

Consortium MUSICA2  
GT2 — Ontologies, thésauri & catalogage  
4 juillet 2023, CESR, Université de Tours

---

OBJECTIFS DU GT2  
&  
INTRODUCTION À L'ONTOLOGIE CIDOC-CRM

---

{ thomas.bottini • augustin.braud • achille.davy-rigaux • marco.gurrieri } @cnrs.fr

*Institut de Recherche en Musicologie — IReMus, UMR 8223 CNRS*

<https://github.com/Amleth/consortium-musica2-gt2-ontologies/>

# PLAN

- I. Énonciation des objectifs du GT2
  - Problèmes identifiés
  - Réponses visées
- II. Généralités sur le Web sémantique
- III. Introduction à l'ontologie CIDOC-CRM
- IV. Cas concrets récurrents
- V. Deux mots sur LRMoo et Doremus
- VI. Une étude de cas : modélisation des partitions encodées



# I. ÉNONCIATION DES OBJECTIFS DU GT2

PROBLÈMES IDENTIFIÉS



## FAIR, RECHERCHE & INGÉNIERIE

- L'époque est au FAIR et au LOD. Afin que l'ouverture des données de la recherche, leur interopérabilité et leur mise en relation avec des sources de données tierces soient correctement traitées, il faut que ces questions soient pensées très en amont des projets de recherche, et finement articulées aux questions méthodologiques, voire scientifique.




## FAIR, RECHERCHE & INGÉNIERIE

- L'époque est au FAIR et au LOD. Afin que l'ouverture des données de la recherche, leur interopérabilité et leur mise en relation avec des sources de données tierces soient correctement traitées, il faut que ces questions soient pensées très en amont des projets de recherche, et finement articulées aux questions méthodologiques, voire scientifique.
- Il faut alors des ingénieurs et ingénieures qui « pensent » les données conjointement avec les chercheurs et chercheuses, dans des situations de travail où la technique ne joue pas un rôle ancillaire :
  - Ces ingénieurs et ingénieures doivent jouer un rôle maïeutique (savoir poser les questions, confronter le chercheur ou la chercheuse à des cas limites pour l'amener à mieux comprendre ses objets d'étude).
  - Le travail d'explicitation, de modélisation, des données doit avoir une fonction heuristique : aider à révéler la structure interne des sources et des phénomènes étudiés.



## FAIR, RECHERCHE & INGÉNIERIE

- L'époque est au FAIR et au LOD. Afin que l'ouverture des données de la recherche, leur interopérabilité et leur mise en relation avec des sources de données tierces soient correctement traitées, il faut que ces questions soient pensées très en amont des projets de recherche, et finement articulées aux questions méthodologiques, voire scientifique.
- Il faut alors des ingénieurs et ingénieures qui « pensent » les données conjointement avec les chercheurs et chercheuses, dans des situations de travail où la technique ne joue pas un rôle ancillaire :
  - Ces ingénieurs et ingénieures doivent jouer un rôle maïeutique (savoir poser les questions, confronter le chercheur ou la chercheuse à des cas limites pour l'amener à mieux comprendre ses objets d'étude).
  - Le travail d'explicitation, de modélisation, des données doit avoir une fonction heuristique : aider à révéler la structure interne des sources et des phénomènes étudiés.
-  Les ressources d'ingénierie sont trop maigres, ce niveau dialogue est rare.

## CAPITALISATION DES CONNAISSANCES TECHNO- MÉTHODOLOGIQUES

- La FAIRisation des données musicologiques suppose une dynamique informationnelle intellectuelle et technique entre les projets de BDD.



## CAPITALISATION DES CONNAISSANCES TECHNO-MÉTHODOLOGIQUES

- La FAIRisation des données musicologiques suppose une **dynamique informationnelle intellectuelle et technique entre les projets de BDD**.
- Pour la bâtir, un réseau d'acteurs et d'actrices est nécessaire, mais :
  - Il faut une complémentarité recherche/ingénierie/SIB car ces **connaissances sont très abstraites** et difficiles à saisir.





## CAPITALISATION DES CONNAISSANCES TECHNO-MÉTHODOLOGIQUES

- La FAIRisation des données musicologiques suppose une **dynamique informationnelle intellectuelle et technique entre les projets de BDD**.
- Pour la bâtir, un réseau d'acteurs et d'actrices est nécessaire, mais :
  - Il faut une complémentarité recherche/ingénierie/SIB car ces **connaissances sont très abstraites** et difficiles à saisir.
  - Les musicologues devant piloter de tels projet manquent d'informations claires sur les **enjeux scientifiques des méthodes et technologies disponibles** pour correctement modéliser les informations scientifiques. Ceci peut conduire à des choix techniques inadaptés qui obèrent les possibilités scientifiques.



## CAPITALISATION DES CONNAISSANCES TECHNO-MÉTHODOLOGIQUES

- La FAIRisation des données musicologiques suppose une **dynamique informationnelle intellectuelle et technique entre les projets de BDD**.
- Pour la bâtir, un réseau d'acteurs et d'actrices est nécessaire, mais :
  - Il faut une complémentarité recherche/ingénierie/SIB car ces **connaissances sont très abstraites** et difficiles à saisir.
  - Les musicologues devant piloter de tels projet manquent d'informations claires sur les **enjeux scientifiques des méthodes et technologies disponibles** pour correctement modéliser les informations scientifiques. Ceci peut conduire à des choix techniques inadaptés qui obèrent les possibilités scientifiques.
  - Les profils techniques sont recrutés sur des contrats courts.




## CAPITALISATION DES CONNAISSANCES TECHNO-MÉTHODOLOGIQUES

- La FAIRisation des données musicologiques suppose une **dynamique informationnelle intellectuelle et technique entre les projets de BDD**.
- Pour la bâtir, un réseau d'acteurs et d'actrices est nécessaire, mais :
  - Il faut une complémentarité recherche/ingénierie/SIB car ces **connaissances sont très abstraites** et difficiles à saisir.
  - Les musicologues devant piloter de tels projet manquent d'informations claires sur les **enjeux scientifiques des méthodes et technologies disponibles** pour correctement modéliser les informations scientifiques. Ceci peut conduire à des choix techniques inadaptés qui obèrent les possibilités scientifiques.
  - Les profils techniques sont recrutés sur des contrats courts.
  - Les prestataires n'ont pas d'intérêt à s'inscrire dans les réseaux HN.



## CAPITALISATION DES CONNAISSANCES TECHNO-MÉTHODOLOGIQUES

- La FAIRisation des données musicologiques suppose une **dynamique informationnelle intellectuelle et technique entre les projets de BDD**.
- Pour la bâtir, un réseau d'acteurs et d'actrices est nécessaire, mais :
  - Il faut une complémentarité recherche/ingénierie/SIB car ces **connaissances sont très abstraites** et difficiles à saisir.
  - Les musicologues devant piloter de tels projet manquent d'informations claires sur les **enjeux scientifiques des méthodes et technologies disponibles** pour correctement modéliser les informations scientifiques. Ceci peut conduire à des choix techniques inadaptés qui obèrent les possibilités scientifiques.
  - Les profils techniques sont recrutés sur des contrats courts.
  - Les prestataires n'ont pas d'intérêt à s'inscrire dans les réseaux HN.
-  Conséquemment, les **connaissances d'ingénierie spécifiques à la modélisation des données de la discipline sont peu capitalisées** ; chaque nouveau développement peine à bénéficier de l'expérience méthodologique et conceptuelle acquise informellement au fil des projets passés.



# TECHNICITÉ DU WEB SÉMANTIQUE & DES ONTOLOGIES

- Les questions techniques soulevées par le FAIRisation des données de la recherche sont non-triviales, les profils techniques compétents sont rares, et notre attractivité est parfois faible :



# TECHNICITÉ DU WEB SÉMANTIQUE & DES ONTOLOGIES

- Les questions techniques soulevées par le FAIRisation des données de la recherche sont non-triviales, les profils techniques compétents sont rares, et notre attractivité est parfois faible :
  - 🚧 Les ontologies **standards** dans les mondes de la recherche, des bibliothèques et des musées (CIDOC-CRM, LRMoo : nous allons y revenir) sont **expressives** (elles aident à ne pas réduire ni trahir les productions analytiques des chercheurs et des chercheuses) et sont maintenues par des communautés fortes, mais **complexes à comprendre et mettre en œuvre**, et manquent d'exemples concrets qui permettraient aux musicologues de mieux saisir leur intérêt.



# TECHNICITÉ DU WEB SÉMANTIQUE & DES ONTOLOGIES

- Les questions techniques soulevées par le FAIRisation des données de la recherche sont non-triviales, les profils techniques compétents sont rares, et notre attractivité est parfois faible :
  - 🚧 Les ontologies **standards** dans les mondes de la recherche, des bibliothèques et des musées (CIDOC-CRM, LRMoo : nous allons y revenir) sont **expressives** (elles aident à ne pas réduire ni trahir les productions analytiques des chercheurs et des chercheuses) et sont maintenues par des communautés fortes, mais **complexes à comprendre et mettre en œuvre**, et manquent d'exemples concrets qui permettraient aux musicologues de mieux saisir leur intérêt.
  - 🚧 Le Web sémantique semble être **le meilleur milieu technique pour représenter des données scientifiques** (nous allons y revenir), mais il est difficile de construire des outils de saisie ergonomiques et génériques s'appuyant sur RDF et sur les ontologies citées infra.

# I. ÉNONCIATION DES OBJECTIFS DU GT2

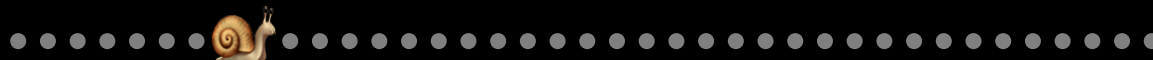
RÉPONSES VISÉES





# IDENTIFIER ET CARACTÉRISER DES SITUATIONS DE MODÉLISATION

- Il existe des **situations récurrentes de modélisation**, se présentant quels que soient les objets de recherche et quelles que soient les périodes étudiées.



## IDENTIFIER ET CARACTÉRISER DES SITUATIONS DE MODÉLISATION

- Il existe des **situations récurrentes de modélisation**, se présentant quels que soient les objets de recherche et quelles que soient les périodes étudiées.
- Dans le GT2, nous ambitionnons d'**identifier** ces situations, de les **décrire**, et de les **formaliser** sous la forme de **fragments de graphes CIDOC-CRM**, afin d'aider les concepteurs et conceptrices bases de données musicologiques à prendre les bonnes décisions lors de l'établissement des modèles conceptuels et relationnels de leurs bases.



