin – Juillet  
2014 :

Stagiaire chez la société Yasmine Engineering System – Ariana, Tunisie

**Sujet 1:** Conception et développement d'une interface web pour les projets immobiliers.

**Technologies :** Sencha Extjs, .Net, SQL Server, SCRUM, Tortoise SVN.

**Sujet 2:** Conception et développement d'un module d'étude de faisabilité d'un projet

**Technologies:** .Net, SQL Server, SCRUM, Tortoise SVN.

## PROJETS ET REALISATIONS

- **Académiques**

- Conception et développement d'une application place de marché (Asp.net MVC, C#, SQL Server, Responsive design)
- Conception et Réalisation d'un workflow (IBM Workflow)

- **Associatifs**

- Réalisation d'un rappel de prise des médicaments (Android)
- Réalisation d'un jeu 2D (HTML5/CSS3, JavaScript(JQuery))
- Réalisation d'un réveil intelligent (Android)

## LOGICIELS / LANGUAGES DE PROGRAMMATION

- **Web front-end:** HTML5/CSS3, EXTJS.
- **Mobile:** Android.
- **JEE:** JSF (PrimeFaces), EJB 3.0, spring, JPA, Hibernate.
- **.NET:** ASP MVC x, Unity, Entity Framework, WCF, WPF.
- **Bases de données:** MySQL, SQL Server.

## LANGUES

- **Arabe :** Langue maternelle.
- **Français :** Courant.
- **Anglais :** Technique.

# Une autre contrainte « presque TAL » : le temps de traitement

---

- Le délai communément admis pour une réponse d'un système est d'une seconde
  - au delà, l'utilisateur pense que le système n'a « pas compris » et généralement relance sa requête
  - ce qui aggrave le problème
- Il faut donc qu'en moins d'une seconde, on puisse traiter
  - l'envoi vers le site, puis le serveur
  - la conversion de document (avec OCR si besoin)
  - l'analyse
  - le retour de la réponse depuis le serveur et l'affichage

# L'extraction d'informations depuis un CV

---

- On sait ce qu'on cherche, mais on ne sait pas forcément où le trouver
  - par exemple, dans un CV, on s'attend à trouver un nom, généralement vers le début du texte
  - un lexique de prénoms peut grandement aider
- On cherche essentiellement à extraire :
  - l'état-civil (nom, prénom, adresse, e-mail, téléphone) et les permis/ habilitations
  - la liste des expériences professionnelles (postes occupés, employeur)
  - la formation (importance inversement proportionnelle à la durée de l'expérience professionnelle)
  - parfois : publications, références.

# Grands principes techniques

---

- L'extraction d'information consiste à repérer des éléments « raisonnablement sûrs » pour délimiter des zones dans lesquelles on va chercher des informations plus précises
  - en affinant la recherche avec des éléments potentiellement plus dangereux
  - cf. on peut se permettre de considérer « Le » comme prénom si on est certain d'être dans une zone d'état-civil
- Utilisation abondante de grammaires locales couplées à des dictionnaires et/ou des techniques d'apprentissage pour identifier ces éléments
  - Techniques similaires à l'extraction d'entités
  - On n'a généralement pas besoin de POS-tagging ou autres traitements standards
- Ensuite, on « rassemble les morceaux » par des outils qui tiennent plus du « raisonneur » que de l'extracteur

## EXPERIENCES

- Juillet-Août 2014 : Réceptionniste trilingue au Camping Paris Est (Groupe Homair [redacted], Chilly-Mazarin (94)).
- Août 2012 et 2013 : Serveur au Bar Club du festival JAZZ IN MARCIAC (32).
- Août 2011 : Bénévolat au Restaurant Le Boeuf Sous La Toile du festival JAZZ IN MARCIAC (32). Tâches diverses en cuisine et salle.
- Juillet 2011 : Stage régie son au festival LES SUDS à Arles (13).
- 2013-2014 : Groupes musicaux NO MAD et WAK: participation à des festivals et concerts dans des salles culturelles.
- 2006-2012 : Groupe musical The Chat Spleen : participation à des festivals et concerts dans des salles culturelles ainsi qu'au Gibus et au New Morning (75) et au festival Jazz In Marciac (32).

