

# NaviQuest : un outil pour naviguer dans une base de questions posées à un Agent Conversationnel

Conférence WACA 2010  
25 - 26 novembre  
Lille

Ph. Suignard  
EDF R&D



Posez une autre question

OK

26 / 11 / 2010



CHANGER L'ÉNERGIE ENSEMBLE

# Plan de la présentation

- ▶ Des « nouvelles » de Laura !
- ▶ Contexte et problématique de cette application et cette étude
- ▶ Présentation du logiciel Naviquest
- ▶ Présentation de quelques résultats
- ▶ Conclusion et perspectives

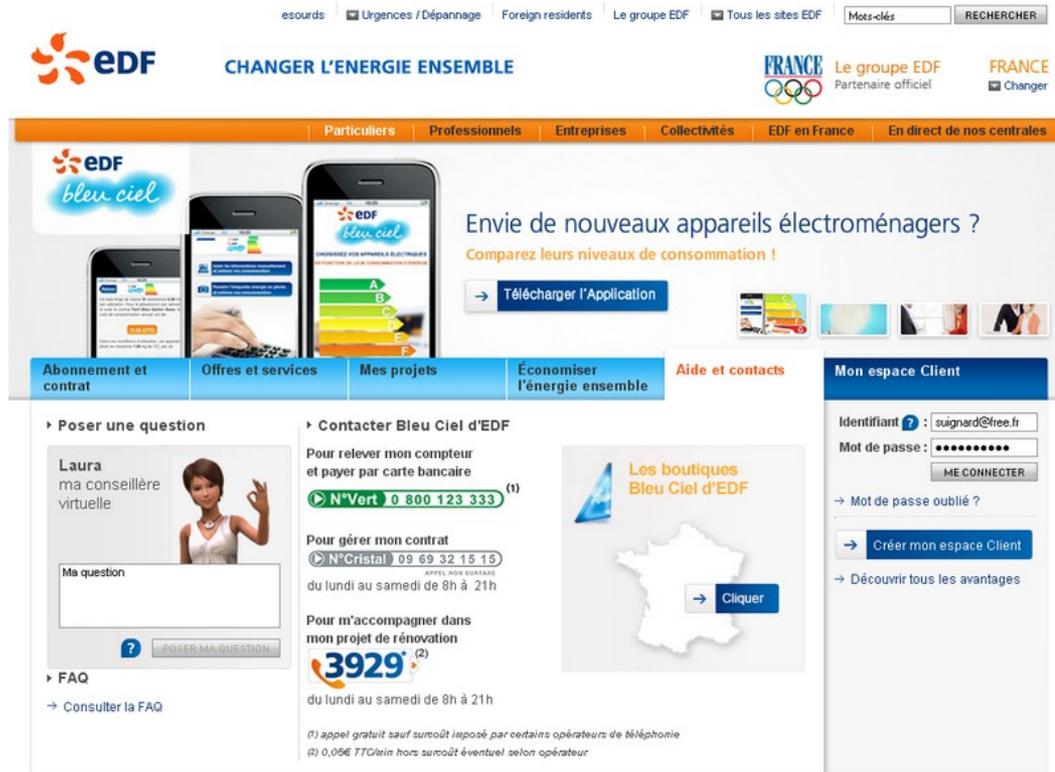


Posez une autre question



OK

# Des « nouvelles » de Laura !



Posez une autre question



- ▶ Laura est la conseillère virtuelle du site EDF Bleu Ciel
- ▶ Répond aux questions des internautes 24h/24h et 7j/7 !
- ▶ Présente depuis plus de 5 ans sur le site dans 3 univers différents
- ▶ Réalisée en partenariat avec les sociétés Cantoche et Sémantia
- ▶ Répond à environ 10 000 questions par jour
- ▶ Base de connaissance mise à jour régulièrement
- ▶ Suivi mensuel par la Direction Marketing

# Contexte et problématique de cette étude



Posez une autre question

## ▶ EDF cherche constamment à améliorer Laura :

- Augmenter la pertinence de ses réponses par rapport aux questions posées
- Augmenter son périmètre
- Augmenter la satisfaction des clients
- Rapports mensuels Sémantia, mais souhait d'aller plus loin

## ▶ Lancement d'une étude pour :

- 1. Connaître plus finement les questions posées par les Internautes
  - Longueur des questions (courtes/longues)
  - Style (familier ou non), manière de s'adresser à Laura (tutoiement/vouvoiement)
  - Longueur des dialogues
  - Détection de nouveautés
  - Sujets de conversation
- 2. Connaître plus finement les réponses fournies par Laura
  - Incompréhensions
  - Trouver plus finement quelles questions amène à une réponse particulière
- 3. Détecter les signaux faibles :
  - Non imaginé au départ
  - Crowdsourcing, profiter de ce corpus pour extraire des informations pertinentes

