

# **Un chatbot pour les bibliothèques : exemple d'une collaboration entre BU et labo**

# L'environnement du projet

Coordination d'un réseau  
de 30 bibliothèques

4 bibliothèques intégrées,  
3 pôles transverses



Un (futur) Learning Center  
avec 2 établissements  
composantes et 3  
directions

# Le laboratoire



Algorithmes,  
Apprentissage et  
Calcul

Interactions avec  
l'Humain

Mécanique  
Énergétique

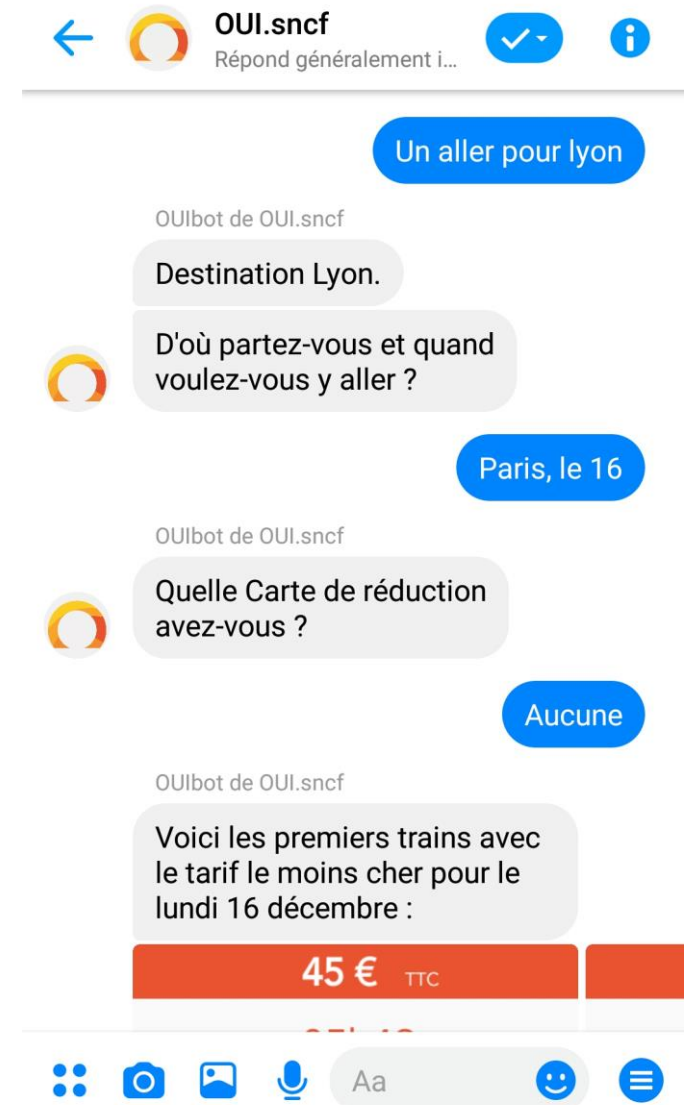
Sciences des  
Données

Sciences et  
Technologies des  
Langues

# Un chatbot : pourquoi ? Et comment ?

# Qu'est-ce qu'un chatbot ?

- Un **programme informatique** qui peut converser avec un humain.
- « chatbot » **terme industriel**, couvre une variété de systèmes différents : système de question-réponse, système de dialogue orienté tâche, système de dialogue conversationnel.
- Généralement **orienté tâche** : donner des informations spécifiques, réaliser une tâche.



OUIBot, le chatbot de OUI.sncf

# Pourquoi un chatbot des bibliothèques ?



Proposer un service en phase avec les usages de nos lecteurs mais pas que...

