

REGISTRE FIAS DES SPÉCIFICATIONS DE RECHERCHE

Profil minimal de provenance pour l'ingestion de partitions MEI dans une bibliothèque de corpus de type NEUMA

IDENTIFIANT DU REGISTRE

FOAR-CON2X-PJ7ZBHF4J

TYPE D'OBJET

Modèle de données ou format d'échange

MATURITÉ DU DOCUMENT

Candidat pour un usage stable

PUBLIÉE

30 mai 2025, 12:00 UTC

VERSION

1.0

DOMAINE

Musicologie et notation

SOUMISE

14 mars 2025, 12:00 UTC

LICENCE

Creative Commons Attribution 4.0

VERSION FAISANT AUTORITÉ

L'enregistrement consigne le résultat de l'examen du Registre pour cette version : la soumission a été suffisamment identifiée, documentée et déterminée en interne pour la procédure indiquée. La portée, l'adoption et les conventions concurrentes sont documentées séparément lorsqu'il y a lieu.

STATUT AU REGISTRE

Acceptée pour enregistrement

30 mai 2025, 12:00 UTC

Description principale

TYPE DE SOUMISSION

Nouveau concept de registre

TITRE COMPLET

Profil minimal de provenance pour l'ingestion de partitions MEI dans une bibliothèque de corpus de type NEUMA

TITRE COURT OU ACRONYME

NPM-MEI

TYPE D'OBJET

Modèle de données ou format d'échange

STATUT REVENDIQUÉ PAR LE DÉPOSANT

Profil d'une norme externe

MATURITÉ DE CETTE VERSION

Candidat pour un usage stable

VERSION

1.0

DOMAINE PRINCIPAL

Musicologie et notation

MOTS-CLÉS

- musicologie numérique
- partitions
- MEI
- NEUMA
- provenance
- qualité des données
- corpus musicaux

PROBLÈME, PORTÉE ET EXCLUSIONS

Les bibliothèques de partitions numériques peuvent recevoir un fichier MEI techniquement valide sans connaître précisément son document source, la méthode de production, la responsabilité de l'encodage, la licence ou le résultat des contrôles qualité. Le profil définit une enveloppe JSON minimale autour d'un fichier MEI 5.1 pour une ingestion de corpus

traçable. Son périmètre concerne la provenance d'ingestion; notation musicale MEI, exactitude musicologique, catalogue bibliographique, audio seul et données personnelles de musiciens vivants relèvent d'autres cadres.

UTILISATEURS ET CONTEXTES VISÉS

- Ingénieurs de données de bibliothèques de partitions numériques
- Musicologues déposant un petit corpus encodé
- Responsables qualité comparant des lots MEI

RELATION AVEC LA PRATIQUE EXISTANTE

Profile une norme existante

CONTEXTE SCIENTIFIQUE, SOURCES ET ÉCHANGES SCIENTIFIQUES

Cette section décrit les fondements scientifiques, les travaux sous-jacents et les sources de la spécification, ainsi que les groupes de recherche et les cadres scientifiques dans lesquels la solution proposée ou sa formalisation a été présentée ou discutée. Elle documente le contexte et la nature de ces échanges scientifiques.

Les bibliothèques numériques de partitions réunissent des fichiers dont la valeur scientifique dépend de leur relation aux documents sources, des opérations d'encodage et des contrôles réalisés. MEI 5.1 fournit un modèle XML et des mécanismes de métadonnées pour représenter les sources musicales. Les travaux de F. Foscarin, P. Rigaux et V. Thion (2021) sur GioQoso développent l'évaluation de la qualité des bibliothèques de partitions et montrent l'intérêt d'associer aux fichiers des informations sur leur production et leur traitement.

NEUMA articule recherche musicologique, partage de partitions et méthodes informatiques au sein d'une collaboration associant notamment l'IReMus et le CEDRIC au CNAM. Les travaux de P. Rigaux et les contributions musicologiques auxquelles participe H. Audéon situent cette démarche à l'interface entre étude des sources et conception d'une bibliothèque de corpus. Cet environnement de recherche fait de la provenance de l'encodage et de la traçabilité des contrôles des conditions de la comparaison et de la réutilisation des partitions.