# IDENTIFIER ET CARACTÉRISER DES SITUATIONS DE MODÉLISATION

- Il existe des **situations récurrentes de modélisation**, se présentant quels que soient les objets de recherche et quelles que soient les périodes étudiées.
- Dans le GT2, nous ambitionnons d'**identifier** ces situations, de les **décrire**, et de les **formaliser** sous la forme de **fragments de graphes CIDOC-CRM**, afin d'aider les concepteurs et conceptrices bases de données musicologiques à prendre les bonnes décisions lors de l'établissement des modèles conceptuels et relationnels de leurs bases.
- Chaque situation sera publiée sur le **dépôt Github du GT2** , et proposera :
  - un fragment de graphe CIDOC-CRM illustrant concrètement la proposition de modélisation ;
  - une problématisation scientifique ;
  - un recensement des usages connus ;
  - une explicitation de l'usage qui est fait des classes et propriétés du CIDOC-CRM mises en jeu ;
  - une description des verrous et des limites éventuelles rencontrées.



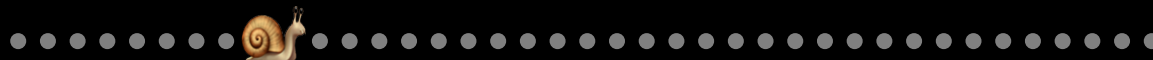
# CONSTITUER UNE COMMUNAUTÉ DE TRAVAIL AUTOUR DE CES SITUATIONS DE MODÉLISATION

- À court terme :
  - Organiser entre six et huit ateliers thématiques de modélisation collective en septembre, octobre, novembre et décembre 2023.
  - Chaque atelier sera consacré à l'exploration à fond d'une situation de modélisation, et convoquera des « experts » de cette question.



# CONSTITUER UNE COMMUNAUTÉ DE TRAVAIL AUTOUR DE CES SITUATIONS DE MODÉLISATION

- À court terme :
  - Organiser entre six et huit ateliers thématiques de modélisation collective en septembre, octobre, novembre et décembre 2023.
  - Chaque atelier sera consacré à l'exploration à fond d'une situation de modélisation, et convoquera des « experts » de cette question.
- À long terme, nous maintiendrons ces ateliers de modélisation avec des séances tenues régulièrement, appuyés par des référents et référentes musicologues pour chaque situation de modélisation.



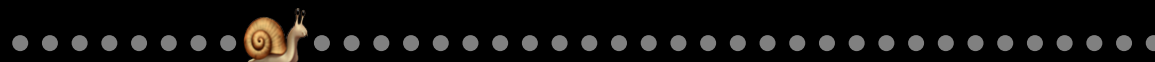
# CONSTITUER UNE COMMUNAUTÉ DE TRAVAIL AUTOUR DE CES SITUATIONS DE MODÉLISATION

- À court terme :
  - Organiser entre six et huit ateliers thématiques de modélisation collective en septembre, octobre, novembre et décembre 2023.
  - Chaque atelier sera consacré à l'exploration à fond d'une situation de modélisation, et convoquera des « experts » de cette question.
- À long terme, nous maintiendrons ces ateliers de modélisation avec des séances tenues régulièrement, appuyés par des référents et référentes musicologues pour chaque situation de modélisation.
- Ces exemples d'utilisation du CIDOC-CRM autour de cas concrets immédiatement pertinents pour la musicologie constitueront une ressource de référence pédagogique, méthodologique, pratique et scientifique de premier plan, qu'il faudra maintenir.



# CONSTITUER UNE COMMUNAUTÉ DE TRAVAIL AUTOUR DE CES SITUATIONS DE MODÉLISATION

- À court terme :
  - Organiser entre six et huit ateliers thématiques de modélisation collective en septembre, octobre, novembre et décembre 2023.
  - Chaque atelier sera consacré à l'exploration à fond d'une situation de modélisation, et convoquera des « experts » de cette question.
- À long terme, nous maintiendrons ces ateliers de modélisation avec des séances tenues régulièrement, appuyés par des référents et référentes musicologues pour chaque situation de modélisation.
- Ces exemples d'utilisation du CIDOC-CRM autour de cas concrets immédiatement pertinents pour la musicologie constitueront une ressource de référence pédagogique, méthodologique, pratique et scientifique de premier plan, qu'il faudra maintenir.
- Les caractérisations de ces situations formeront une partie du guide de bonnes pratiques du GT2. qui contiendra également un guide introductif à la modélisation des données musicologiques avec le CIDOC-CRM.



## IDENTIFIER DES SOLUTIONS TECHNIQUES POUR LA MANIPULATION DES DONNÉES SÉMANTIQUES

- Recenser l'ensemble des problèmes techniques et méthodologiques soulevés par la manipulation de données sémantiques :
  - Ergonomie de la saisie.
  - Conformation des données existantes au CIDOC-CRM.
  - Alignement avec des données tierces.
  - Création de points d'accès (SPARQL endpoint).
  - Définition d'une politique pour la publication des données représentant des concepts ou des termes dans des vocabulaires contrôlés/thésauri.
  - Création d'interfaces de consultation génériques.



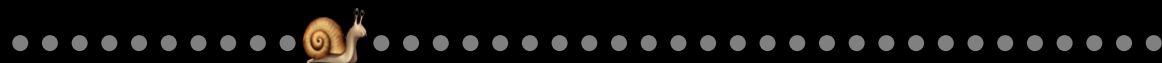


## IDENTIFIER DES SOLUTIONS TECHNIQUES POUR LA MANIPULATION DES DONNÉES SÉMANTIQUES

- Recenser l'ensemble des problèmes techniques et méthodologiques soulevés par la manipulation de données sémantiques :
  - Ergonomie de la saisie.
  - Conformation des données existantes au CIDOC-CRM.
  - Alignement avec des données tierces.
  - Création de points d'accès (SPARQL endpoint).
  - Définition d'une politique pour la publication des données représentant des concepts ou des termes dans des vocabulaires contrôlés/thésauri.
  - Création d'interfaces de consultation génériques.
- Là aussi, nous formerons un groupe de réflexion technique pérenne.

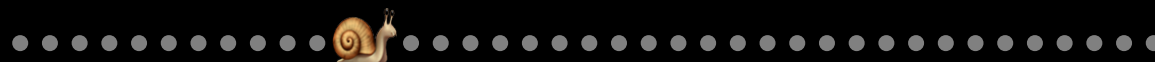


## II. GÉNÉRALITÉS SUR LE WEB SÉMANTIQUE



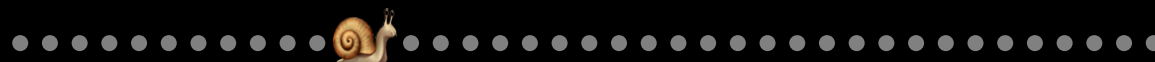
# LE WEB SÉMANTIQUE, EN UNE SLIDE

- Promesse d'une base de données à l'échelle du Web. Le Web initial (Tim Berners Lee, 1991) était un Web de documents liés (hypertexte), le Web sémantique est un Web de données liées.



# LE WEB SÉMANTIQUE, EN UNE SLIDE

- Promesse d'une base de données à l'échelle du Web. Le Web initial (Tim Berners Lee, 1991) était un Web de documents liés (hypertexte), le Web sémantique est un Web de données liées.
- Chaque donnée est identifiée par une URL.



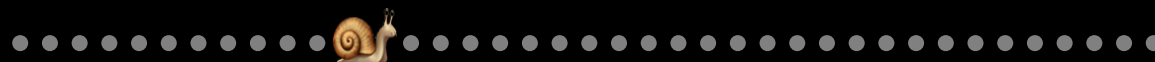
# LE WEB SÉMANTIQUE, EN UNE SLIDE

- Promesse d'une base de données à l'échelle du Web. Le Web initial (Tim Berners Lee, 1991) était un Web de documents liés (hypertexte), le Web sémantique est un Web de données liées.
- Chaque donnée est identifiée par une URL.
- Toute information s'exprime sous la forme d'un triplet (sujet/prédicat/objet) dans un langage de description, le RDF.



# LE WEB SÉMANTIQUE, EN UNE SLIDE

- Promesse d'une base de données à l'échelle du Web. Le Web initial (Tim Berners Lee, 1991) était un Web de documents liés (hypertexte), le Web sémantique est un Web de données liées.
- Chaque donnée est identifiée par une URL.
- Toute information s'exprime sous la forme d'un triplet (sujet/prédicat/objet) dans un langage de description, le RDF.
- La connexion de ces triplets RDF forme un graphe.



# LE WEB SÉMANTIQUE, EN UNE SLIDE

- Promesse d'une base de données à l'échelle du Web. Le Web initial (Tim Berners Lee, 1991) était un Web de documents liés (hypertexte), le Web sémantique est un Web de données liées.
- Chaque donnée est identifiée par une URL.
- Toute information s'exprime sous la forme d'un triplet (sujet/prédicat/objet) dans un langage de description, le RDF.
- La connexion de ces triplets RDF forme un graphe.
- Chaque prédicat est également identifié par une URL.

[https://data.bnf.fr/13962206/morton\\_feldman\\_for\\_philip\\_guston/](https://data.bnf.fr/13962206/morton_feldman_for_philip_guston/)

<http://purl.org/dc/terms/creator>

[https://data.bnf.fr/fr/13928795/morton\\_feldman/](https://data.bnf.fr/fr/13928795/morton_feldman/)

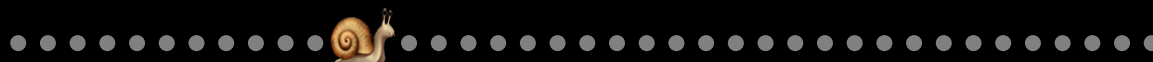


# LE WEB SÉMANTIQUE, EN UNE SLIDE

- Promesse d'une base de données à l'échelle du Web. Le Web initial (Tim Berners Lee, 1991) était un Web de documents liés (hypertexte), le Web sémantique est un Web de données liées.
- Chaque donnée est identifiée par une URL.
- Toute information s'exprime sous la forme d'un triplet (sujet/prédicat/objet) dans un langage de description, le RDF.
- La connexion de ces triplets RDF forme un graphe.
- Chaque prédicat est également identifié par une URL.

