



# Music Copyright Infrastructure

---

D5.1 BOOKLET



Funded by  
the European Union



## Fair MusE research and industry partners



# Table of Contents

|                                                                  |    |                                                                                     |    |
|------------------------------------------------------------------|----|-------------------------------------------------------------------------------------|----|
| Why Music Data Matters Today .....                               | 4  | Why EMMI is Different from Past Initiatives .....                                   | 26 |
| How the Digital Music System Really Work .....                   | 5  | What Data Should Be Mandatory: Core Data Set 2.0 and<br>the "Golden Standard" ..... | 27 |
| The Data Behind Music .....                                      | 7  | Core Data Set 2.0: What Changes Compared to Today .....                             | 29 |
| How We Got Here: Why Europe's System Became Fragmented .....     | 9  | How EMMI Would Work: Architecture and Governance .....                              | 30 |
| Why EU Regulation Unintentionally Reinforced Fragmentation ..... | 11 | Implementation Roadmap: a Phased Approach .....                                     | 32 |
| Who Controls Music Data Today .....                              | 12 | Incentives, Compliance and a Regulatory Safety Net .....                            | 33 |
| Publishing Hubs as De Facto Music Data Infrastructures .....     | 14 | Incentives, Soft Law and a Regulatory Backstop .....                                | 34 |
| Where the System Breaks .....                                    | 15 | Benefits for the Music Community .....                                              | 35 |
| Why Performers' Metadata is not Optional Under EU Law .....      | 17 | Looking Ahead: Metadata for AI and Synthetic Uses .....                             | 36 |
| MSS vs Social Media: Two Divergent Data Cultures .....           | 18 | Building a Fairer Digital Music Ecosystem .....                                     | 37 |
| Who Pays the Price .....                                         | 20 | References .....                                                                    | 38 |
| Where The System Works And Where It Fails .....                  | 21 | List of acronyms .....                                                              | 39 |
| What Has Been Tried Before and What Europe Can Learn .....       | 22 |                                                                                     |    |
| The European Music Metadata Infrastructure .....                 | 24 |                                                                                     |    |

## INTRODUCTION

# Why Music Data Matters Today

Music in the digital age is not only sound. It is also data.

Every stream, upload, share or reuse of a track depends on information that identifies who created it, who owns the rights, and how revenues should be distributed. This information, commonly referred to as music metadata, is the invisible infrastructure that supports licensing, reporting and remuneration across the music ecosystem.

Yet, despite the centrality of data, today's digital music markets operate on fragmented and uneven information systems. Rights are documented in separate databases, identifiers are not consistently aligned, and access to high-quality data is unevenly distributed. As a result, transparency is limited, creators struggle to verify how their works are used, and the principle of appropriate and proportionate – therefore 'fair' – remuneration established in EU law is often undermined in practice.

To address these long-standing structural problems, the Fair MusE project convened a Task Force of legal, technical, industry and academic experts in 2025. Their mandate was not merely to diagnose inefficiencies, but to design a realistic framework for improving how music data is produced, shared and governed in Europe.

This booklet presents the outcome of that work in an accessible format. It explains:

- How music data currently functions
- Why the system was born and keeps being fragmented
- Who controls information flows
- How a European Music Metadata Infrastructure could rebalance transparency and accountability and promote fairness.

# How the Digital Music System Really Works

Digital music licensing is built on **three distinct layers of rights and data**. Each layer is governed and managed by different organisations and identified through different infrastructures. For remuneration to work correctly, these layers must be properly interconnected. In practice, as we will see, they are not.

## THE THREE RIGHTS LAYERS

### **Musical works (compositions and songs)**

This layer covers the underlying musical work, namely, the melody, harmony, and lyrics of each musical composition, as fixed through notated music and published (i.e., sheet music). Rights in these purely intangible creative works are typically managed by Collective Management Organisations (CMOs), Independent Management Entities (IMEs), and music publishers. Works are identified through a standard system of identifiers known as 'International Standard Musical Work Code' (ISWC) and linked to creators through party identifiers.

### **Sound recordings**

This layer protects the recording of a performance of a musical work. Rights are usually acquired and held by their respective (i.e., the record labels) and managed in global catalogues. Recordings are identified by the International Standard Recording Code (ISRC), which is widely used across the recording industry and digital platforms.