# Un exemple de valorisation : la fête de la science 2020

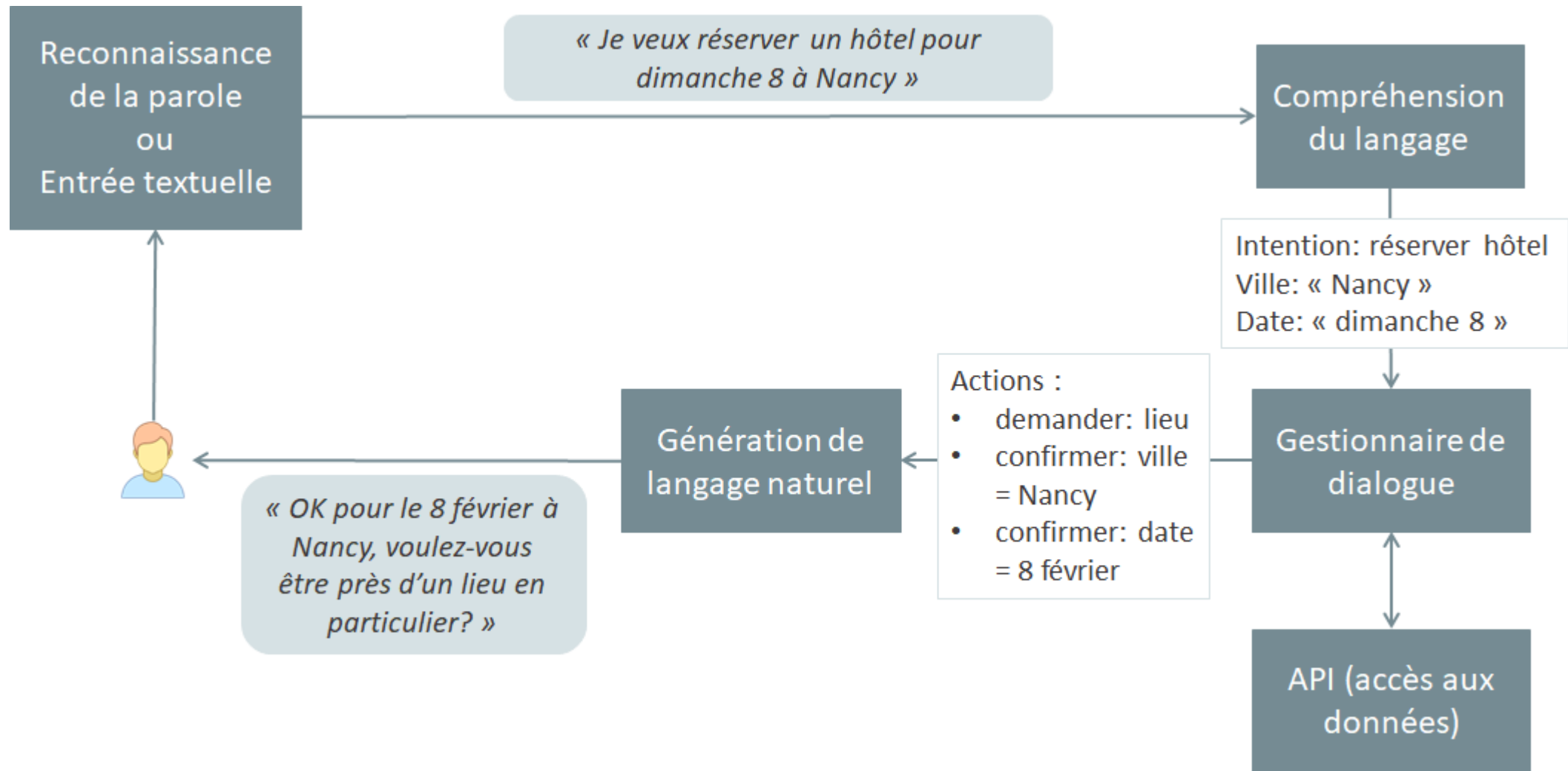




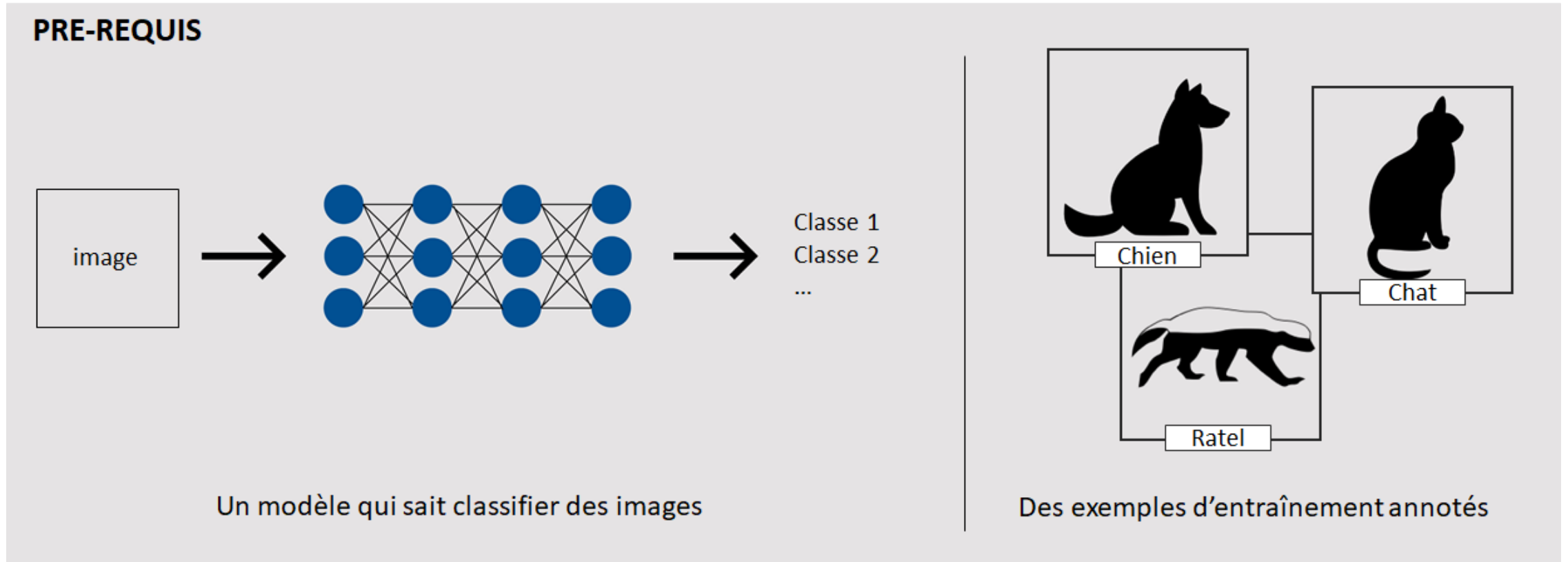
# Les chatbots, toujours de la recherche ?

- Présence de chatbots sur une multitude de plateformes.
- Mais certaines limitations :
  - Discrimination (e.g. biais de genre)
  - Fiabilité des réponses
  - Beaucoup d'ingénierie
  - Difficulté à adapter hors de son domaine d'application
- **Avantage d'une collaboration avec la DiBISO :**
  - Cas d'étude réel, présence d'experts du domaine pour l'**annotation de données** (Deep learning) et d'utilisateurs finaux pour l'évaluation.