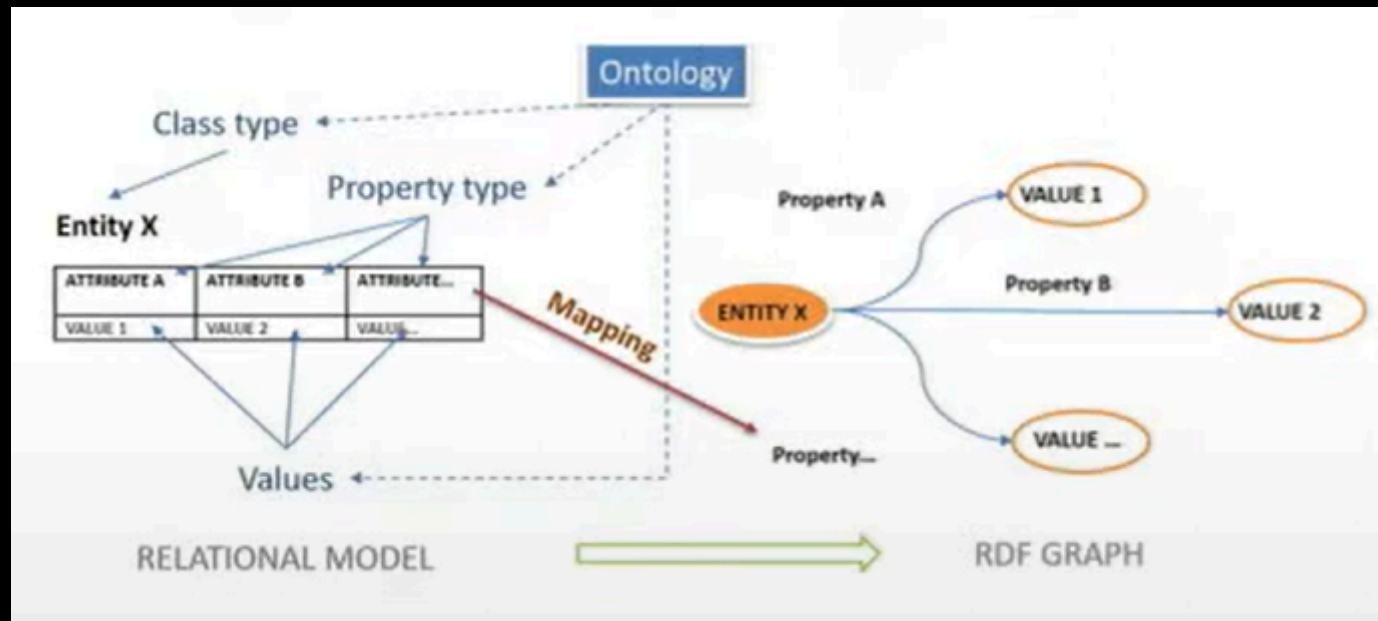
[https://data.bnf.fr/13962206/morton\\_feldman\\_for\\_philip\\_guston/](https://data.bnf.fr/13962206/morton_feldman_for_philip_guston/) <http://purl.org/dc/terms/creator> [https://data.bnf.fr/fr/13928795/morton\\_feldman/](https://data.bnf.fr/fr/13928795/morton_feldman/)

- C'est le milieu technique idéal pour des données FAIR, pour l'expression et à la diffusion des données de la recherche.





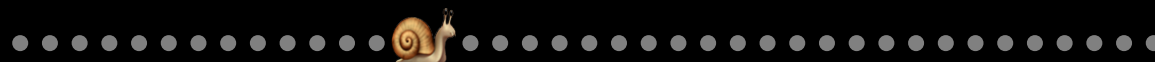
# DONNÉES RELATIONNELLES VS GRAPHE RDF



*Corago in LOD - Seminar by Angelo Pompilio and Paolo Bonora, Digital Humanities and Digital Knowledge, Università di Bologna, 2017.*

# QU'EST CE QU'UNE ONTOLOGIE ?

- Formalisation d'un modèle conceptuel pour un domaine identifié proposant des :
  - **Classes** : types d'entités peuplant le domaine, possiblement organisées selon des relations d'héritage (spécificité). On appelle **individu** une ressource qui est du type d'une classe.
  - **Propriétés** : aspects, caractéristiques, attributs possibles de ces classes, qui peuvent soit pointer vers une valeur, soit vers un individu.
- Utiliser les classes et les propriétés d'une ontologie confère ainsi une sémantique partagée aux données RDF (les individus identifiés par des URL seront des sujets ou des objets, les propriétés des classes seront des prédicats).
- Vous connaissez peut-être déjà une ontologie : SKOS (pour construire des thésauri).

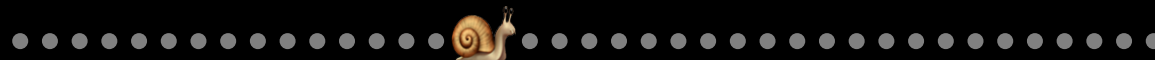


# III. INTRODUCTION À L'ONTOLOGIE CIDOC-CRM



# LE CIDOC-CRM EN BREF

- Le CIDOC-CRM est une ontologie qui documente le patrimoine matériel et immatériel ainsi que les processus de production de connaissances à son propos.
- <https://www.cidoc-crm.org/>
- Venant du monde des musées, elle est désormais utilisée dans tous les domaines des HN.
- Elle est extrêmement abstraite et générique.
- Ontologie centrée événement (nous y reviendrons dans les exemples...)
- Classes et propriétés : [https://cidoc-crm.org/html/cidoc\\_crm\\_v7.1.2.html](https://cidoc-crm.org/html/cidoc_crm_v7.1.2.html)

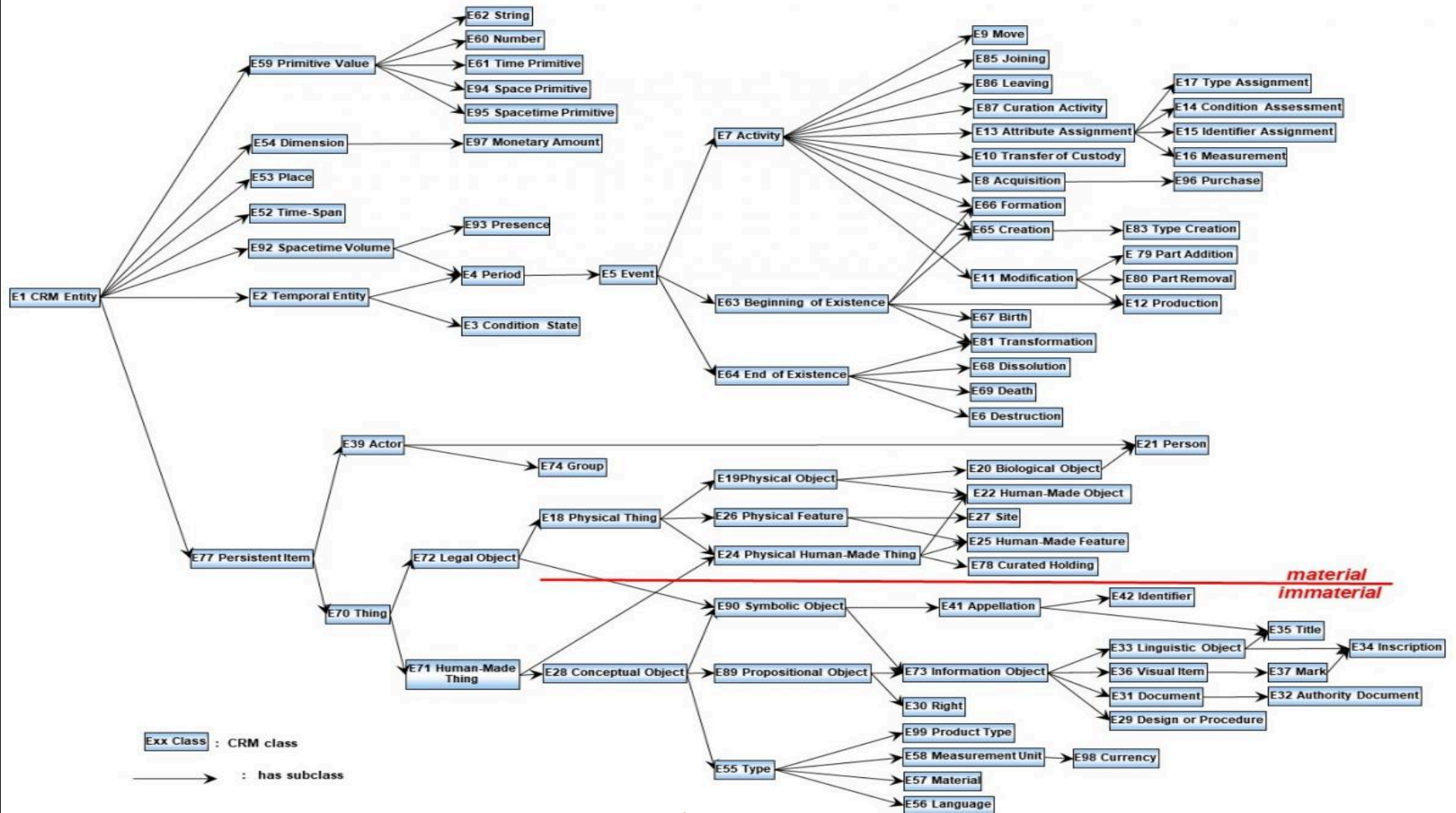


# HIÉRARCHIE DES CLASSES

## CRM Class Hierarchy

*Martin Doerr*

2/16/2020



Opinion : En dépit du nombre de classes centrées sur les usages de musées, le CIDOC-CRM propose des classes génériques permettant de rendre compte de l'ensemble des problématiques de modélisation de la structure et du contenu des sources, ainsi que des processus analytiques qui les prennent pour cible.