### **Performances**

This layer relates to the rights of featured and non-featured performers. Because performers typically transfer their exclusive rights in their performances to record producers at the time of recording, performers' organisations have emerged primarily to collect and distribute revenues arising from statutory rights to remuneration. These rights subsist independently of the transfer of performers' exclusive rights and apply to specific uses of recordings identified by law, such as their communication in public venues or via radio and television broadcasting. Performers' organisations have traditionally managed these territorially limited rights using their own separate databases. Standard identifiers for performers were the last to emerge in the music industry, with the introduction of so-called International Performer Numbers (IPNs), which are designed to facilitate cross-border payments among performers' organisations.

Historically, for the reasons outlined above, this layer has been the least documented and the most fragmented.

Each digital use of music normally engages all three layers at once. Appropriate and proportionate remuneration, therefore, depends on the ability to link recordings (ISRCs) to the underlying musical works (ISWCs) and the relevant performers (IPNs) involved in each recording.

### WHO OPERATES IN THE SYSTEM

The music data ecosystem involves multiple categories of actors:

- **Creators:** composers and performers for their works and performances
- **Publishers and record labels:** manage their catalogues and commercialise their respective rights
- **CMOs and IMEs:** collectively represent rightholders and license uses of their works by collecting and distributing royalties to them
- **Music streaming services (MSSs):** Spotify, Apple Music, Deezer, etc.
- **User-generated content platforms (UGC platforms):** YouTube, TikTok, Instagram, Facebook, etc.
- **Distributors and data intermediaries:** companies that ingest, match and enrich music metadata, monitoring usage and reconciling datasets

Each actor in this ecosystem holds or generates part of the information necessary to identify music uses and translate them into data for estima-

ting the use of recordings, musical works, and performances. However, no single actor controls or sets standards for usage metadata or, at the other end of the spectrum, for the full chain of music metadata interconnecting works, recordings, and performances.

#### **The recorded-music layer: why it is more structured**

The main reason the recording layer is easier to track than other layers is that it developed earlier forms of international coordination. The sector has invested in shared technical practices that support large-scale exchange of usage and metadata and developed a global standard identifier: the ISRC.

In practice, global recording catalogues and reporting formats are more consistent across markets than their equivalents for musical works or performances because usage reports from digital platforms refer directly to recordings and their ISRCs.

D5.1 highlights that this is not “perfect data”, but a comparatively more standardised environment, supported by industry-wide interoperability efforts and common specifications.

# The Data Behind Music

Behind every digital music use lies a complex layer of information that determines how a song or recording is identified, licensed and paid for. While listeners interact only with tracks and playlists, the functioning of the entire digital music economy depends on how accurately this underlying data is created, stored and exchanged.

Understanding what music metadata is, and why it plays such a central role in rights management, is essential to grasp why today's system produces errors, opacity and unfair outcomes for many creators.

## WHAT IS MUSIC METADATA?

Music metadata includes all information needed to describe a piece of music and its rights:

- **Identifiers for recordings** (ISRC)
- **Identifiers for musical works** (ISWC)
- **Identifiers for creators and performers** (IPI, ISNI, IPN)
- **Ownership shares and licensing arrangements**
- **Usage data:** how often, where and by whom a track is streamed, downloaded or reused

This data enables platforms to report usage, allows rightholders to track exploitation, and forms the basis for royalty allocation.

## WHY ACCURATE METADATA IS ESSENTIAL

Reliable and interoperable metadata is necessary to:

- Link recordings to the underlying musical works and performances
- Identify all relevant rightholders
- Allocate revenues to the correct parties
- Enable creators to verify how their music is used
- Support the enforcement of EU transparency and remuneration obligations

When metadata is incomplete, inconsistent or inaccessible, digital uses cannot be reliably matched to rights.

This results in delayed payments, misallocations and so-called "black box" revenues that cannot be distributed.

## TIMING AND STRUCTURAL MISALIGNMENT

A key structural problem in the digital music ecosystem lies in the fact that different types of data are created:

- Recording identifiers (ISRCs) are assigned before release.
- Work identifiers (ISWCs) are often assigned after use, once rights information has been processed by collecting societies.
- Performer information is frequently incomplete or added only retrospectively.

This sequencing mismatch makes it difficult to establish early and reliable links between recordings, works and performances. As a result, platforms' usage reports are often based on recording data alone, and the task of "pairing" recordings with the correct works and performers is shifted downstream to CMOs, publishers and data intermediaries.