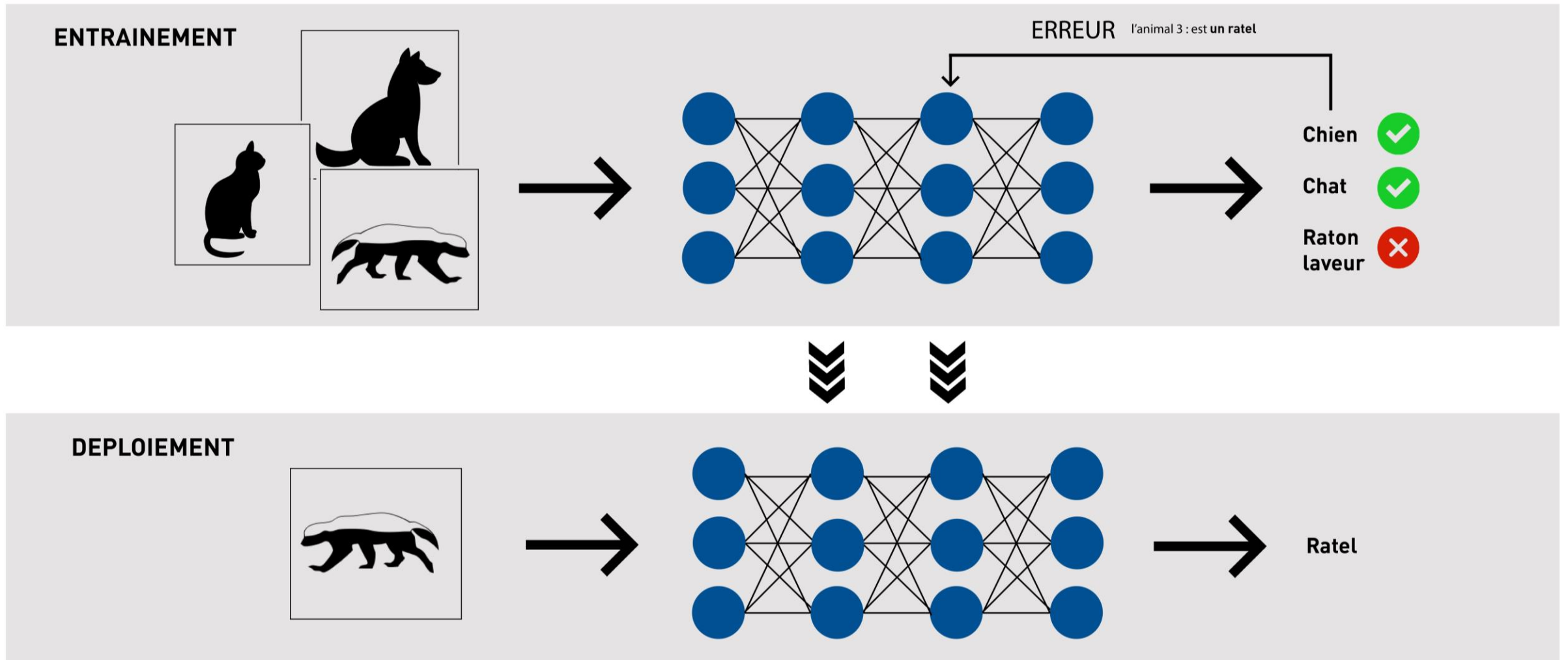
# Systeme de dialogue orienté tâche



# Comment on entraîne un modèle de DL ? (1/2)



# Comment on entraîne un modèle de DL ? (2/2)



source : <http://carrementbechtle.com>

# Construction de données d'entraînement

## Plusieurs étapes : ( ! ) Qualité

1. Rédaction du guide d'annotation
2. Collecte des données
3. Filtrage et nettoyage des données
4. Annotations des données
5. Vérification et correction des annotations
6. Normalisation des données

# Le choix des données d'Ubib

**Découvrez les établissements membres du réseau**

sélectionnez une ville



Bienvenue sur le chat

Ubib !



Posez votre question et appuyez sur Envoyez pour démarrer un chat.

Saisir votre message ici...

Envoyer

# Données Ubib et classes

876 « Sans objet »

692 « Informations pratiques »

467 « Aide à la recherche »

216 « Test »

204 « Problèmes d'accès »

136 « Prêt »

122 « Autres »

66 sans label

2 « Question en doublon »

**Total : 2781**

# La classification des données en thématiques

Données annotées = paire (question, thématique)

- Combien de temps puis-je garder un roman emprunté ?
- Combien d’ouvrages puis-je emprunter en même temps ?
- Vous êtes ouverts demain ?
- A quelle heure ferme la bibliothèque ce soir ?
- Puis-je venir travailler dans une salle de travail en groupe demain ?
- Avez-vous des box à la BU ?

**Thématique :**  
**Prêt**

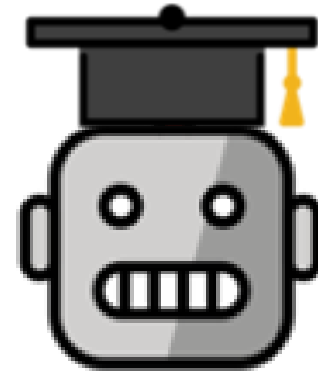
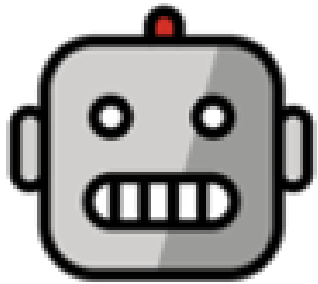
**Thématique :**  
**Horaires**

**Thématique :**  
**Réservation  
salle**

==> **1477 questions annotées** en thématique (**53%** des données brutes)



# Le choix d'un modèle



Les données collectés puis annotées permettent d'entraîner un modèle déjà existant pour aboutir à un prototype

# Premiers tests du prototype par des bibliothécaires

# Capture d'écran du prototype du chatbot utilisé pour le test

Page du Chatbot

Bonjour, je suis un chatbot. Posez-moi une question et je vais tenter de définir le thème de la question.