# LE TEMPS DANS LE CRM

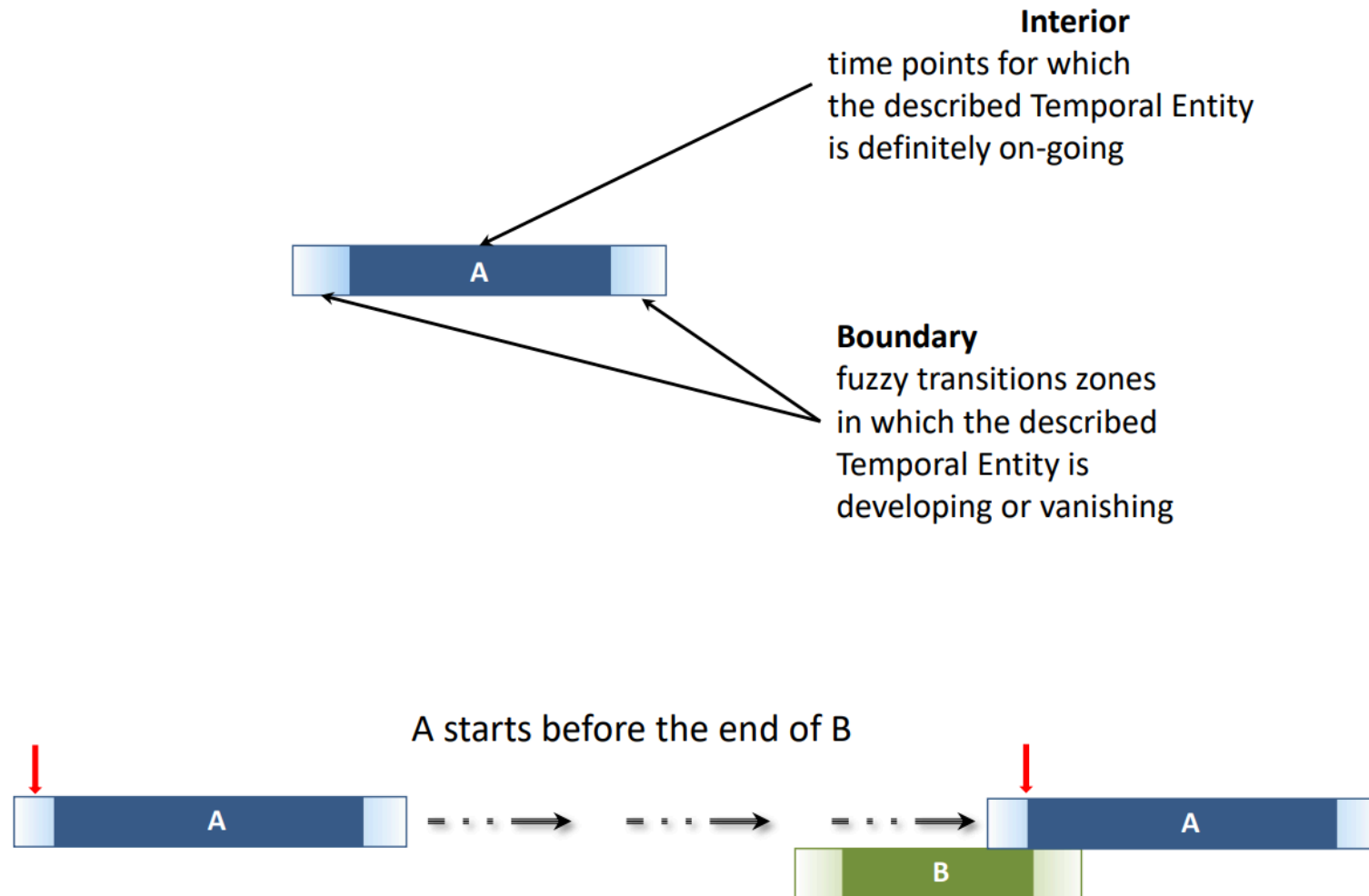


Figure 7: Explanation of Interior and Boundary and an Example of Use from P174 starts before the end of (ends after the start of).

# IV. CAS CONCRETS RÉCURRENTS





# QUE MODÉLISE T-ON ?

- Les sources, et leur organisation logique et matérielle.
- Les faits sociaux, les concepts, les objets matériels, les objets sémiotiques... dénotés ou connotés par les sources et/ou observables sur le terrain.
- Les savoirs nouveaux produits à propos des sources.
- Les ressources analytiques engagées dans l'étude des sources.
- Le contexte de production des savoirs (qu'est ce qui fait qu'une donnée est « scientifique » ? → l'explicitation du qui, du comment, du pourquoi).



# QUE MODÉLISE T-ON ?

- Les sources, et leur organisation logique et matérielle.
- Les faits sociaux, les concepts, les objets matériels, les objets sémiotiques... dénotés ou connotés par les sources et/ou observables sur le terrain.
- Les savoirs nouveaux produits à propos des sources.
- Les ressources analytiques engagées dans l'étude des sources.
- Le contexte de production des savoirs (qu'est ce qui fait qu'une donnée est « scientifique » ? → l'explicitation du qui, du comment, du pourquoi).
- Avoir le CIDOC-CRM en tête lors de l'analyse des pratique de production d'informations scientifiques a des vertus heuristiques.



# CONVENTIONS

Dans les exemples, on utilisera les préfixes suivants pour les URL :

PREFIX `crm:` `<http://www.cidoc-crm.org/cidoc-crm/>`

PREFIX `rdf:` `<http://www.w3.org/1999/02/22-rdf-syntax-ns#>`

PREFIX `m2:` `<http://musica2.huma-num.fr/id/>`

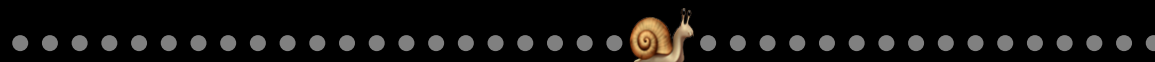
Ainsi, l'URL :

`http://musica2.huma-num.fr/id/tralala`

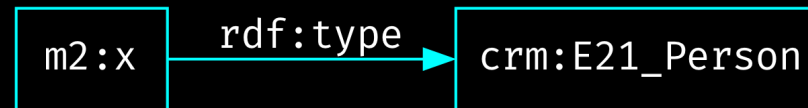
sera notée :

`cm2:tralala`

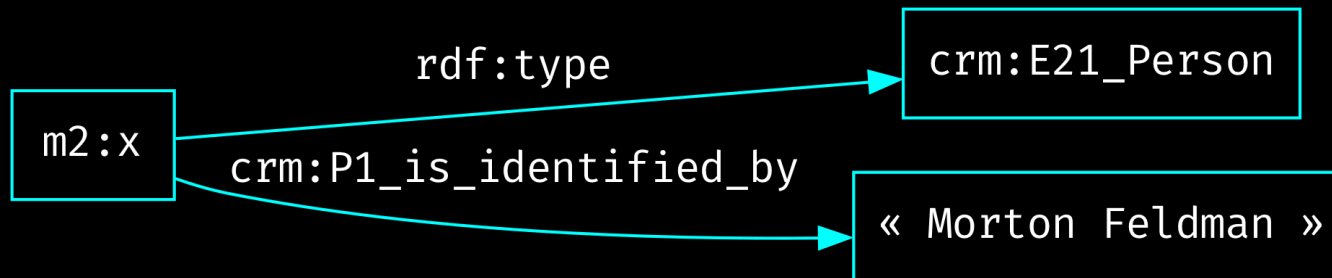
...ce qui évitera d'avoir des exemples illisibles, mais il ne faudra pas perdre de vue que toute ressource (en incluant les classes et les propriétés du CRM) est identifiée par une URI sur le Web.



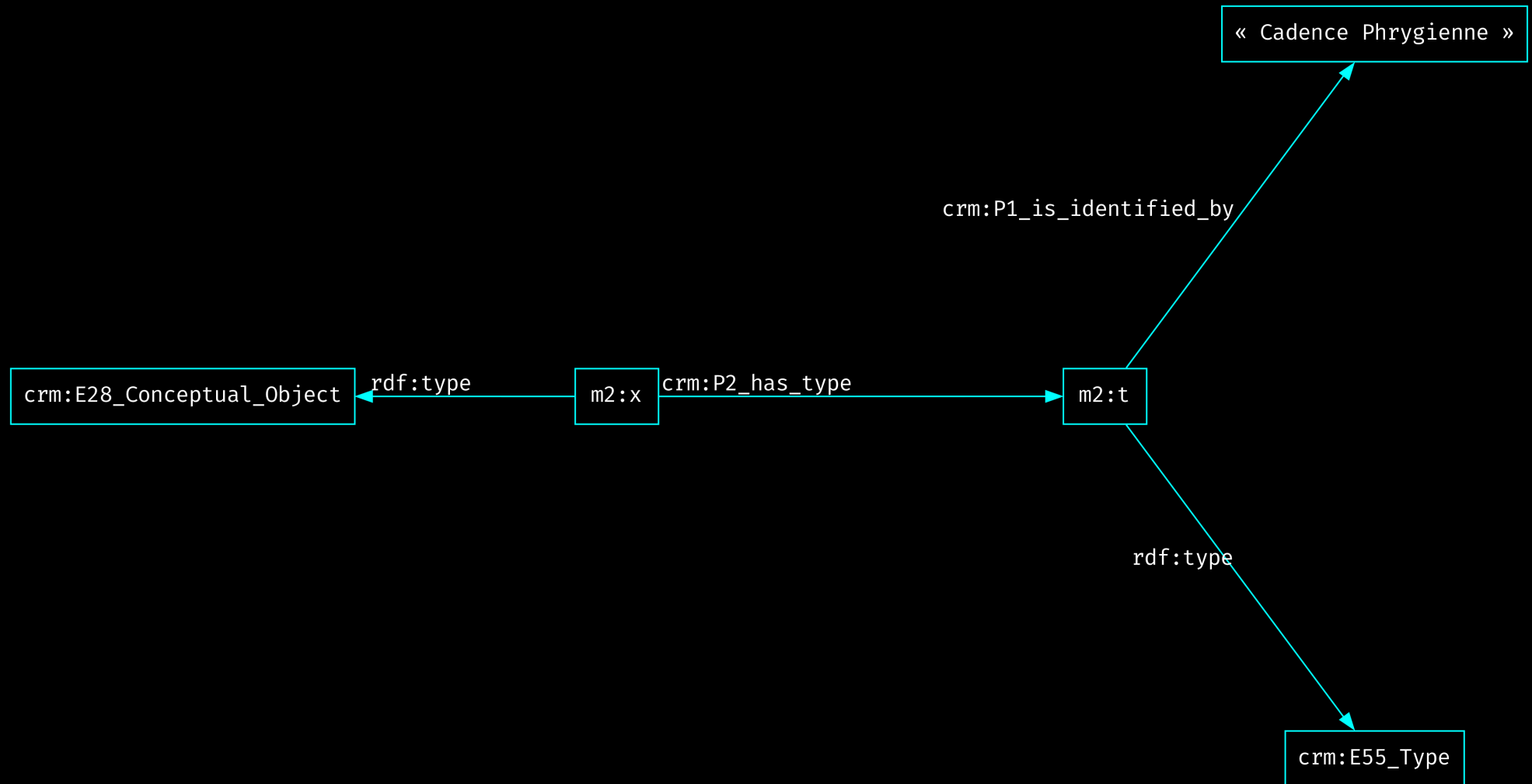
- Typing a person :