## ETUDES

- 2014 : Prépa D'IM Jazz, EDIM (Ecole de musiques actuelles), Cachan (94).
- 2012-2013 : Etudes de Sciences pour l'ingénieur (parcours acoustique), Université du Maine (72).
- 2011 : Obtention du Bac TND avec mention Assez bien.
- 2009-2011 : Etudes musicales au conservatoire de Créteil (94) en basse continue et en piano jazz.
- 2004-2009 : Etudes musicales à l'école de musique de Joinville-le-Pont (94) en clavecin.

## LANGUES

- Bilingue Français-Espagnol.
- Anglais : Très fluide oral et écrit.
- Italien et Allemand : Bonne compréhension.
- Russe : Notions.

## COMPETENCES INFORMATIQUES

- Logiciels : Word, Excel, Cubase 7, Magix Music Maker, Octave.

## DIVERS

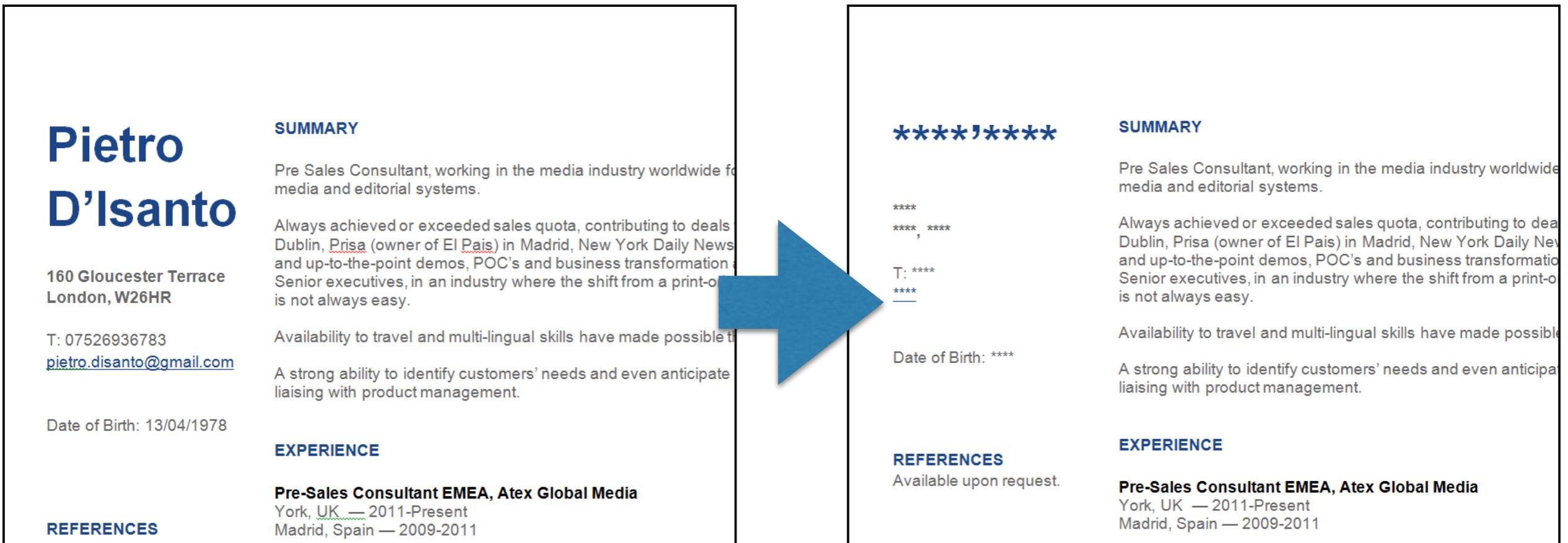
- Association des Orchestres Universitaires du Maine (AOUM).
- Voyages : Séjours en Amérique du Sud (Uruguay et Brésil).