Comment faire pour emprunter un livre?

pret

Je souhaite connaitre les horaires?

horaire

Quels sont les modalités pour réserver une salle

droits d'entrée

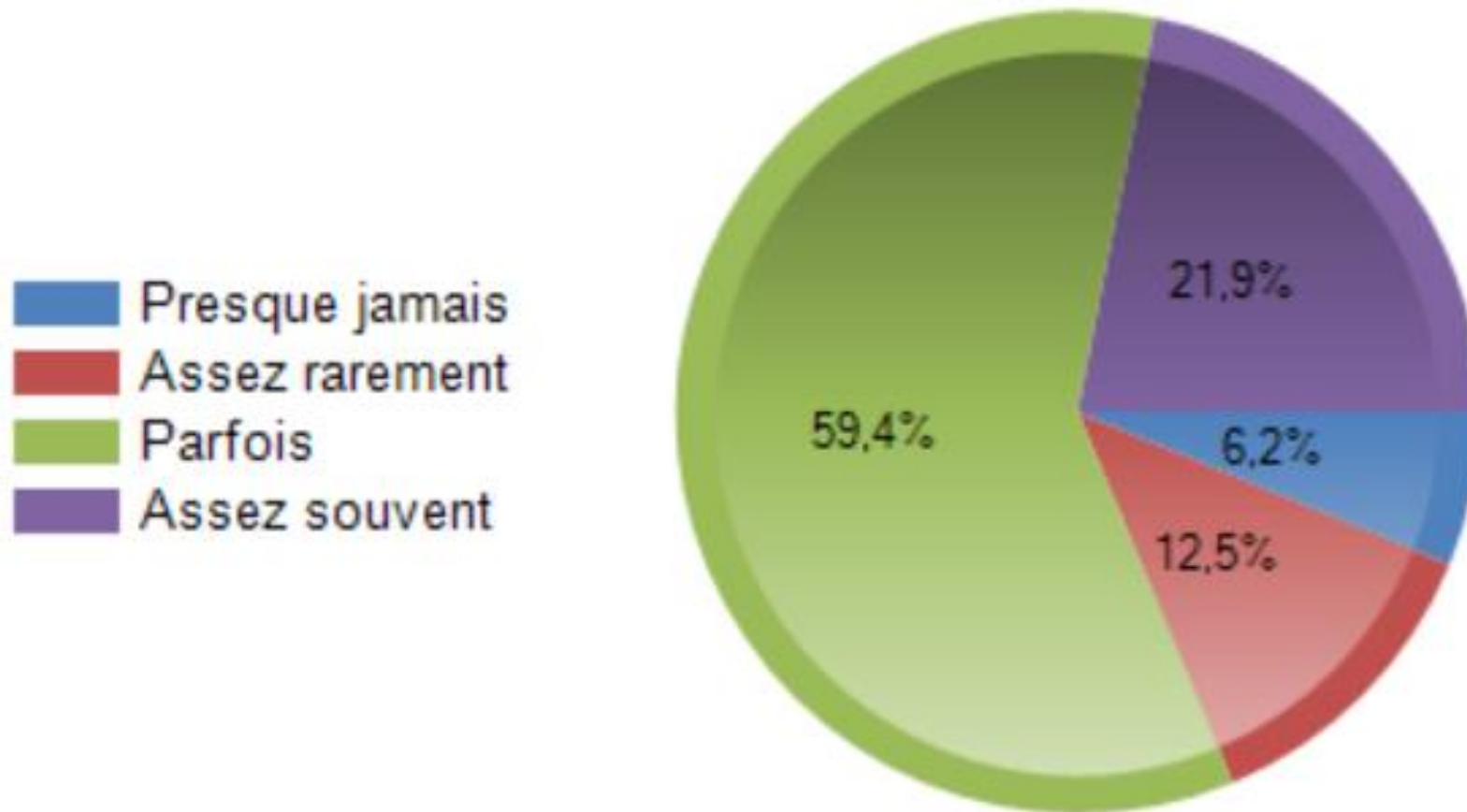
**Question**

Envoyer

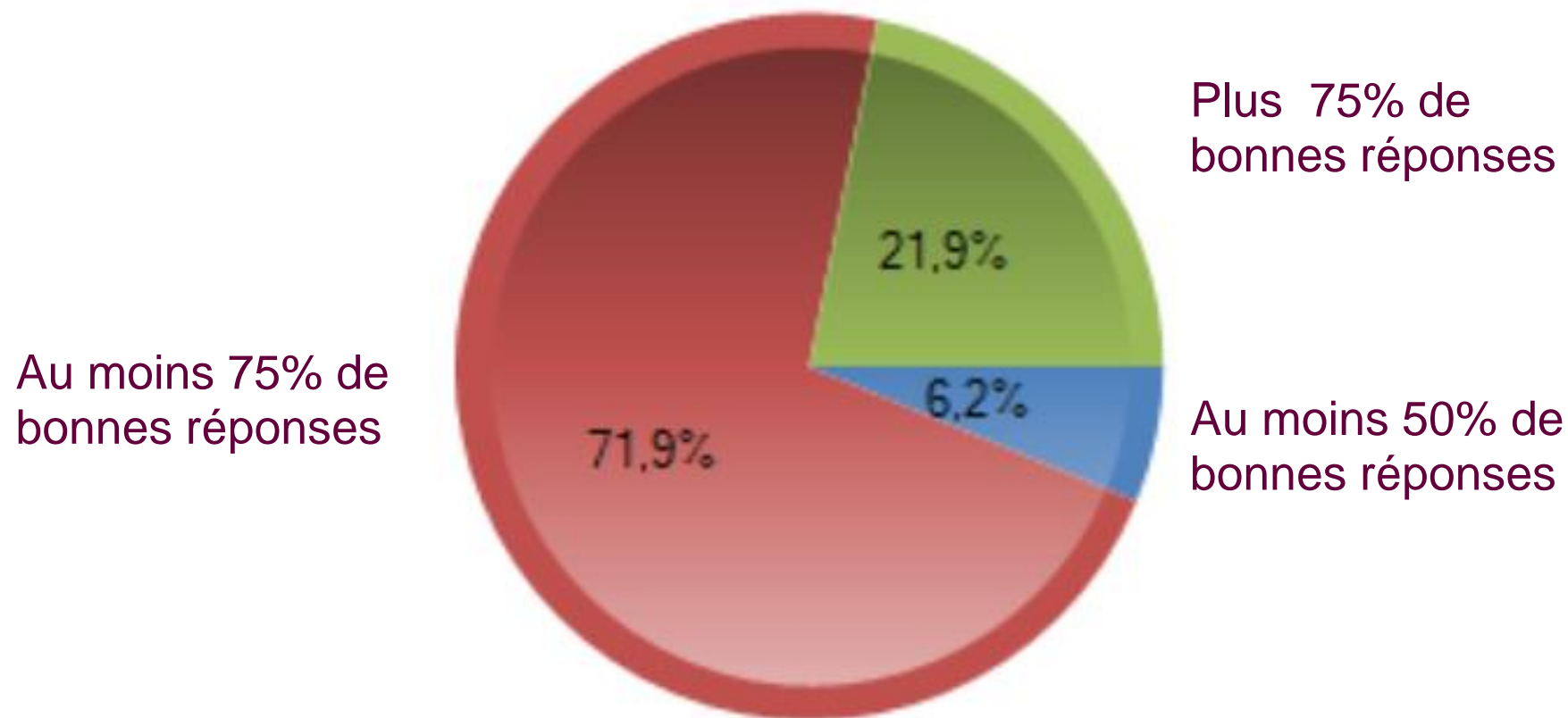
# Liste des thématiques en fonction des données d'entraînement disponibles

aide a la recherche documentaire	459
pret	232
problemes d'accès aux ressources ou espaces numériques	220
horaire	112
contacter un bibliothécaire	111
prolongation	73
reservation livre	68
droits entree	66
consultation des ressources numeriques	58
reservation salle	52
peb	45
impression-reprographie	43
compte lecteur	38
quitus	35
formation	26
navette	20
retard	18
authentification	17
inscription	13
localisation	6
handicap-accessibilite	3
Name: Wrapup, dtype: int64	