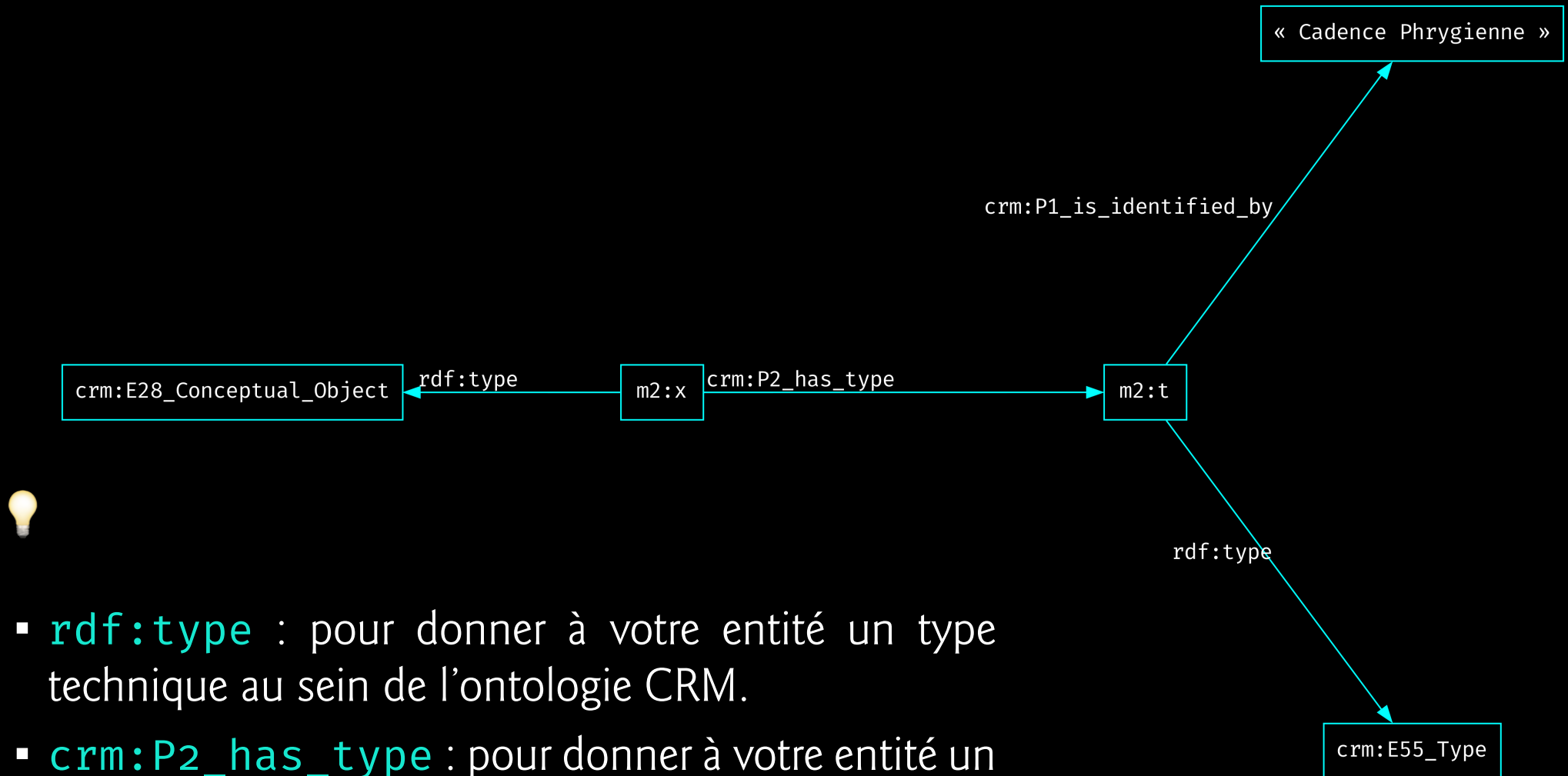
- Naming a person with a simple character string :



- Typer quelque chose :

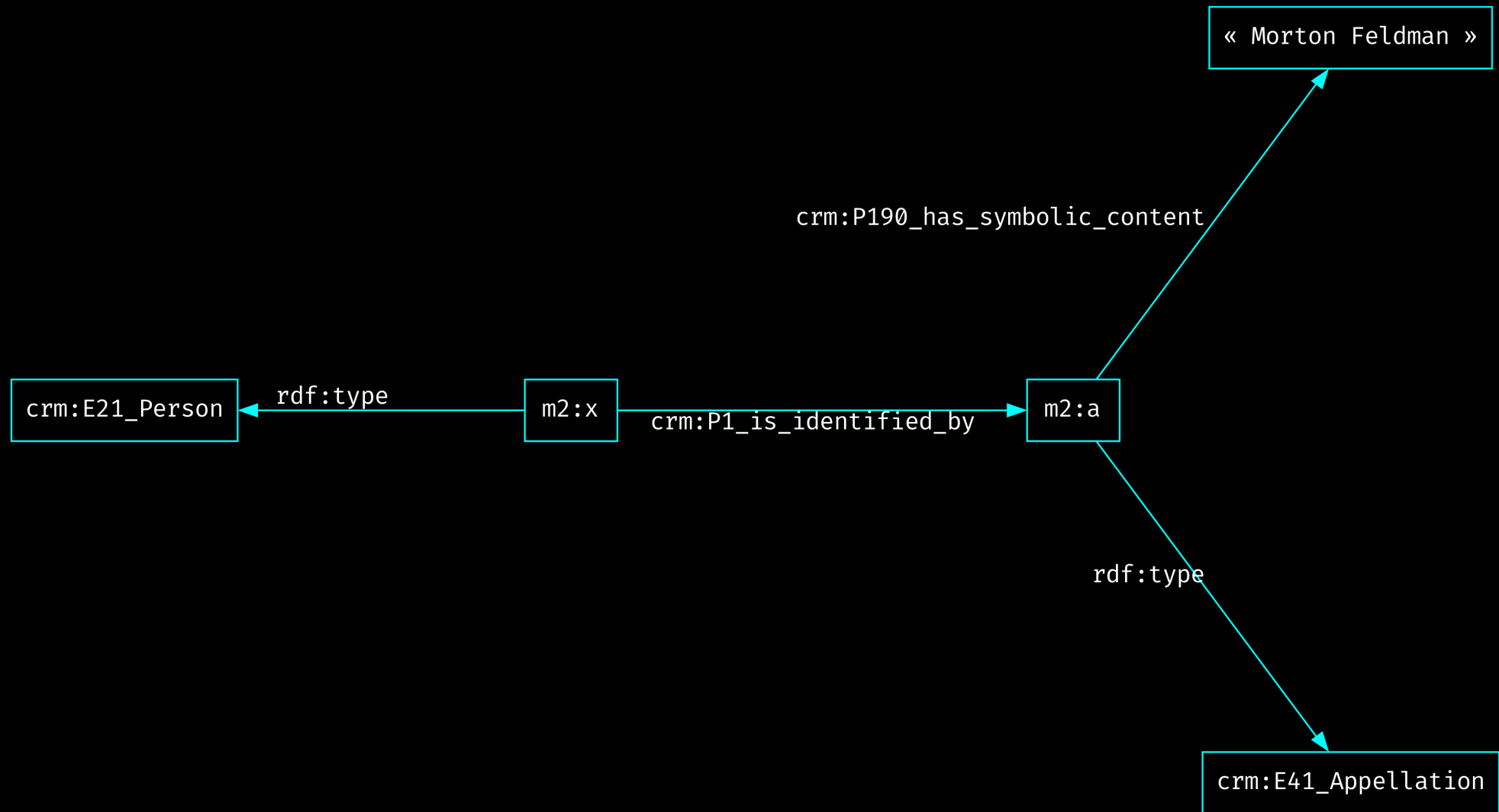


- Typer quelque chose :

