

Formations autour de la diffusion de la recherche et de la science ouverte
Les humanités numériques en pratique
16 février 2023

MODÉLISER LES CONNAISSANCES EN SHS :
POURQUOI, COMMENT, JUSQU'OU ?

Thomas Bottini

thomas.bottini@cnrs.fr

Institut de Recherche en Musicologie — IReMus, UMR 8223 CNRS

PLAN



Le rôle du modélisateur/de la modélisatrice dans un projet de SHS



Quelques mots sur le Web sémantique comme milieu technique adapté à l'expression et à la diffusion des données de la recherche |



Les fondements de l'ontologie CIDOC-CRM



Quelques cas concrets



MODÉLISER ?

- Modèle conceptuel vs. ontologie ?
- Entités + propriétés.



DE L'ACTIVITÉ DE RECHERCHE À LA DONNÉE

Comment « l'ingénierie des données » intervient-elle aux différentes étapes d'un projet de recherche en SHS :



ÉCOUTER

- Les chercheur·se·s en situation de production de connaissances susceptibles d'être traduites en données informatiques structurées ont sur ce point des degrés de réflexivité différents. Le premier rôle de l'ingénieur est d'ordre maïeutique.
- En SHS, ces connaissances peuvent résulter d'une activité descriptive (sources) ou interprétative (choses dites à propos des sources).
- Dès lors, avoir un modèle conceptuel de type ontologique permet d'organiser l'écoute et l'analyse des pratiques de production de connaissances.



ACCOMPAGNER

- L'ingénieur devient de facto un collaborateur méthodologique :
- Il aide à révéler la structure interne des sources et des phénomènes par lesquels est appréhendé l'objet de recherche (dimension maïeutique + heuristique).
- L'idéal est de mettre en place des interfaces de saisie ergonomiques, mais c'est là où « le bât blesse » : les ontologies sont souples, riches, expressives, mais pas instrumentées pour les humains (formulaires...) par rapport aux données relationnelles SQL.
- Dans l'expression des objectifs scientifiques. Révéler la structure des observables et des connaissances produites à leur endroit permet de repenser les finalités scientifiques.
- C'est une situation de travail typique des HN dans laquelle la technique n'a pas un rôle ancillaire.



MODÉLISER

- Les connaissances nouvelles s'incarnant dans des « données scientifiques ».
- Les sources auxquelles elles se rapportent.
- Les faits sociaux, conceptuels, matériels & sémiotiques dénotés ou conotés par les sources et/ou le terrain.
- Le contexte de production des connaissances (qu'est ce qui fait qu'une donnée est « scientifique » ?).
- C'est l'objet de cette formation !



GÉNÉRER, GÉRER & DIFFUSER

techniquement les données, en les rendant conformes au modèle. [← ce n'est pas l'objet de cette formation, mais je vous dirai tout de même deux mots du Web Sémantique et du RDF]



GÉNÉRER, GÉRER & DIFFUSER

techniquement les données, en les rendant conformes au modèle. [← ce n'est pas l'objet de cette formation, mais je vous dirai tout de même deux mots du Web Sémantique et du RDF] 5. DIFFUSER techniquement les données : 1. Accès par les machines (API, SPARQL Endpoint) 2. Consultation par les humains (sites Web)