## ▶ Tout cela dans une optique d'amélioration

- Adapter les dialogues au contexte par exemple

# Présentation du logiciel Naviquest



- ▶ Réaliser un outil mixant des techniques de visualisation et d'analyse textuelle (préoccupation actuelle)
  - IVITA - premier workshop sur le thème des « interfaces visuelles intelligentes pour le texte »
  - Session « Text Visualization » des conférences InfoVis
  - Session « Text Analytics » des conférences Vast
- ▶ Basé sur un logiciel développé dans le cadre du projet Infomagic/Callsurf
  - Callsurf : fouille dans des conversations téléphoniques automatiquement retranscrites entre client et conseillers
  - Conversation téléphonique = suite de tours de parole entre client et conseiller
  - Dialogue avec un ACA = suite de questions réponses
  - Techniquement similaire, donc réutilisation possible du logiciel
  - Logiciel basé sur des technologies Open Source : Lucene (moteur de recherche), JFreeChart (graphes), JTreeMap (tree map).
- ▶ Deux modules
  - Indexation (questions)
  - Fouille et recherche

# Présentation du logiciel Navigest



Barre de saisie pour  
rechercher un ou  
plusieurs mots

Liste des questions  
posées à Laura  
contenant le ou les mots  
cherchés

Ventilation  
des  
questions  
selon les  
méta-  
données

Conversation Viewer

Mots :  Conversation :

Page : 1/17

**Corpus total**

**Répartition des réponses par :**

Thèmes	Occurences	%
Question	169++	100
Réponse	0	0

Jour	Occurences	%
Lundi	57++	33
Dimanche	45+++	26
Mercredi	24	14
Mardi	20-	11
Jeudi	15--	8
Vendredi	4	2
Samedi	4	2

Semaine	Occurences	%
200953	1	0
201001	4	2
201002	1	0
201003	3	1
201004	1	0
201005	1	0
201006	1	0
201007	1	0
201008	40++++	23
201009	85++++	50
201010	12	7
201011	9	5

169 réponses (segments). [Exporter la liste](#)

334464 - 2,405 - Question - 5 ème segment  
coupure de courant du a la **tempete**

342385 - 2,405 - Question - 3 ème segment  
retablissement de la ligne après **tempete**

346701 - 2,405 - Question - 1 ème segment  
**tempete** arbre couché sur ligne Electrique

367919 - 2,405 - Question - 1 ème segment  
probleme de lumiere suite a la **tempete**

395321 - 2,405 - Question - 1 ème segment  
peut on avoir une indemnité pour une **tempete**

333973 - 2,004 - Question - 3 ème segment  
quand l' électricité sera rétabli chez moi suite a la **tempete**

bonjour - cas\_de\_tempête - cause - combien\_de\_temps - Compteur - compteur\_edf\_extérieur -  
compteur\_electrique - coupure - **coupure\_de\_courant** - courant - edf -  
edf\_sont\_arrachées - envolé\_avec\_tempête - ligne - lirez\_de\_propriété - ménagers\_et\_hi - nuit -  
porte - porte\_de\_coffret - porte\_du\_compteur - portes\_des\_armoires - reparation\_suite\_tempete -  
**tempete** - tempete\_de\_dimanche - tempete\_plusieurs\_porte -  
**tempête** - tempête\_xynthia - vis\_de\_fermeture - électricité - électrique



OK

# Présentation du logiciel Navigest

## ► Fonctionnalités

- Recherche « Full » texte, recherche « floue »
- Export / Extraction
- Recherche avancée : liste de mots obligatoires, facultatifs et exclus
- Ventilation des résultats selon les méta-données
- Raffinement des recherches
- Analyses temporelles
- Nuage de mots
- Détection des nouveautés et évolutions (unigramme, bigramme ou trigramme)
- Visualisation sous la forme de Tree Map