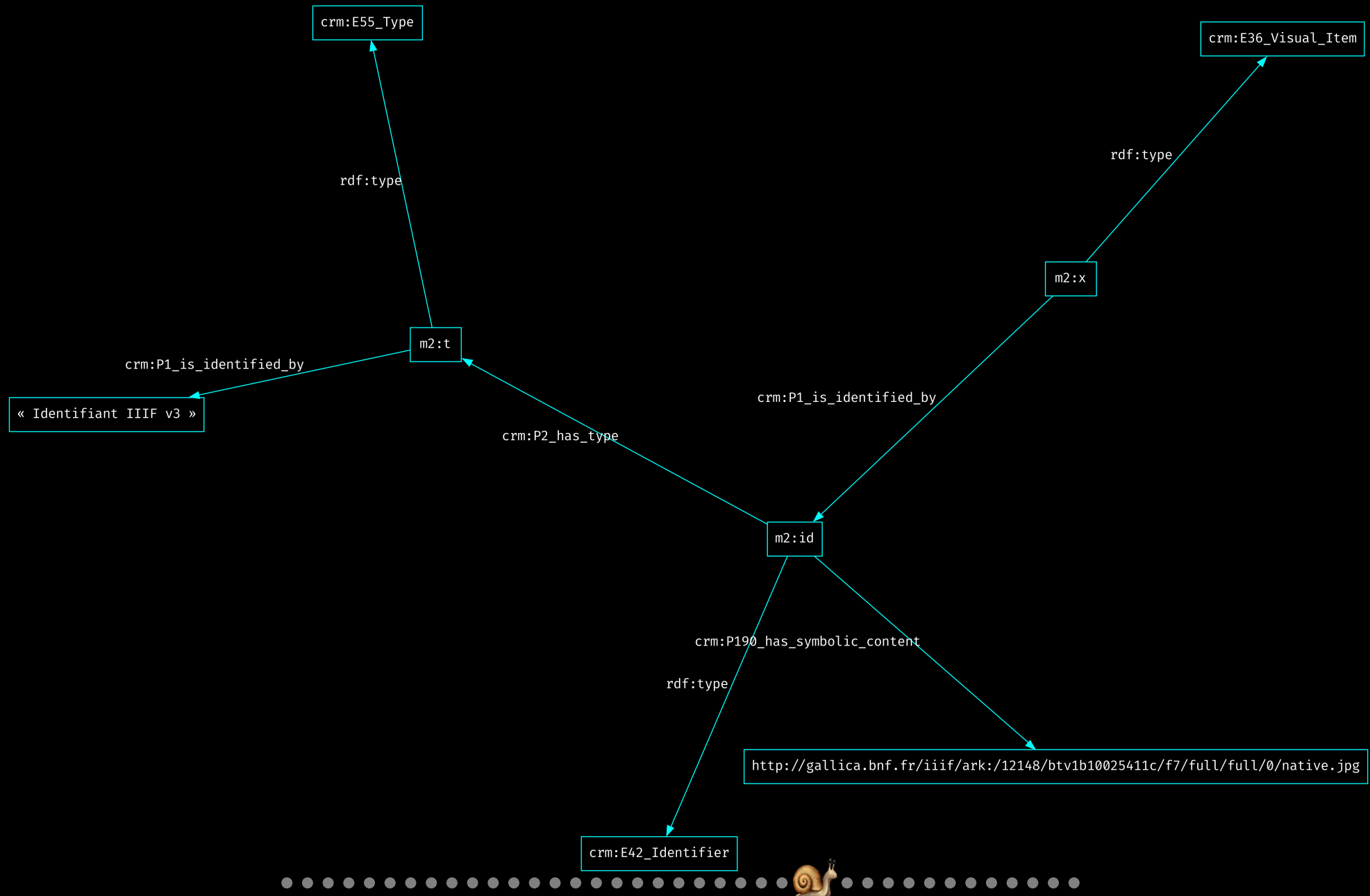


- **`rdf:type`** : pour donner à votre entité un type technique au sein de l'ontologie CRM.
- **`crm:P2_has_type`** : pour donner à votre entité un type métier (défini par vous et en rapport avec les catégories convoquées par l'activité de recherche).

- Nommer une personne avec une entité appellation :



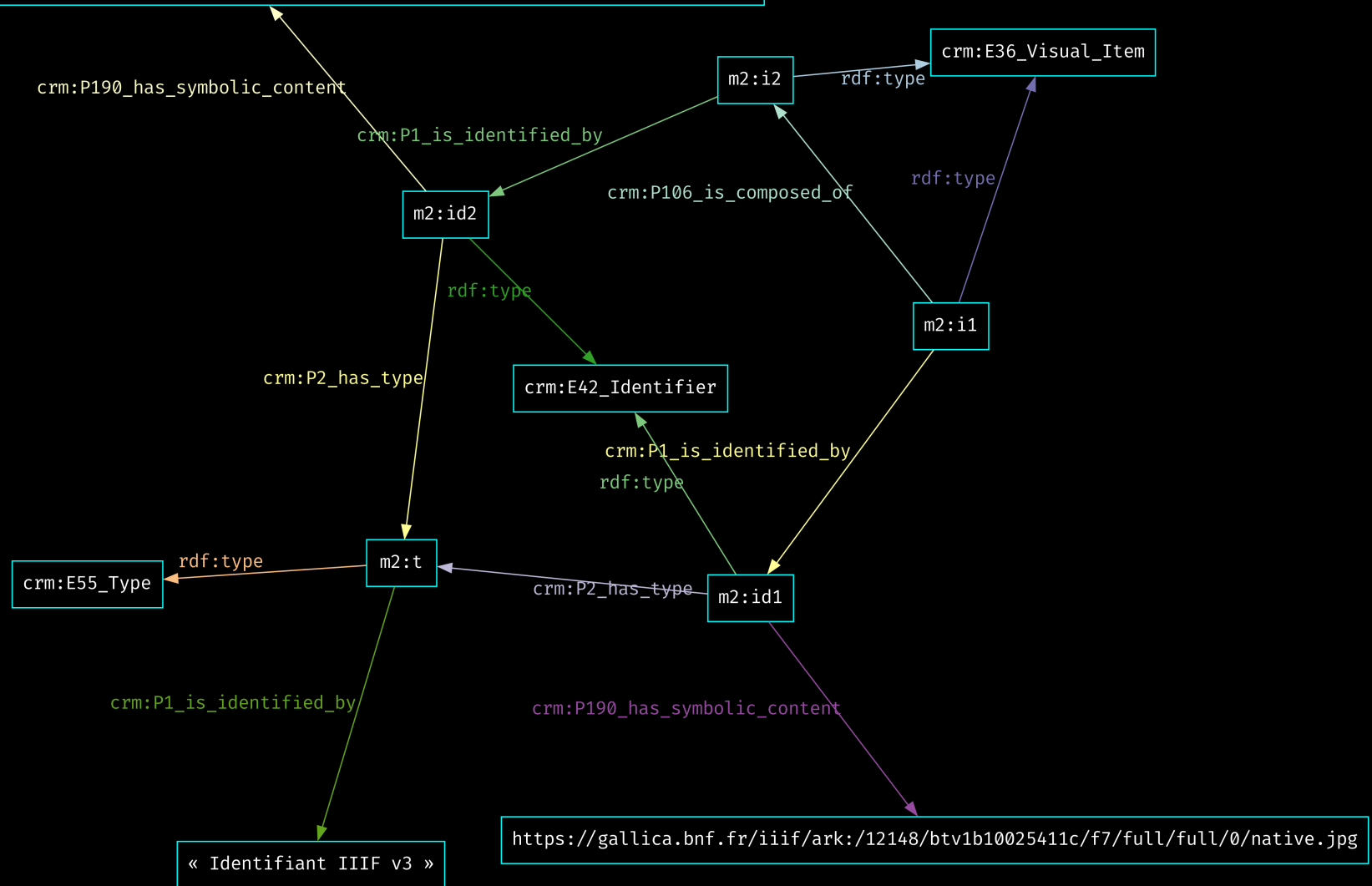
- Typer un identifiant :



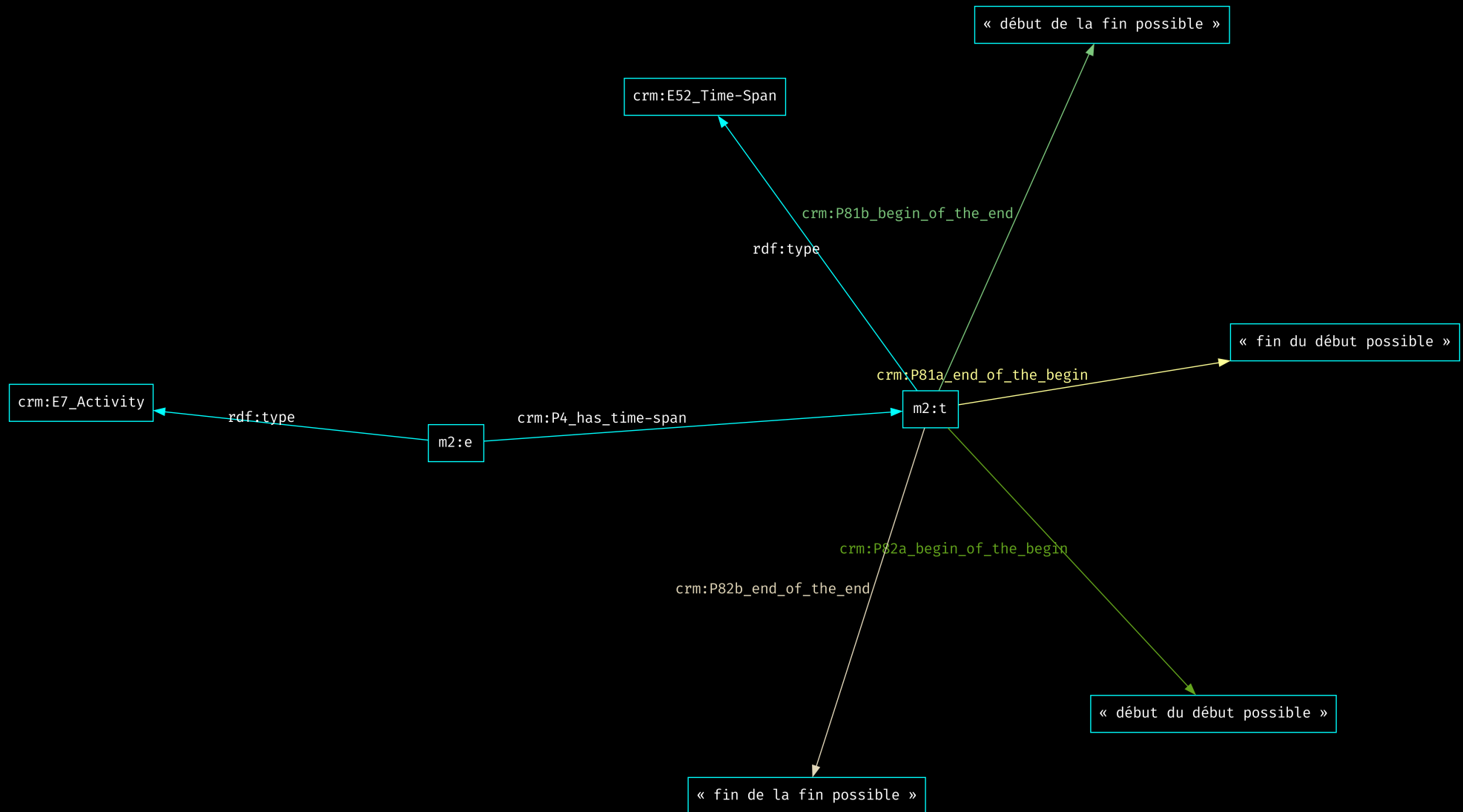


- Organisation interne d'un contenu (ici, d'une image) :

<https://gallica.bnf.fr/iiif/ark:/12148/btv1b10025411c/f7/1000,1000,2000,1000/full/0/native.jpg>