### Intermediaries: the "patch layer" of the current system

Because the music data environment is fragmented, many actors rely on intermediaries to fill gaps: matching recordings to works, enriching missing metadata, tracking usage across platforms, and reconciling conflicting datasets. In other words, intermediaries often provide the operational fixes that the underlying infrastructure does not deliver by default.

D5.1 treats this as a structural sign: when core systems do not interoperate, the market creates a parallel layer to make things work. This can improve functionality, but it also raises questions about accountability, uneven access, and cost allocation. If only some actors can afford high-quality reconciliation and monitoring, the system can reproduce information asymmetries and reinforce unequal bargaining positions.

**No data,  
no pay.**

**Without accurate metadata,  
fair remuneration is impossible.**

# How We Got Here: Why Europe's System Became Fragmented

Today's fragmented music data landscape is not the result of technical failure alone. It is the outcome of historical choices, regulatory design and market dynamics that shaped how rights were managed long before digital distribution made interoperability essential.

## **COLLECTIVE MANAGEMENT AND TERRITORIAL LOGICS**

For much of the twentieth century, composers' and publishers' rights were administered through national Collective Management Organisations (CMOs). Each society maintained its own repertoire database and licensed uses within a defined territory. Information about works and rightholders was exchanged between societies through bilateral agreements rather than through a shared, central system.

This model functioned reasonably well in the analogue era, when licensing was territorial and usage volumes were manageable. However, it produced structurally separate national datasets that were never designed to operate as a single, integrated information environment.

When digital distribution removed geographic boundaries, the underlying data architecture remained territorial. This mismatch between global exploitation and national data systems became a central source of fragmentation.

## **THE RISE OF MULTI-TERRITORIAL LICENSING AND DATA-DRIVEN MANAGEMENT**

In response to cross-border digital services, the EU promoted multi-territorial licensing of rights in musical compositions, notably through the Collective Rights Management (CRM) Directive (2014). This policy shift required organisations to demonstrate technical capacity to process large volumes of usage data and to manage rights across multiple territories.

As a result, the licensing of musical works became increasingly data intensive. Rights management entities began to operate not only as collecting bodies but also as data processors, building proprietary infrastructures to handle complex reporting from platforms.

This transition accelerated differentiation within the sector: organisations with advanced technological capacity expanded their role, while others struggled to keep pace.

### **INDEPENDENT MANAGEMENT ENTITIES AND REGULATORY ASYMMETRIES**

Alongside CMOs, Independent Management Entities (IMEs) emerged as commercially driven alternatives for rights management. Unlike CMOs, which operate under collective governance and detailed regulatory obligations, IMEs function in more lightly regulated environments that vary across Member States.

While IMEs can offer flexibility and customised services, their growth has added further layers of data fragmentation. Repertoires are increasingly managed across parallel systems, each with its own metadata practices and access conditions. This regulatory asymmetry complicates the creation of shared standards and weakens the overall coherence of the data ecosystem.

### **COMPETITION, PLURALISM AND THE LIMITS OF CENTRALISATION**

European policy has consistently favoured pluralism and competition in

rights management rather than centralisation. This approach reinforced diversity of actors but also entrenched decentralised data infrastructures.

Earlier attempts to establish a single, authoritative repertoire database, most notably the Global Repertoire Database (GRD), failed due to:

- Conflicting governance interests
- Disputes over data ownership and control
- Technical challenges in reconciling heterogeneous datasets

The lesson is not that coordination is unnecessary, but that purely centralised, privately governed solutions are politically and institutionally unsustainable in the European context.

# Why EU Regulation Unintentionally Reinforced Fragmentation

## Competition, pluralism and unintended data silos

In the past two decades, EU digital policy has consistently fostered pluralism, innovation, competition and decentralisation in collective rights management. The Collective Rights Management Directive strengthened reporting and transparency obligations for CMOs while encouraging competition among them in the multi-territorial licensing of rights in musical works for digital uses.

While this approach promoted efficiency and rightholders' freedom of choice, it also had unintended consequences for data governance. CMOs, Independent Management Entities and private licensing hubs developed parallel repertoire databases, each of which found different solutions to pair musical work identifiers (ISWCs) with the recording ones (ISRCs), with a subsequent fragmenta-

tion of data infrastructures.