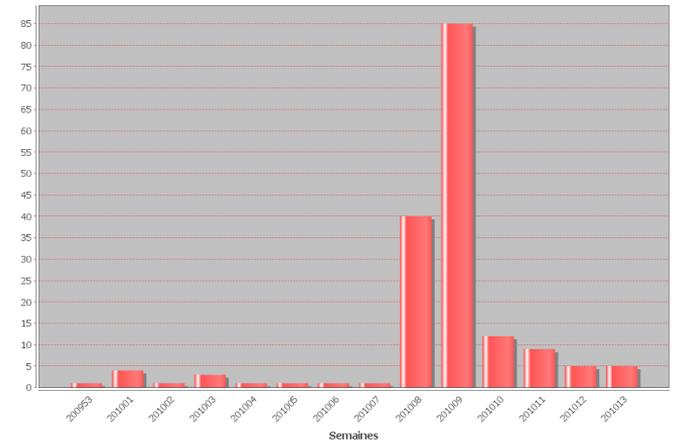
Semaine	Occurrences	%
200953	1	0
201001	4	2
201002	1	0
201003	3	1
201004	1	0
201005	1	0
201006	1	0
201007	1	0
201008	40++++	23
201009	85++++	50
201010	12	7
201011	10	5
201012	2	2
201013	5	2



Posez une autre question



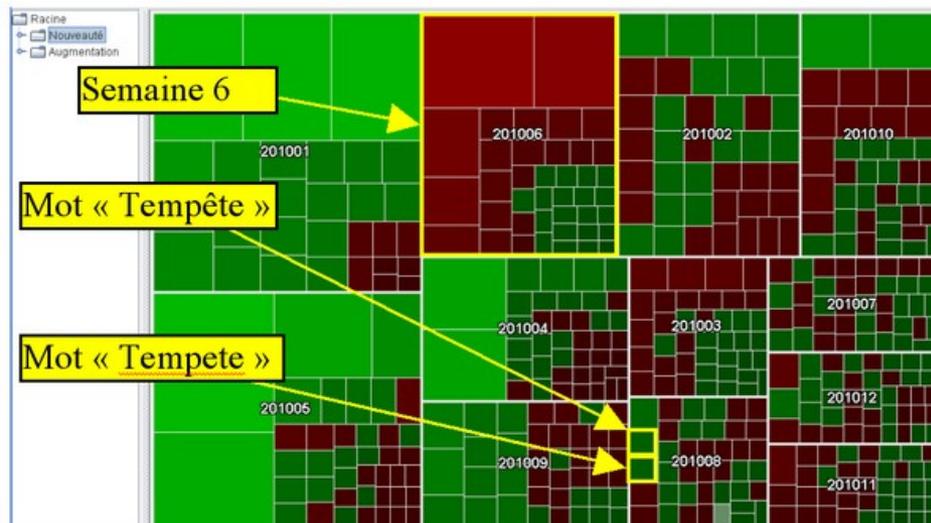
OK



bonjour - cas\_de\_tempête - cause - combien\_de\_temps - compteur - compteur\_edf\_extérieur - compteur\_électrique - coupure - coupure\_de\_courant - courant - edf - edf\_sont\_arrachées - envolé\_avec\_tempête - ligne - limite\_de\_propriété - ménagers\_et\_hi - nuit - porte - porte\_de\_coffret - porte\_du\_compteur - portes\_des\_armoires - réparation\_suite\_tempete -

**tempete** - tempete\_de\_dimanche - tempete\_plusieurs\_porte -

**tempête** - tempête\_xynthia - vis\_de\_fermeture - électricité - électrique



# Présentation de quelques résultats



Posez une autre question

- ▶ Lundi est le jour où il y a le plus de questions
- ▶ Certaines questions très longues sont en fait des mails :
  - On pourrait les rediriger vers les « vrais » mails
- ▶ Les évolutions de vocabulaire permettent de détecter des mots :
  - « tempête » Xinthia
  - Campagne publicitaire avec « Pitt » : récupération des verbatims
- ▶ Détection des réappels :
  - Préoccupation des centres d'appel : les clients qui appellent, puis qui ré-appellent
  - Diversité des canaux : ne doit pas augmenter le nombre total des contacts
  - Déport de canaux et Transferts de canaux
  - Travaux développés pour Callsurf testés sur Laura
  - La trajectoire client=>Web => Laura => conseiller est logique mais paradoxalement la réciproque existe, client => conseiller => Laura
- ▶ Indexation des réponses

# Conclusion et perspectives



Posez une autre question

- ▶ Tous ces verbatims constituent une mine d'or !
- ▶ Travail avec Direction Marketing
  - Intéressé par le logiciel suite à présentation
  - Ajout de fonctionnalités
  - Souhaite tester / utiliser le logiciel en 2011
  - Leur retour permettra d'améliorer le logiciel et surtout Laura !
- ▶ Recherche Stagiaire
  - Pour 2011
  - « Fouille dans un corpus de questions posées à un agent conversationnel »
  - Comparaison questions posées en 2010 et 2011.
  
- ▶ Merci pour votre attention...!