

# Intention et proposition autour de la gestion des données de protéomique

C. Bruley, J.Garin, A.Viari

1er Juin 2004

# Intention

Rhône-Alpes  
Genopole

## Concevoir et développer un système de gestion des données de la plate-forme.

- Le système de gestion des données doit :
  1. Permettre de conserver une trace de l'activité de la plate-forme
    - suivi des échantillons et des analyses
  2. Permettre de conserver les données générées
  3. Être un support aux analyses
  4. Être un système modulaire

# L'activité

Rhône-Alpes

Genopole

- Concevoir une base de données relationnelle
  - Suivi de l'activité
    - Modèle inspiré de propositions existantes (Pedro)
    - Augmenté d'éléments inhérents au fonctionnement de la plate-forme :
      - » Notion de programme de recherche, projet, étude
      - » Notion d'acteur de la plate-forme
  - Exigences
    - Fidèle
      - » Le système doit être adapté aux pratiques et non l'inverse
    - Souple
      - » L'analyse ne suit pas toujours une procédure standard
    - Pragmatique
      - » Le travail de développement est directement proportionnel à la richesse du modèle

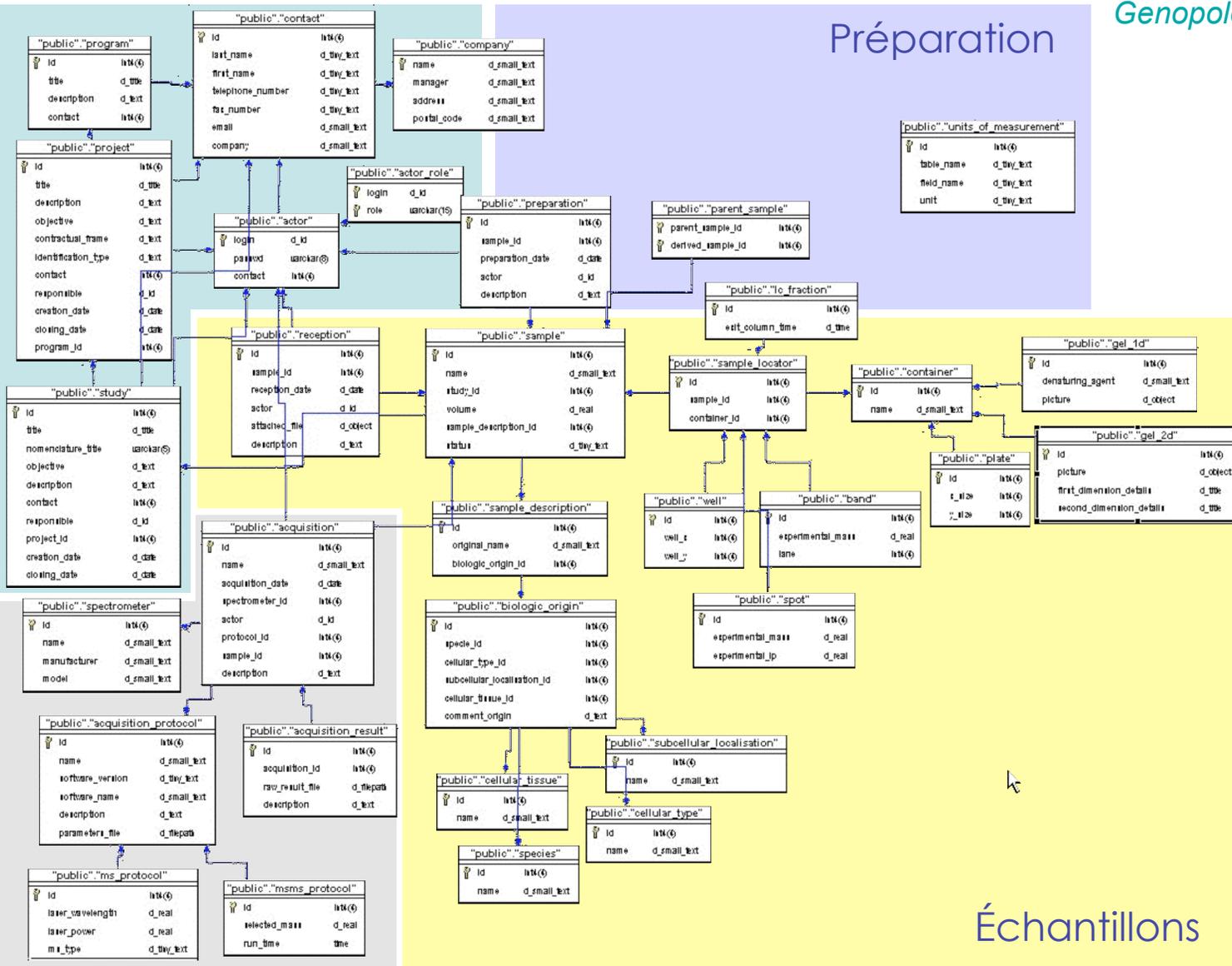
# L'activité

## Projets

## Préparation

## Spectrométrie

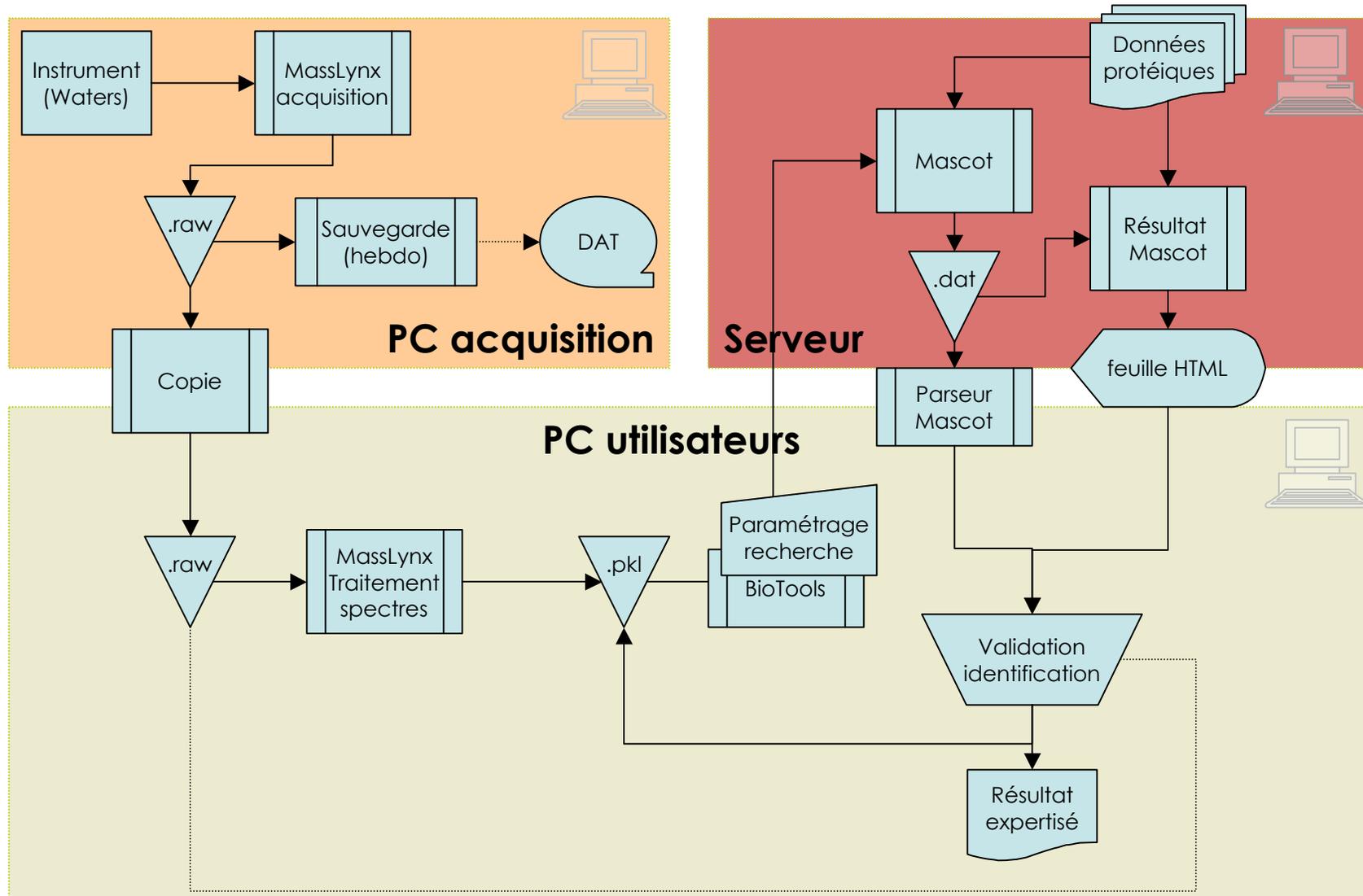
## Échantillons



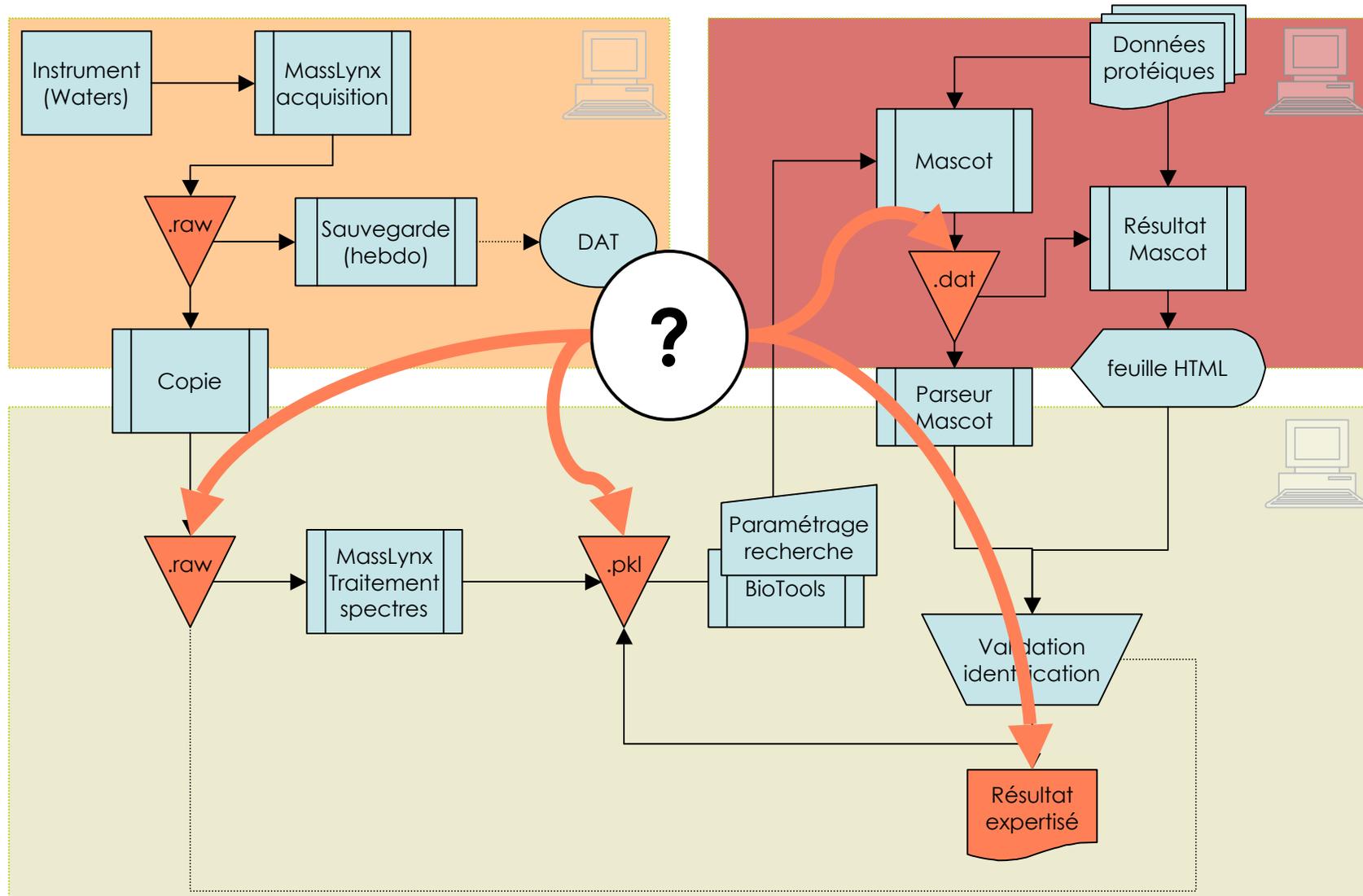
# Intention

- Le système de gestion des données doit :
  1. Permettre de conserver une trace de l'activité de la plate-forme
  2. Permettre de conserver les données générées
    - Centraliser les données pour les lier entre elles
    - Relier les données au suivi de l'activité
    - Stocker puis archiver les données
  3. Être un support aux analyses
  4. Être un système modulaire