As regards musical works, despite the improvement of CISAC's CIS-Net portal for CMOs' and IMEs' access to each other's repertoire databases, the European Union ended up with multiple high-quality but disconnected data infrastructures where no cross-sectoral data-sharing standard applies.

Fragmentation is therefore not a technical accident, but a predictable outcome of regulatory design choices that prioritised market pluralism and competition among rights licensors, without addressing interoperability at the infrastructure level.

# Who Controls Music Data Today

In the contemporary digital music market, control over metadata is no longer a purely technical issue. It has become a strategic asset that shapes bargaining power, market access and the distribution of revenues.

## **PUBLISHING AND LICENSING HUBS**

Over the past decade, large-scale licensing and data-processing hubs have emerged to manage musical works across multiple territories. These hubs aggregate repertoires from major publishers and collecting societies and process high volumes of platform usage data.

By concentrating rights administration and metadata management, these entities have:

- Built extensive, high-performance databases
- Developed advanced matching and reporting capabilities
- Gained significant influence over data standards and licensing workflows

This concentration has improved efficiency for major repertoires but has also created new asymmetries between large and small actors.

## **METADATA AS A SOURCE OF MARKET POWER**

The Task Force identified control over data infrastructures as a form of market power. Actors who operate sophisticated databases and linking systems are better able to:

- Match recordings with works and performers
- Process platform reports accurately and at scale
- Audit usage and resolve disputes

By contrast, organisations without access to such infrastructures depend on intermediaries or external services to interpret usage data. This reinforces structural inequalities in the market.

## **TRANSPARENCY AND ACCOUNTABILITY GAPS**

Despite their central role, major licensing hubs and data platforms are often not subject to the same transparency and governance obligations as CMOs. They are neither fully embedded in collective governance structures nor clearly regulated as independent management entities.

As a result:

- The conditions under which they manage and process external repertoires are not always publicly visible
- Smaller organisations may have limited insight into how their data is handled
- Decisions about data standards and matching practices are effectively made by a small number of powerful actors

This creates a system in which access to high-quality metadata is uneven, and accountability for data accuracy is difficult to enforce.

## **THE ROLE OF INTERMEDIARIES AND DISTRIBUTORS**

Because no single layer of the ecosystem provides complete and interoperable data, a growing number of intermediaries have emerged to fill structural gaps. These include:

- Data-matching and monitoring companies
- Digital distributors and aggregators serving independent and DIY artists
- Analytics providers that reconcile usage across platforms

These actors play a crucial functional role, but their services are typically proprietary. This means that improved data quality is often available only to those who can afford it, further entrenching unequal access to information.

# Publishing Hubs as De Facto Music Data Infrastructures

## Publishing hubs: private solutions with public impact

Over the past decade, large-scale publishing and licensing hubs have emerged to manage musical works and process usage data across multiple territories. These entities aggregate repertoires from major publishers and collecting societies and operate high-performance meta-data infrastructures capable of matching and processing vast volumes of platform reports.

In practice, these hubs already perform many of the functions that earlier initiatives, such as the Global Repertoire Database, sought to achieve. They link works data to recordings at scale, resolve claims, and determine how usage information is translated into remuneration. However, unlike public or collectively governed infrastructures, these hubs are privately operated and selectively accessible.

Crucially, most publishing hubs are neither Collective Management Organisations nor Independent Management Entities under EU law. As a result, they are not subject to the transparency, governance and accountability obligations imposed by the Collective Rights Management Directive. This creates a structural asymmetry: entities that play a quasi-infrastructure role in the music market are not governed as infrastructures.

The existence of these hubs shows that centralisation is already happening in practice, but without public oversight, shared standards or equal access. Any European solution must therefore address not only fragmentation, but also the governance gap created by the rise of private data infrastructures.

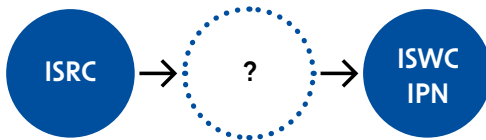
# Where the System Breaks

The fragmentation of Europe's music data ecosystem produces recurring failures that affect how rights are identified, licensed and remunerated. These failures occur at the points where data from different rights layers should connect but does not.

## MISALIGNMENT BETWEEN IDENTIFIERS

The digital music economy relies on multiple identifiers, each developed for a specific purpose:

- **ISRC** identifies sound recordings
- **ISWC** identifies musical works
- **IPN** identifies performers



In theory, these identifiers should be linked to reflect the fact that a single recording embodies a musical work and one or more performances. In practice, such links are often missing or created only after use.