NPM-MEI formalise ces informations dans une enveloppe JSON accompagnant les fichiers MEI. Elle relie chaque partition à son document source, à sa méthode de production, aux agents intervenus, aux droits déclarés et aux résultats de validation. L'enregistrement des versions, des empreintes et des rapports permet de retrouver l'état exact d'un fichier lors de son ingestion et de suivre les transformations successives. Le profil fournit ainsi une base commune pour examiner la qualité documentaire des lots échangés entre bibliothèques de corpus.

Références:

- Music Encoding Initiative, MEI Guidelines 5.1: Texte: <https://music-encoding.org/guidelines/>

v5/content/introduction.html

- MEI Guidelines 5.1, dépôt Zenodo du 3 février 2025: Texte: <https://zenodo.org/records/14795519>

- IReMus, projet NEUMA, « Recherche et indexation de sources musicales partagées »:
Texte: <https://www.iremusc.cnrs.fr/node/376>

- Francesco Foscari, Philippe Rigaux, Virginie Thion, "Data Quality Assessment in Digital Score Libraries: The GioQoso Project," International Journal on Digital Libraries 22 (2021), 159-173, DOI 10.1007/s00799-021-00299-7: DOI: <https://doi.org/10.1007/s00799-021-00299-7>

- Sorbonne Université, Institut de recherche en musicologie. Texte: <https://lettres.sorbonne-universite.fr/unites-de-recherche/institut-de-recherche-en-musicologie>

STATUT DE NORMALISATION EXTERNE

Aucune prévue

LICENCE DE LA SPÉCIFICATION SOUMISE

Creative Commons Attribution 4.0

DÉCLARATION DE PUBLICATION

Oui

Modèle de données ou format d'échange

ENTITÉS, ENREGISTREMENTS ET RELATIONS

Entités : CorpusRelease, ScoreItem, SourceDocument, DigitalSurrogate, EncodingFile, EncodingActivity, Agent, ValidationRun et RightsStatement. CorpusRelease agrège un ou plusieurs ScoreItem. Chaque ScoreItem décrit exactement un fichier MEI principal et pointe vers au moins un SourceDocument ou explique l'absence de source. EncodingActivity relie l'agent, la date, la méthode (saisie, OMR, conversion, correction) et le fichier produit. ValidationRun relie un validateur, sa version, une date, un résultat et un rapport conservé.

SÉRIALISATION ET TYPE DE MÉDIA

Un manifeste UTF-8 JSON, recommandé sous le nom manifest.json, accompagne chaque lot. Chaque EncodingFile est un document XML MEI 5.1 séparé. Le manifeste est servi comme application/json ; le fichier MEI comme application/mei+xml lorsqu'accepté par l'infrastructure, sinon application/xml avec suffixe .mei. Les chemins sont relatifs au lot et ne contiennent ni « .. » ni URL avec identifiants incorporés.

CHAMPS OBLIGATOIRES ET CARDINALITÉS

CorpusRelease (1): profile_version[1], release_id[1], released_at[1], title[1], licence[1], items[1..n].
ScoreItem (1..n): item_id[1], title[1], mei_path[1], mei_sha256[1], source[1], encoding_activities[1..n], validation_runs[1..n].
SourceDocument (1): source_type[1], citation[1], shelfmark_or_id[0..1], holding_institution[0..1], surrogate_uri[0..n], source_note[0..1].
EncodingActivity (1..n): activity_id[1], method[1], agent_id[1], started_at[0..1], ended_at[1], software[0..n], note[0..1].
Agent (1..n): agent_id[1], display_name[1], agent_type[1], persistent_id[0..1].
ValidationRun (1..n): validator[1], validator_version[1], run_at[1], result[1], report_path[0..1].