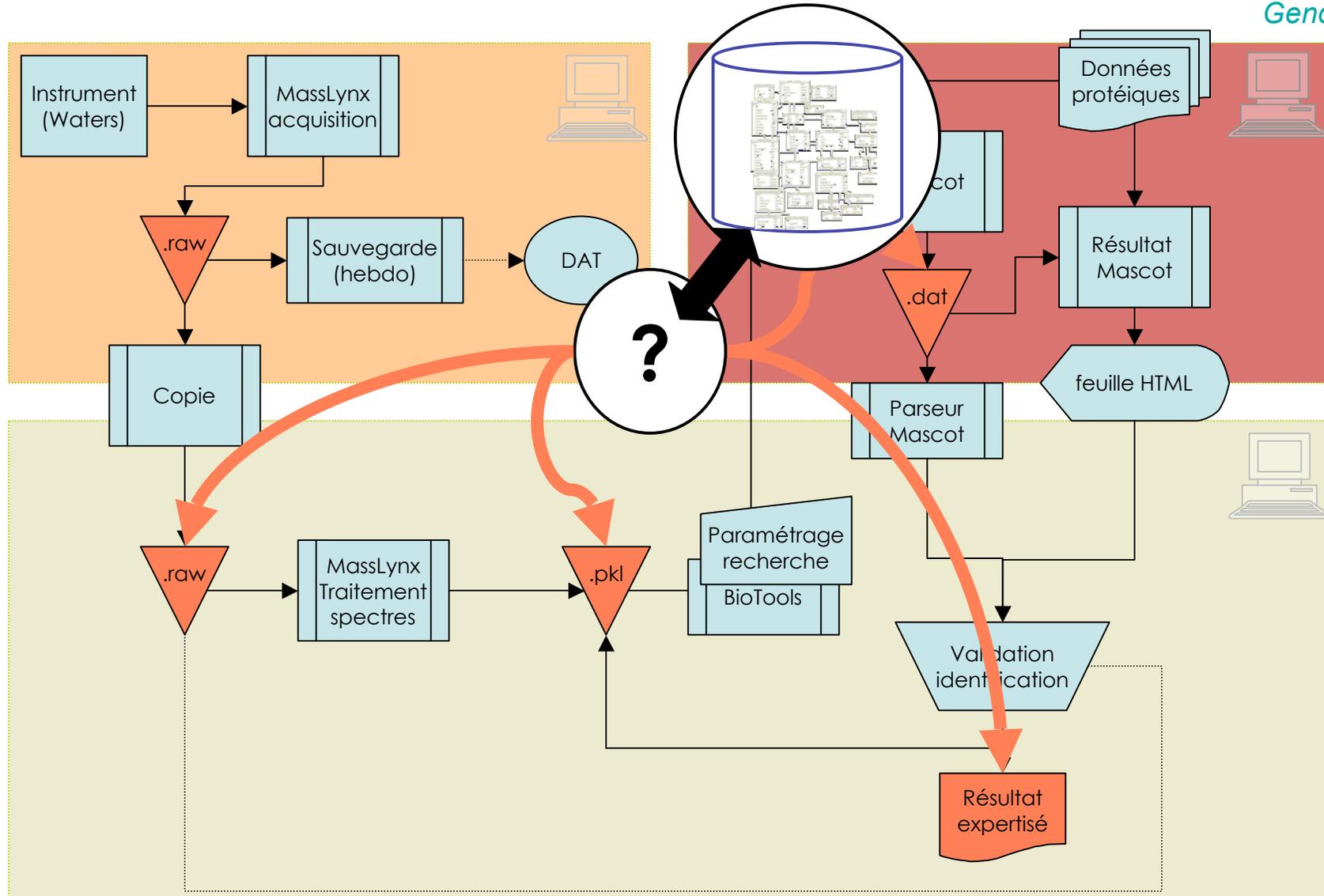
# Les données générées



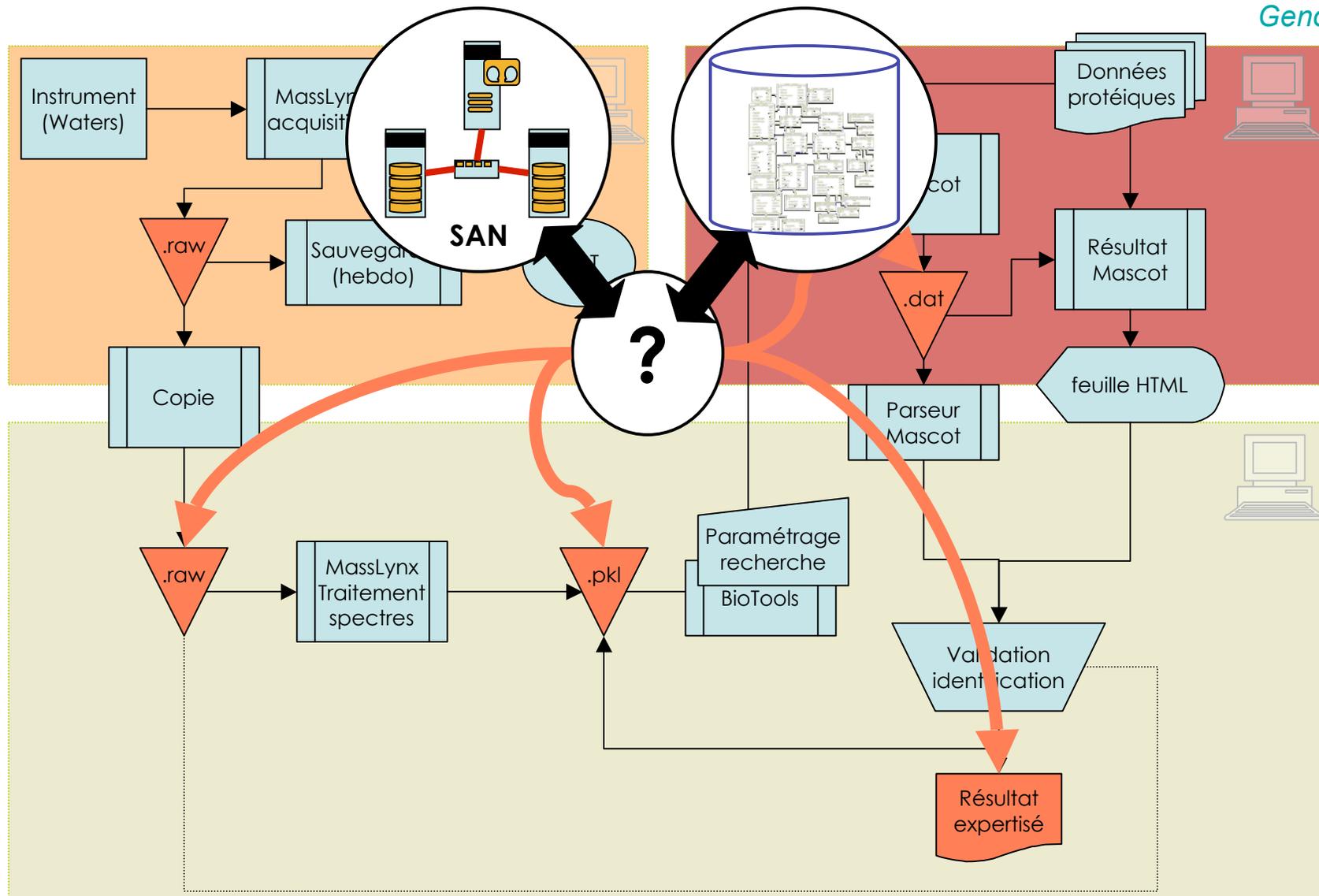
# Les données générées