Because usage reports from platforms are primarily ISRC-based, the absence of reliable ISRC-ISWC-IPN links makes it difficult to determine which authors and performers are entitled to remuneration for a given use.

## TEMPORAL GAPS IN DATA CREATION

A major structural problem lies in the sequencing of metadata production:

- Recording data is created and delivered to platforms before release.
- Work and authorship data is often completed only after release and initial exploitation.
- Performer data is frequently incomplete and added retrospectively, if at all.

This temporal gap shifts the burden of data reconciliation downstream, where CMOs, publishers and intermediaries must reconstruct links after the fact. The later data is added, the more costly and unreliable the process becomes.

### **INCOMPLETE PERFORMER DOCUMENTATION**

Performer information represents the weakest link in the chain. Many recordings lack comprehensive data on featured and non-featured musicians, particularly for older catalogues.

As a result:

- Performers' statutory remuneration rights are difficult to enforce
- Additional remuneration mechanisms established under EU and national law cannot operate effectively
- Performers' organisations invest significant resources in reconstructing historical data with limited returns

This structural under-documentation reflects decades of studio practices that did not prioritise performer metadata.

### **Who creates errors vs Who pays to fix them**

#### **Responsibility without incentives**

One of the Task Force's key findings is that responsibility for data accuracy is diffuse, while incentives to invest in quality are uneven.

Errors introduced early in the chain often remain invisible to those who create them but generate costs for actors downstream. Smaller publishers, performers' organisations and independent artists are frequently required to correct data they did not produce in order to recover relatively small amounts of revenue.

Meanwhile, actors with greater market power can continue operating even when data quality is imperfect.

# Why Performers' Metadata is not Optional Under EU Law

## Performers' data as a legal requirement

The inclusion of performers' metadata is not merely a matter of fairness or best practice. It is required to give effect to existing EU and national remuneration rights.

EU law establishes specific remuneration mechanisms for performers, including additional remuneration linked to the extension of the term of protection for sound recordings. Several Member States also provide statutory or collectively managed remuneration rights for performers in digital uses.

Without reliable links between recordings and performer identifiers, these rights cannot be enforced in practice. Establishing ISRC-IPN linkages is therefore not optional. It is a technical prerequisite for compliance with existing legal obligations.

## Performers' rights: the most operationally burdened layer

Performers' rights have historically been documented later and less consistently than works and recordings. This creates a practical problem: even when a recording and a musical work are well identified, the performer information may still be incomplete, fragmented, or stored in separate data systems.

D5.1 explains that this is not only a technical issue. It affects how remuneration flows are verified, how performers can validate usage, and how costs are distributed across the chain. Where documentation is incomplete, money can accumulate in ways that are hard to allocate transparently, and performers may face additional administrative barriers to claim what they are entitled to.

# MSS vs Social Media: Two Divergent Data Cultures

Digital music exploitation takes place across two very different types of platforms, each with its own data culture and reporting practices.

## **MUSIC STREAMING SERVICES: STRUCTURED REPORTING**

Music streaming services (MSSs) operate within established licensing frameworks and rely on structured data flows. They generally:

- Ingest metadata using industry standards
- Report usage through structured formats
- Base remuneration on identifiable recordings

While not without flaws, this model allows for a relatively high level of auditability and predictability in royalty distribution.

## **USER-GENERATED CONTENT PLATFORMS: OPAQUE SYSTEMS**

User-generated content (UGC) platforms developed under different legal and technical conditions. Historically shielded by copyright liability exemptions, they built proprietary content identification systems designed to manage user uploads rather than professional music catalogues.

Key characteristics include:

- Heavy reliance on user-supplied metadata, which is often inaccurate
- Limited use of standard music industry identifiers, at least at the beginning of the social media industry
- Reporting practices that are less granular and harder to audit

Even after the introduction of EU rules on platform liability, these platforms remain the most opaque part of the music data ecosystem.

## **CONSEQUENCES FOR TRANSPARENCY AND REMUNERATION**

The coexistence of two divergent data cultures has concrete consequences:

- Works and performances used on social media are harder to identify and track
- Flat-rate or blanket arrangements replace detailed usage reporting
- Creators and performers face greater difficulty in verifying remuneration

This divergence undermines the effectiveness of EU transparency and remuneration provisions and reinforces asymmetries between platforms and rightholders.