# Question : Le chatbot vous a-t-il paru capable d'associer les bonnes thématiques à vos questions ?



**Question : Le chatbot ne sera jamais capable répondre à toutes les questions des utilisateurs. Quel taux minimum de réponse positive vous paraît acceptable ?**



# **Les solutions envisagées et perspectives de collaborations futures**

# Redirection vers les pages web de la DiBISO

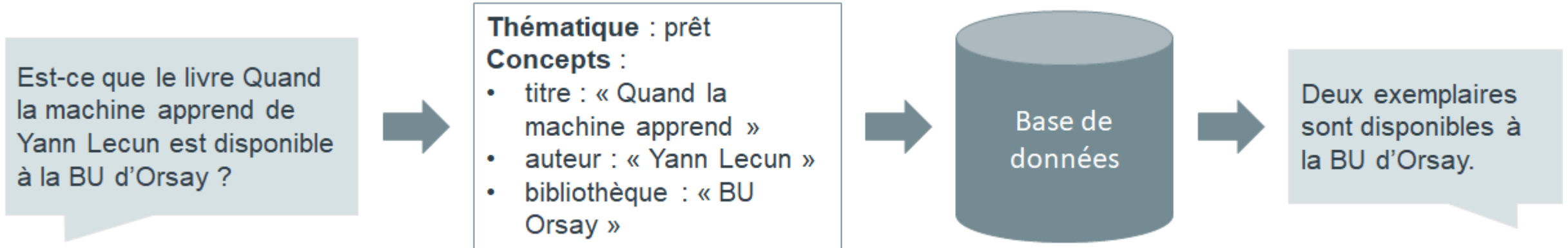


## Limitations :

- Non adapté à toutes les questions
- Expérience utilisateur limitée
- Aucune perspective niveau recherche



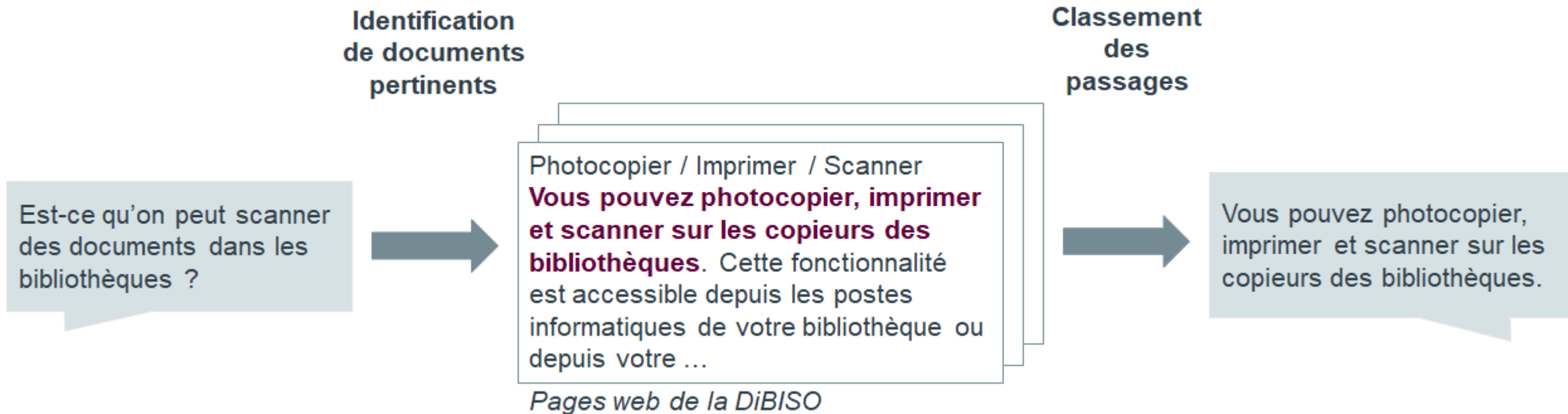
# Accès à une base de données



## Limitations :

- Non adapté aux données type FAQ
- Beaucoup d'ingénierie
- Maintenance régulière des bases de données nécessaire

# Systeme de Question-Réponse



## Limitations :

- Pas adapté à des questions factuelles
- Déjà une question de recherche largement étudiée

# Solution hybride : une perspective de collaboration



## Intérêt :

- Adapté à la diversité des questions
- Question de recherche encore peu étudiée