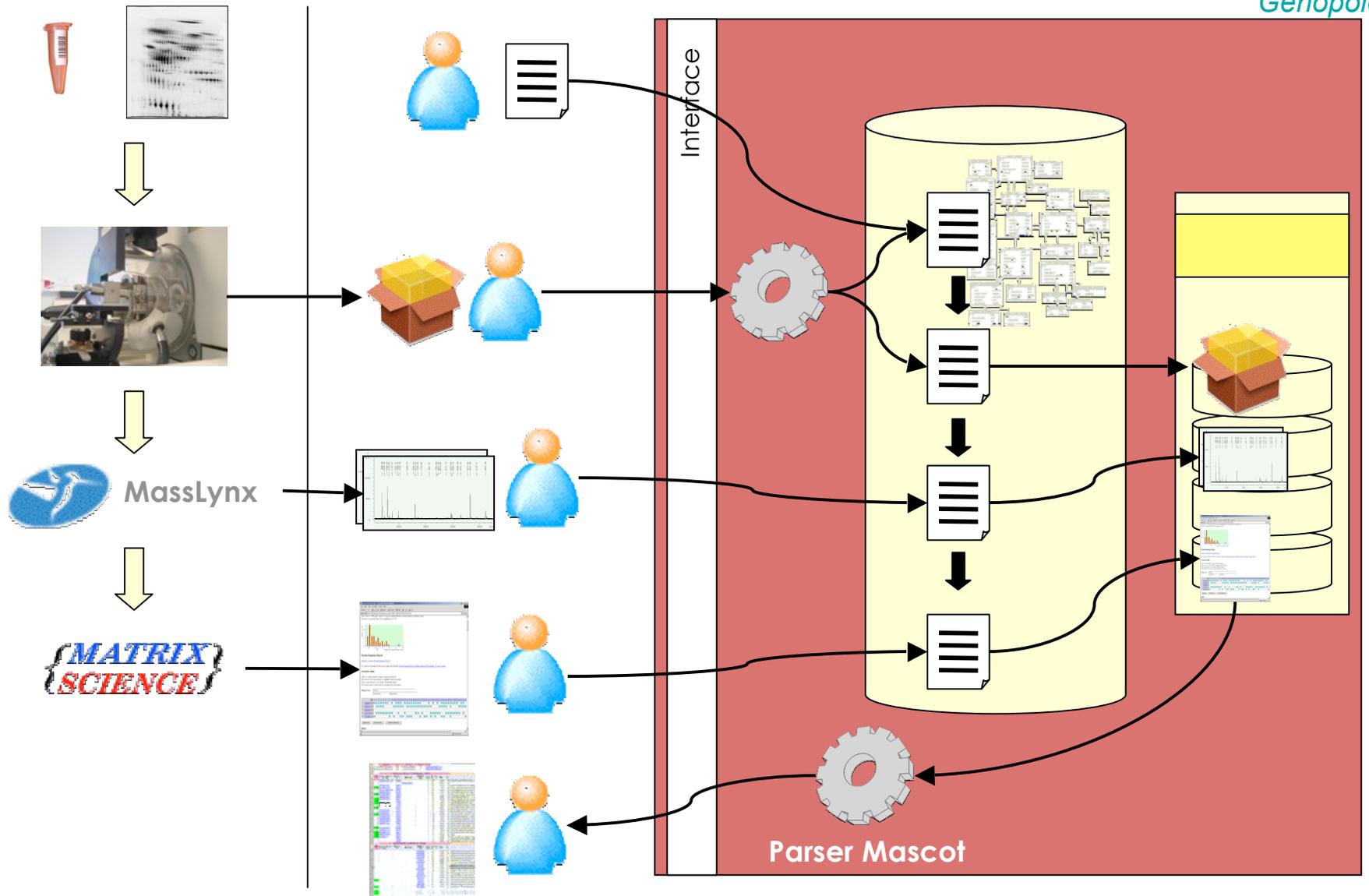
# Les données générées



# Les données générées



# Scénario



# Intention

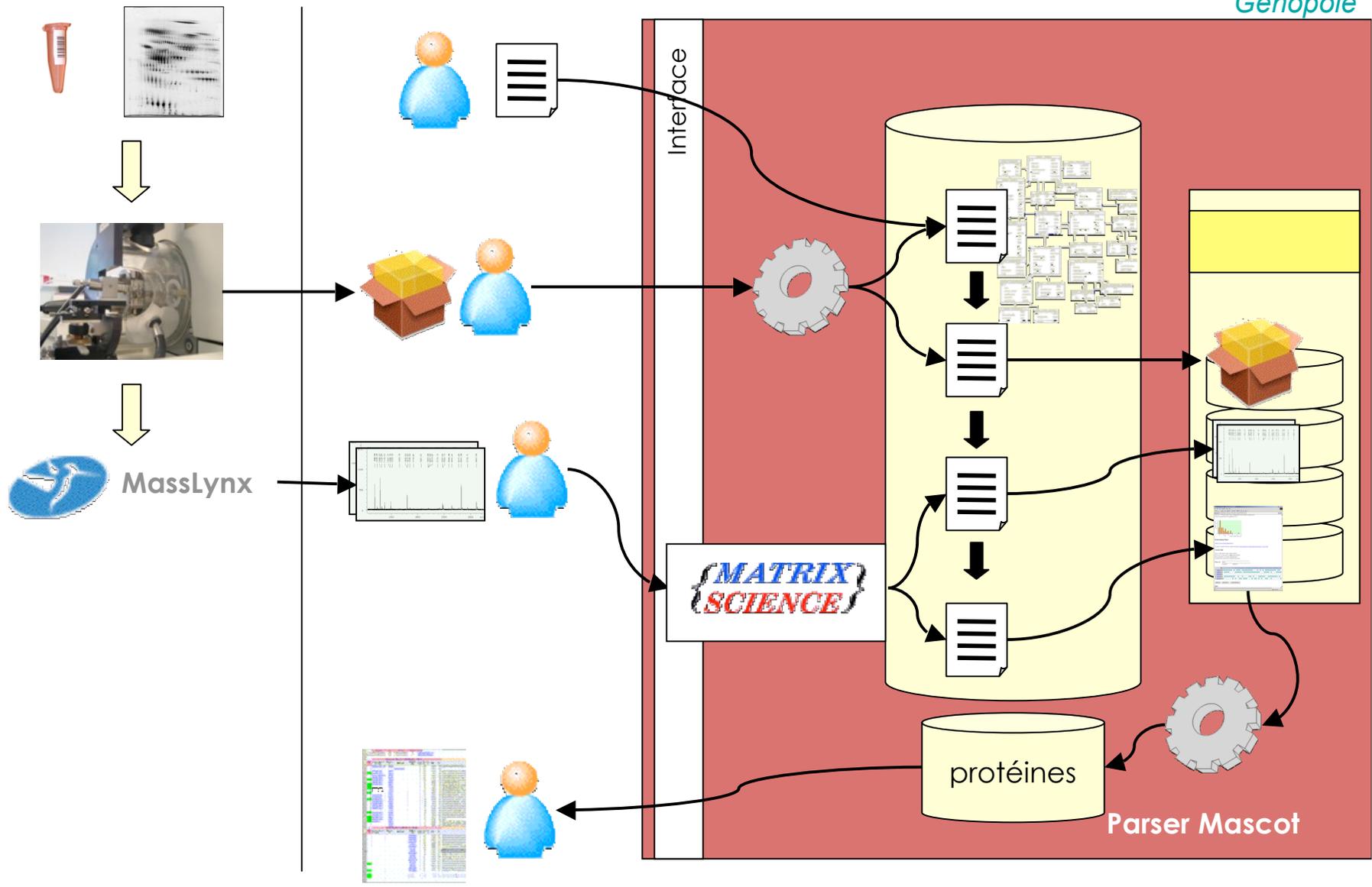
- Le système de gestion des données doit :
  1. Permettre de conserver une trace de l'activité de la plate-forme
  2. Permettre de conserver les données générées
  3. Être un support aux analyses
    - Intégrer des outils d'aide à l'analyse et à la validation des résultats
  4. Être un système modulaire