# Qualité des systèmes

---

- Les systèmes fonctionnent plutôt bien: environ 98% de bonne reconnaissance de l'état-civil (nom, prénom, adresse, téléphone)
  - les problèmes viennent essentiellement des mises en forme « non standard »
- Environ 80% de bonne reconnaissance de l'ensemble de l'expérience (période, fonction, employeur)
  - problèmes de « frontière droite »
  - cf. « *2014-2015 serveur au bar chez Mimile à Pont-à-Marcq* »
    - (serveur au bar) chez Mimile
    - serveur au (bar chez Mimile)
    - Si on remplace « chez Mimile » par « chez Michou », il n'y a plus d'ambiguïté
  - Problème : comment connaître l'intégralité des dénominations alors qu'on ne connaît pas le domaine a priori (et qu'on ne le connaîtra certainement jamais !)
  - On revient à la notion de référentiel qui recenserait l'ensemble des entités potentiellement présentes dans un CV

# Un produit dérivé : l'anonymisation des CV



# L'analyse d'offres d'emploi

---

Les outils TAL pour l'emploi

# Pourquoi analyser les offres d'emploi ?

---

- Pour aider au matching (voir plus loin)
- Pour éviter les mentions discriminatoires
  - On constate qu'environ 10% des offres ont des mentions discriminatoires (sexe, âge, origine)
    - *XXX recherche un(e) assistant(e) (H/F). Elle aura pour mission de ...*
    - Les types de discrimination ont été codifiés par la Halde (reprise par le Défenseur des Droits)
    - Officiellement, les sanctions encourues pour une société :
      - 225.000€ d'amende
      - interdiction d'exercer l'activité
      - à titre individuel :
        - jusqu'à 3 ans d'emprisonnement
  - Pour faire en sorte qu'elles incitent les candidats à postuler

# Exemple

## Secrétaire commercial(e)

Au sein d'une PME, vous serez chargé(e) de l'accueil physique et téléphonique, de la gestion du courrier, de la gestion des factures et devis, et autres travaux administratifs.

Vous avez un fort sens commercial pour intervenir sur le showroom en place, accueil, orientation et conseil clients.

Le profil :

Diplômé(e) d'un BTS Assistant de gestion PME-PMI ou Assistant de Direction, vous avez une expérience de 2 années minimum en tant qu'**assistante** administrative en PME, avez de solides connaissances du pack office (Word, Excel, internet et messagerie électronique).

Vous êtes dynamique, et vous avez une bonne élocution. Votre **langue maternelle est le français**, mais la maîtrise de l'anglais est indispensable.

Merci d'envoyer CV et lettre de motivation à  
contact@mycompany.fr

## Conformité juridique

|                     |    |
|---------------------|----|
| Âge                 | OK |
| Expérience maximale | OK |
| Sexe                |    |

|                          |    |
|--------------------------|----|
| Origine géographique     | OK |
| <b>Langue maternelle</b> |    |
| Apparence physique       | OK |
| Durée du contrat         | OK |

En orange : critère non conforme

## Formateur de français langue étrangère H/F

Ecole de langues cherche formateur en FLE cours individuel en face à face.

Les élèves sont de langue maternelle asiatique.

Vous savez vous adapter aux différents élèves, vous êtes pédagogue.

Qualifications et expérience de formateur(trice) en FLE demandées.

Si vous êtes intéressé(e) par cette mission de formation, veuillez nous adresser votre candidature (CV détaillé) par mail à l'adresse suivante

[fle@enseignement-langues.com](mailto:fle@enseignement-langues.com)

# Pas toujours simple

La prise en compte du contexte est nécessaire pour éviter les « faux positifs »