TYPES DE DONNÉES, VALEURS CONTRÔLÉES, VALEURS NULLES ET INCONNUES

Les identifiants internes sont des chaînes ASCII correspondant à `^[A-Za-z][A-Za-z0-9._-]{2,63}$`. Les dates sont ISO 8601/RFC 3339 avec fuseau. Un SHA-256 est une chaîne hexadécimale minuscule de 64 caractères. method appartient à manual, omr, conversion, mixed; agent_type à person, organisation, software; result à pass, fail, warning. Une valeur inconnue obligatoire utilise un objet `{"unknown": true, "reason": "..."}` , jamais null. Un champ optionnel inconnu est omis. La langue d'un titre libre est portée par un objet `{"value": "...", "lang": "fr"}` .

MÉCANISME DE VALIDATION MACHINE

JSON Schema

DÉTAILS DU VALIDATEUR, DU SCHÉMA OU DE LA GRAMMAIRE

La validation du manifeste repose sur JSON Schema Draft 2020-12. Chaque `mei_path` doit exister, rester dans le lot et correspondre à `mei_sha256`. Le fichier MEI est validé séparément contre le schéma officiel MEI 5.1 ; le nom et la version du validateur sont consignés dans `ValidationRun`. Un contrôle applicatif vérifie les références `agent_id` et l'unicité de tous les identifiants internes.

COMPATIBILITÉ ASCENDANTE ET DESCENDANTE

Un lecteur 1.x doit ignorer tout champ inconnu et refuser seulement un changement de version majeure. Les champs obligatoires de 1.0 ne changent pas dans une version 1.x ; un nouveau champ obligatoire exige 2.0. La mise à jour du fichier MEI crée un nouveau `ScoreItem` ou une nouvelle version explicitement liée et un nouveau hachage. La capacité d'un logiciel à rendre tous les éléments MEI 5.1 fait l'objet d'une vérification séparée.

EXEMPLE MINIMAL COMPLET

```
{
  "$schema": "https://example.invalid/npm-mei/1.0/schema.json",
  "profile_version": "1.0",
  "release_id": "example-air-2025-01",
  "released_at": "2025-03-14T10:00:00Z",
  "title": { "value": "Exemple de corpus", "lang": "fr" },
  "licence": "CC-BY-4.0",
  "agents": [
    { "agent_id": "team01", "display_name": "Équipe de l'exemple",
      "agent_type": "organisation" }
  ],
  "items": [{
    "item_id": "score001",
    "title": { "value": "Exemple de partition", "lang": "fr" },
    "mei_path": "scores/score001.mei",
    "mei_sha256":
"0123456789abcdef0123456789abcdef0123456789abcdef0123456789abcdef",
    "source": {
      "source_type": "printed_score",
      "citation": "Exemple, édition de travail, p. 1",
      "source_note": "Exemple de référence à une source."
    },
    "encoding_activities": [{
      "activity_id": "enc01", "method": "manual",
      "agent_id": "team01", "ended_at": "2025-03-10T17:00:00Z"
    }]
  }],
}
```

```
"validation_runs": [{  
  "validator": "jing", "validator_version": "20220510",  
  "run_at": "2025-03-11T09:00:00Z", "result": "pass",  
  "report_path": "reports/score001-validation.txt"  
}]  
}]  
}
```

Historique d'examen du Registre

Soumission reçue

14 mars 2025, 12:00 UTC

Soumission reçue.

Acceptée pour enregistrement

30 mai 2025, 12:00 UTC

Accepté pour enregistrement comme profil minimal d'ingestion lié à la provenance MEI. Le modèle sépare correctement la validité XML/MEI, l'intégrité du fichier, la provenance de l'encodage et l'évaluation musicologique. Cette séparation est cohérente avec les travaux GioQoso sur la qualité des bibliothèques de partitions. La version publiée devra toutefois fournir un hachage correspondant au fichier dans un jeu d'essai exécutable et inclure le JSON Schema annoncé.