# Support aux analyses

Rhône-Alpes  
Genopole

- L'activité est soumise aux contraintes imposées par les outils existants.
  - Format des données brutes
  - Format et contenu des fichiers de spectres
  - Format et contenu des résultats d'identification
- Nécessité de développer ses propres outils
  - Pour extraire d'autres informations des fichiers existants
  - Pour adapter l'existant aux pratiques
  - Pour compléter le contenu de certains résultats
  - Pour faciliter la validation des résultats d'identification
  - Pour obtenir d'autres types de résultats
- Une mosaïque d'outils
  - Multiplication des technologies
    - ➡ maintenance, déploiement et mise à jour
  - Multiplication des OS
    - ➡ accès difficile à certains outils

# Scénario

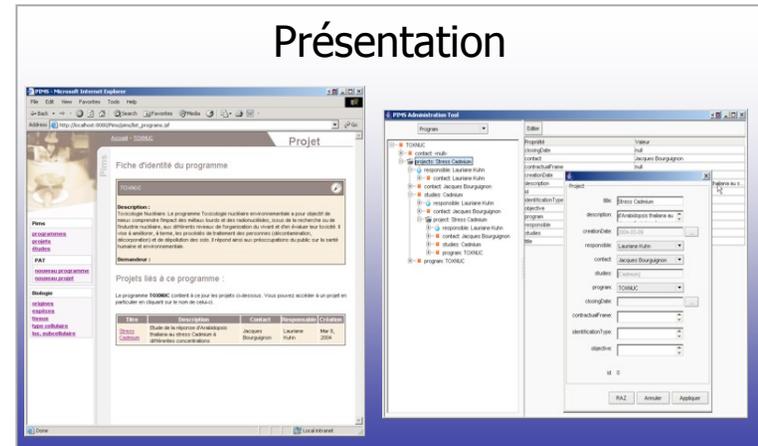


# Intention

- Le système de gestion des données doit :
  1. Permettre de conserver une trace de l'activité de la plate-forme
  2. Permettre de conserver les données générées
  3. Être un support aux analyses
  4. Être un système modulaire
    - Enrichir le modèle de données
    - Étendre la couverture fonctionnelle

# Architecture Logicielle

- Architecture n-tiers
  - Présentation :
    - Client lourd
    - Client léger (navigateur web)
  - Logique applicative :
    - Application des règles métier
      - » Réceptionner un gel
      - » Clore une étude
  - Gestion des données :
    - Objets métier
    - Accès aux données



# Technologie

- Java
  - Présentation :
    - JSP
  - Logique applicative :
    - JSF / composants J2EE
  - Gestion des données :
    - Postgresql
    - Hibernate (mapping relationnel / objet)
  
- Effort de cohérence technologique vis à vis des développements en cours
  - répondre aux besoins immédiats
  - facilité d'intégration

# Avancement

- Abouti
  - Choix technologiques
  - Conception du modèle
  - Mapping relationnel / objet
  - Outils divers et variés
    - Qualification des spectres, recherche de masses ... etc
- En cours
  - Interface de saisie des données
- A venir
  - Transfert des données
  - Modules
    - Base de données de protéines identifiées (AFPdb ? PROTIcDb ?)

# Proposition

Rhône-Alpes  
*Genopole*

- la richesse du système dépend directement des forces de développement en présence
  - ➔ Mutualiser les efforts de développement de différentes plates-formes