# Marché

---

- Dans les faits, peu de sociétés recourent à ce type d'outils
  - ~~flans... fondations pour la communication interne~~
- **Profil recherché**
  - Vous êtes une personne rigoureuse, bien organisée, exigeante avec elle-même.
  - Vous êtes motivé(e) et vous avez envie d'évoluer dans un univers ultra dynamique.
  - Si possible pas juif(ve).
  - Vous maîtrisez parfaitement la chaîne graphique et les techniques d'imprimerie.
  - Compréhension rapide et globale des projets.
  - Proactif(ve) et autonome, vous n'hésitez pas à chercher l'information par vous-même, avant d'avoir recours à l'équipe, mais savez vous faire aider lorsque cela est nécessaire.
  - Sens du travail en équipe et des responsabilités.
  - Un esprit méthodique : rigueur, réactivité et capacité à gérer les priorités.
  - Expérience de 2 à 4 ans exigée.
- Mais quelques grandes sociétés (Verona, ...) ou amuseurs institutionnels (Apec) mettent en œuvre de tels systèmes

# Le matching

---

Les outils TAL pour l'emploi

# La problématique

---

- Faire correspondre un CV à une offre
  - pour recruter efficacement quand on est recruteur
  - pour « pousser » des offres à des candidats de la part d'un site d'emploi
- Nécessite une base de profils (« CV-thèque »)

# Un profil est plus qu'un CV

---

- De nombreuses informations importantes pour une recherche d'emploi ne sont pas dans le CV
  - salaire recherché
  - zone géographique
  - contrat recherché (CDI, CDD, Interim, stage, ...)
  - poste recherché par le candidat

# Le processus de matching

---

- Faire correspondre les critères factuels
  - lieu, type de contrat, ...
  - on peut avoir un peu de souplesse (cf. proposer un CDD plutôt qu'un CDI si tous les autres critères matchent)
- et mettre en correspondance les critères d'une offre avec ceux d'un CV
  - métier, compétence, expérience, ...
  - sachant que ces éléments ne sont pas toujours exprimés avec les mêmes mots
  - c'est ici que le TAL intervient principalement

# Les problématiques

---

- Trouver les compétences requises dans l'offre, les distinguer de celles qui sont souhaitables
  - *la connaissance de l'environnement java est requise, la maîtrise du framework spring est un plus*
  - induire les informations qui permettent de faire correspondre des pré-requis
    - *... recherche développeur C++ banque*
    - « matche » avec
    - *... compétences : C++, Calypso ...*

# Cela nécessite

---

- Un bon analyseur de CV pour extraire les compétences et les expériences professionnelles avec leur date
  - une expérience de C++ il y a 10 ans n'a pas grand intérêt
- Un bon analyseur d'offres pour extraire les compétences/expériences requises et distinguer celles qui sont nécessaires des autres
  - domaine d'expérience, qualifications demandées (permis, CACES, par exemple)
- Un modèle permettant de faire correspondre des connaissances
  - Zend => php
  - Unitex => TAL
- Un bon algorithme pour gérer tous ces critères
  - L'anglais est-il plus ou moins important qu'une expérience réussie dans le domaine de la banque pour un informaticien ?

# Techniques mobilisables

---

- Les référentiels fournissent des compétences/conditions d'accès à l'emploi (métiers, formations)
  - mais elles ne sont que partielles
- L'analyse de grands volumes permet d'identifier des contextes pour extraire des données (par apprentissage ou par grammaires locales)
  - *La connaissance de <NP> est requise*
- Des techniques de « word-embeddings » peuvent aider à créer des liens entre compétences/domaines, etc.
  - à partir des offres, mais aussi des CV

# GPEC, trajectoires, évolutions

---

- On peut déduire des évolutions potentielles en étudiant les suites de postes dans les CV
- On peut ainsi prévoir des évolutions pour un profil à l'intérieur d'une société
  - ou lui proposer des scénarios d'évolution
- Cela demande de formaliser les métiers
  - Classification / référentiel
  - Travaux essentiellement de R&D sur le sujet (technologie pas mûre d'un point de vue industriel)

# Conclusion

---

# Un formidable terrain de jeu pour le TAL

---

- De nombreux problèmes non-résolus
- Des enjeux sociaux
- Des données à foison (presque big-data)
- Beaucoup d'outils mobilisables