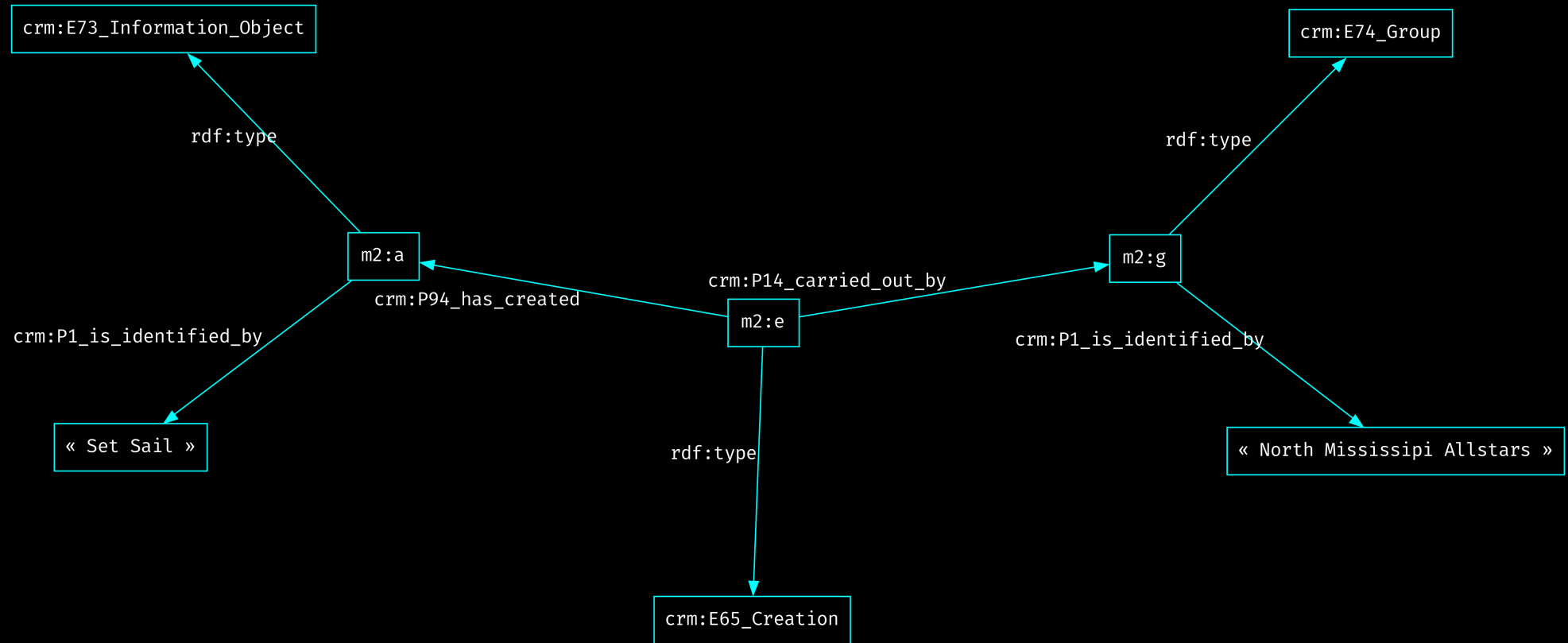


## ■ Dater un événement :

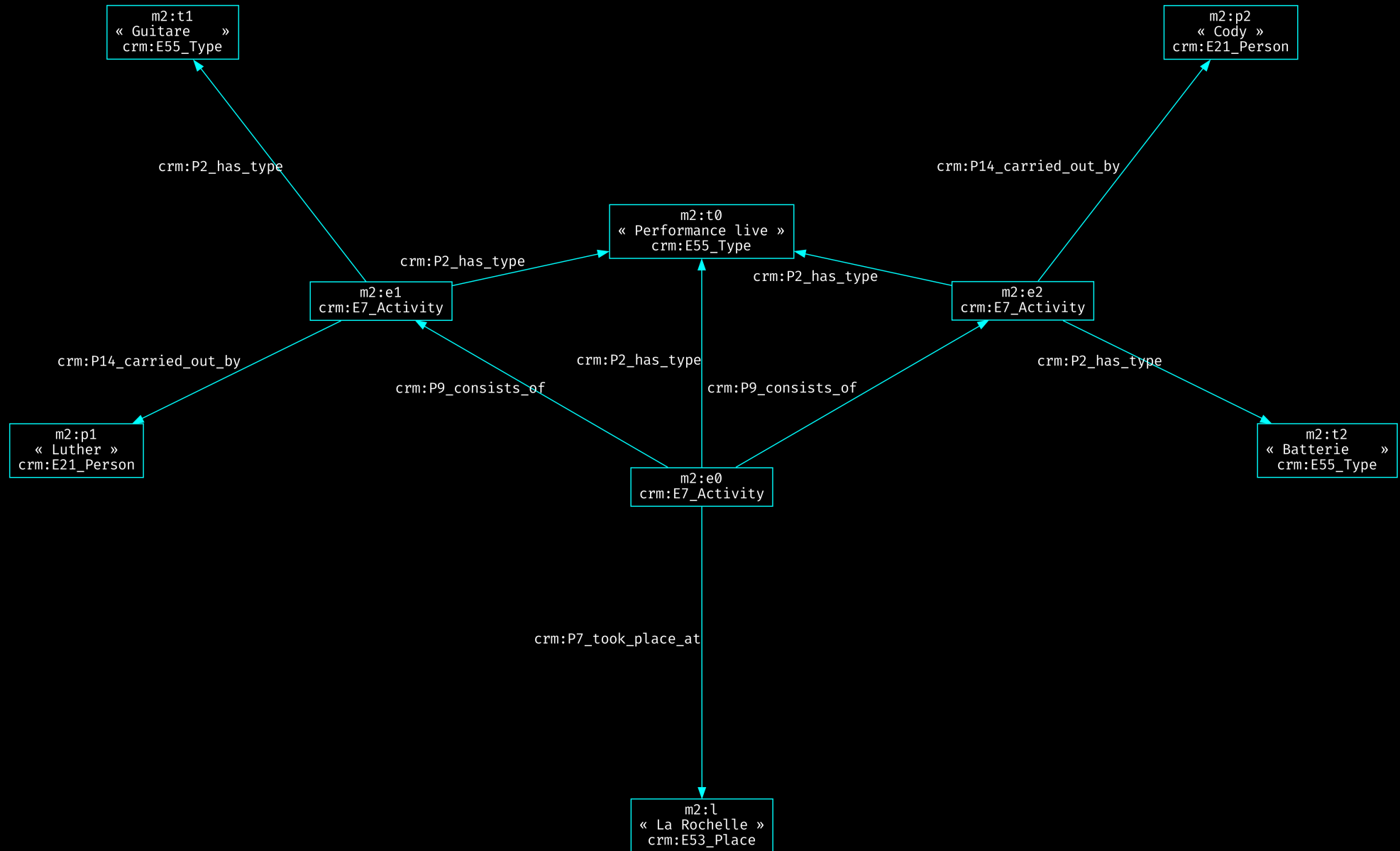


On retrouve la possibilité de définir des incertitudes aux bornes.

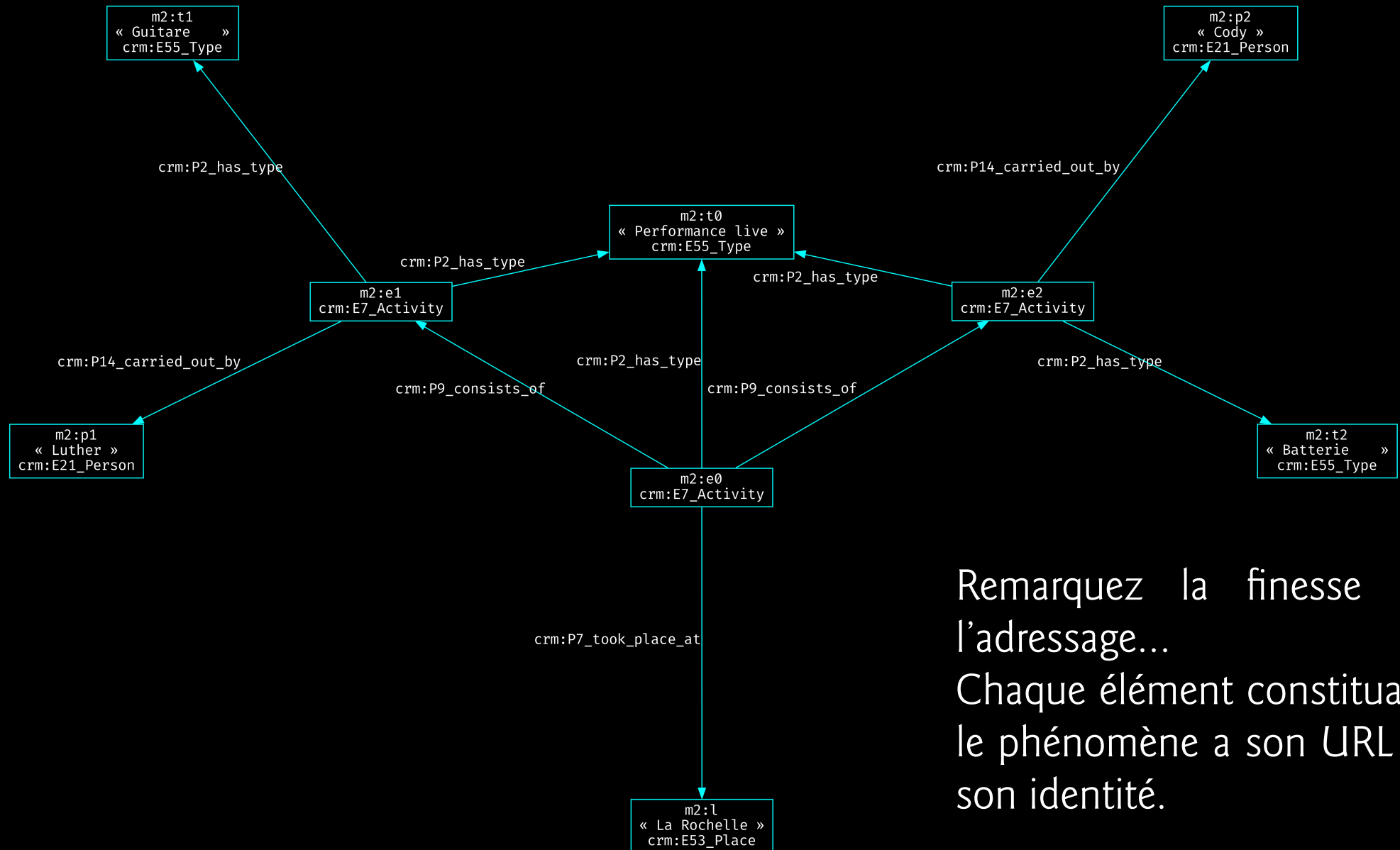
- Exprimer un événement de création simple :



- Exprimer un événement de création complexe :



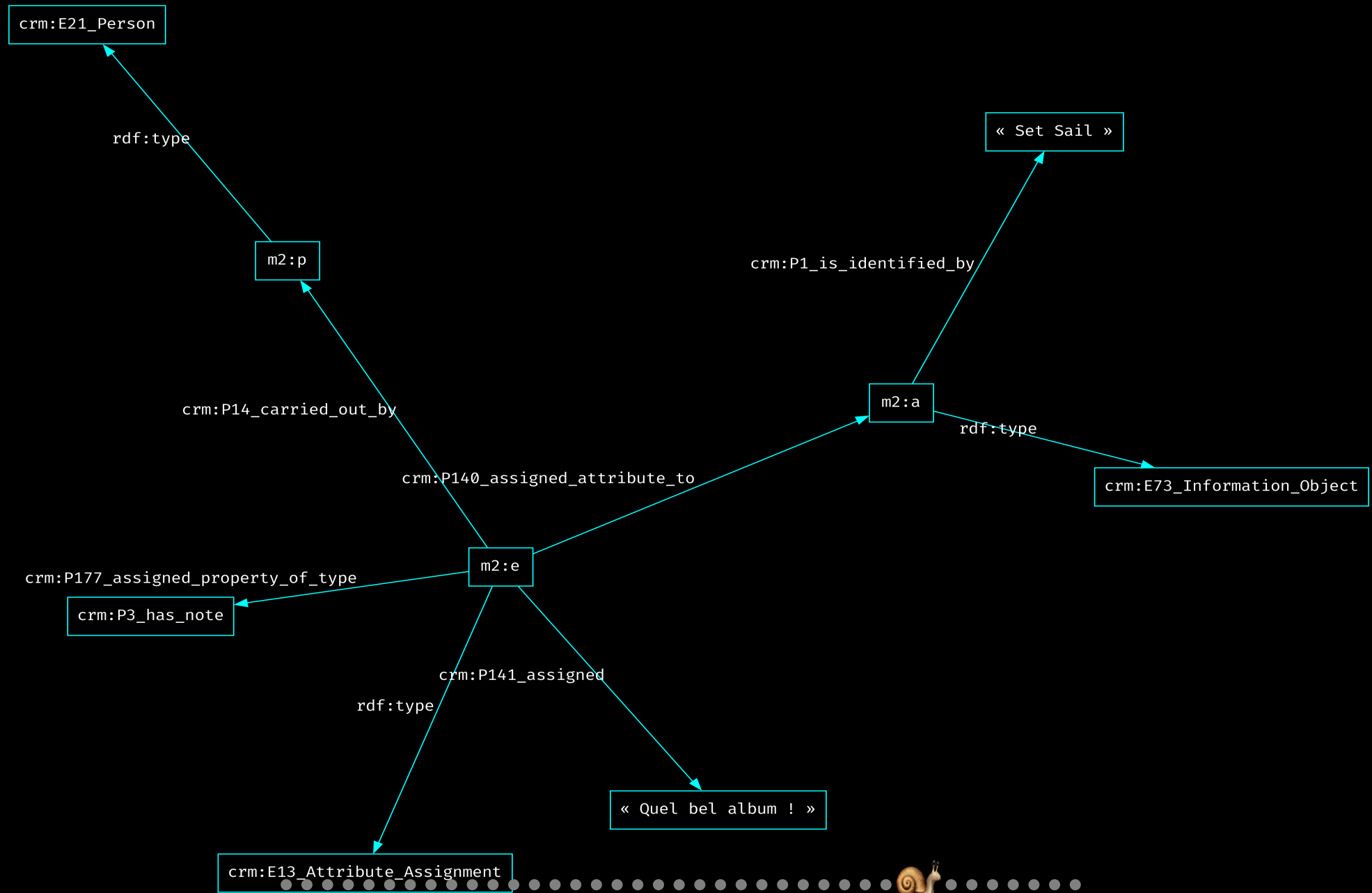
- Exprimer un événement de création complexe :



Remarquez la finesse de l'adressage...  
Chaque élément constituant le phénomène a son URL et son identité.



## ■ Exprimer une annotation :



V. DEUX MOTS SUR LRMOO ET  
DOREMUS



# LRM○○

- Basé sur l'**IFLA LRM** (auparavant FRBR).
- Convergence du monde des musées (CIDOC-CRM) et du monde des bibliothèques (IFLA LRM) : l'ontologie LRMoo spécialise certaines classes du CIDOC-CRM (qui sert donc d'ontologie de base).
- Quatre entités fondamentales — Work, Expression, Manifestation, Item (WEMI) — représentant les quatre modes d'existence des sources (conceptuel, sémiotique, éditorial, matériel).
- **Documentation**





# Entités WEMI : exemple

Un avion sans elle (français). Pocket 2014. 844 BUS

**Œuvre**  
(Work)

Création de  
l'esprit par  
Michel  
Bussi

Le contenu en  
français

**Expression**

Une des  
publications  
commerciales de  
ce texte chez  
Pocket

**Manifestation**

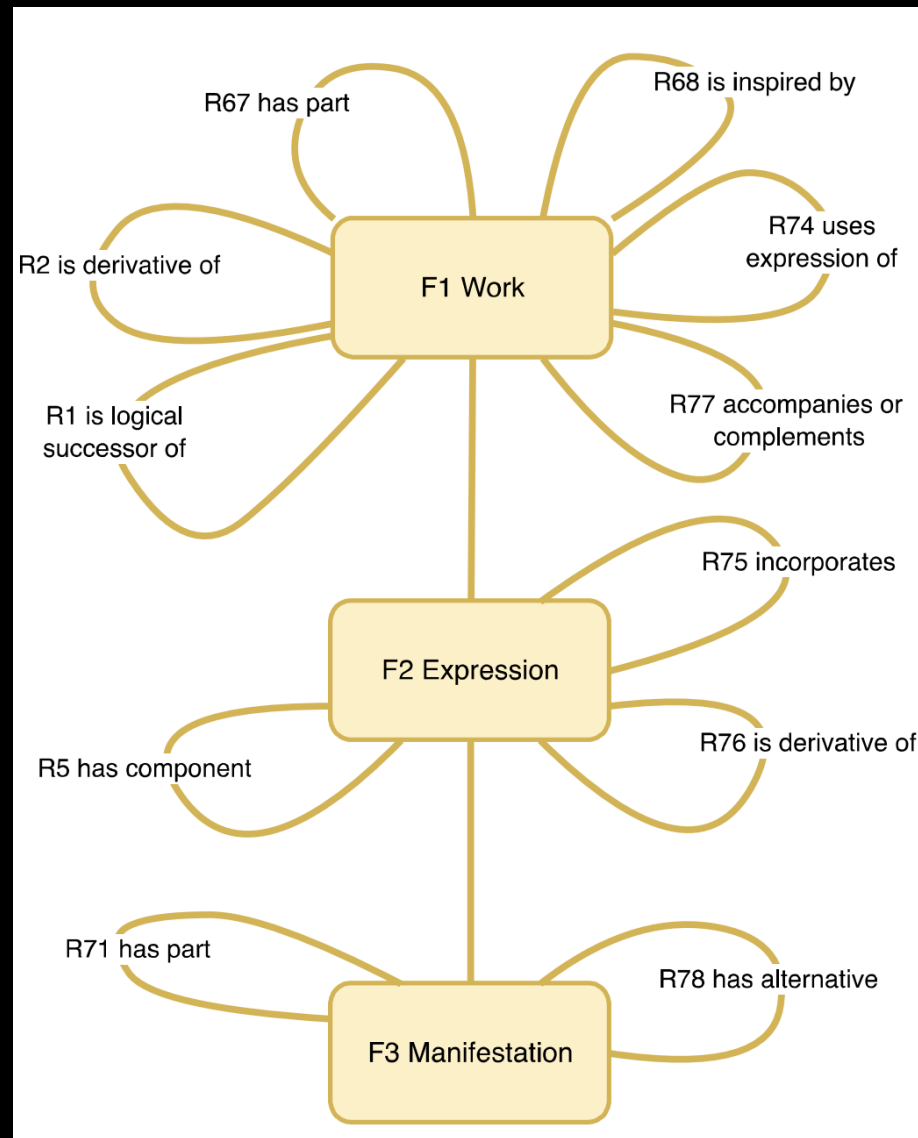


Mon exemplaire  
personnel

Exemplaire  
Médiathèque Valais

**Item**

# DES RELATIONS RICHES ENTRE LES ENTITÉS



# DOREMUS : POINTS SAILLANTS

- Des vocabulaires utiles pour la musique : [data.doremus.org/vocabularies](https://data.doremus.org/vocabularies) + [github.com/DOREMUS-ANR/knowledge-base/tree/master/vocabularies](https://github.com/DOREMUS-ANR/knowledge-base/tree/master/vocabularies)



# DOREMUS : POINTS SAILLANTS

- Des vocabulaires utiles pour la musique : [data.doremus.org/vocabularies](https://data.doremus.org/vocabularies) + [github.com/DOREMUS-ANR/knowledge-base/tree/master/vocabularies](https://github.com/DOREMUS-ANR/knowledge-base/tree/master/vocabularies)
- Une ontologie sans équivalent pour la modélisation des effectifs prévus dans l'œuvre, des concerts, et des effectifs réalisés durant les performances : <https://data.doremus.org/ontology/>

# DOREMUS : POINTS SAILLANTS

- Des vocabulaires utiles pour la musique : [data.doremus.org/vocabularies](https://data.doremus.org/vocabularies) + [github.com/DOREMUS-ANR/knowledge-base/tree/master/vocabularies](https://github.com/DOREMUS-ANR/knowledge-base/tree/master/vocabularies)
- Une ontologie sans équivalent pour la modélisation des effectifs prévus dans l'œuvre, des concerts, et des effectifs réalisés durant les performances : <https://data.doremus.org/ontology/>
- Un parti prix (désormais techniquement obsolète) sur la manière dont on peut modéliser une situation de composition : [data.doremus.org/ontology](https://data.doremus.org/ontology)

# DOREMUS : POINTS SAILLANTS

- Des vocabulaires utiles pour la musique : [data.doremus.org/vocabularies](https://data.doremus.org/vocabularies) + [github.com/DOREMUS-ANR/knowledge-base/tree/master/vocabularies](https://github.com/DOREMUS-ANR/knowledge-base/tree/master/vocabularies)
- Une ontologie sans équivalent pour la modélisation des effectifs prévus dans l'œuvre, des concerts, et des effectifs réalisés durant les performances : <https://data.doremus.org/ontology/>
- Un parti pris (désormais techniquement obsolète) sur la manière dont on peut modéliser une situation de composition : [data.doremus.org/ontology](https://data.doremus.org/ontology)
- Un énorme dataset (BNF + Radio France + Philharmonie) d'œuvres, partitions, concerts... [data.doremus.org](https://data.doremus.org)

# VI. UNE ÉTUDE DE CAS : MODÉLISATION DES PARTITIONS ENCODÉES