# Who Pays the Price

The fragmentation of music data does not affect all actors equally. Instead, it redistributes costs and risks across the ecosystem in ways that systematically disadvantage weaker participants.

## STRUCTURAL WINNERS AND LOSERS

The Task Force identified a clear asymmetry:

- **Well-resourced actors** (major labels, major publishers, large platforms and advanced licensing hubs) can absorb data imperfections. Their scale, technical capacity and bargaining power allow them to maintain stable revenue flows even when metadata is incomplete or mismatched.
- **Weaker actors** (independent publishers, smaller CMOs, performers' organisations and 'Do It Yourself' or DIY artists) depend on accurate data but lack the resources to control or correct it.

As a result, those who most need reliable metadata are often those with the least ability to influence how it is produced and shared.

## RESPONSIBILITY SHIFTED DOWNSTREAM

Errors and omissions are often introduced early in the chain, during recording sessions, release preparation or initial metadata submission. However, the responsibility for fixing these errors is typically shifted downstream. In practice:

- Smaller organisations invest time and money to correct data they did not create
- Performers' organisations reconstruct missing documentation decades after recordings were made
- Independent creators must navigate opaque systems to claim relatively small sums

This creates a structural imbalance between effort and reward.

## THE ACCUMULATION OF "BLACK BOX" REVENUES

When usage cannot be matched to identifiable rights holders, revenues accumulate without being distributed. These "black box" sums are often redistributed according to proxy mechanisms or retained within systems until claims can be resolved. This situation:

- Reduces transparency
- Undermines trust in licensing and remuneration systems
- Disproportionately affects creators and performers with limited visibility

### FOCUS 5

# Where The System Works And Where It Fails

A system with asymmetric outcomes

| Where the system works                                 | Where the system fails                        |
|--------------------------------------------------------|-----------------------------------------------|
| Major labels and publishers with advanced data systems | Independent creators and performers           |
| Large licensing hubs with proprietary databases        | Smaller CMOs and local repertoires            |
| Platforms able to operate with aggregate reporting     | Session musicians and non-featured performers |
| Actors who control metadata infrastructures            | Actors who depend on downstream corrections   |

The system functions well enough to keep money flowing for dominant actors but fails structurally where accurate attribution and detailed data are required. Those who benefit most from opacity bear the least cost of fixing it.

# What Has Been Tried Before and What Europe Can Learn

Europe is not starting from scratch. Several initiatives have attempted to address music data fragmentation, offering valuable lessons for future reform.

## **THE GLOBAL REPERTOIRE DATABASE (GRD)**

The Global Repertoire Database aimed to create a single, authoritative database of musical works and rights ownership. Despite strong initial support, it failed due to:

- Technical difficulties in reconciling heterogeneous datasets
- Disagreements, mostly among CMOs, over governance and control
- Concerns about data ownership and competitive advantage

The GRD demonstrated that centralisation without clear, trusted governance is unlikely to succeed in Europe.

## **THE US MECHANICAL LICENSING COLLECTIVE (MLC)**

The United States adopted a different approach through the Mechanical Licensing Collective established under the 2018 Music Modernization Act and funded by streaming services.

Key features include:

- The creation and maintenance of a public musical works database for their 'mechanical rights' (traditionally acquired, owned, and licensed by music publishers, directly or through a specialised agency)
- Mandatory data submission obligations
- Regulatory oversight by a public authority

While the MLC improves transparency for rights in musical works, its scope remains limited to mechanical rights and reflects a legal and institutional context that does not exist in the EU. Moreover, the Music Modernization Act excludes social media companies and their UGC platforms from the

scope of the MLC's licensing and data consolidation mechanisms. This is because such UGC services continue to benefit from copyright liability exemptions for "hosting providers" laid down in the 1998 Digital Millennium Copyright Act.

## **LESSONS FOR EUROPE**

From these experiences, the Task Force drew several conclusions:

- Centralised databases are politically and institutionally difficult to sustain
- Voluntary cooperation must be supported by incentives and oversight
- Solutions must respect Europe's pluralistic rights management landscape

The challenge is to design a framework that enables coordination without imposing uniformity or transferring control to a single actor.

# A New Proposal for Europe: the European Music Metadata Infrastructure

In light of the structural problems identified by the Task Force, Fair MusE proposes the creation of a **European Music Metadata Infrastructure (EMMI)**. This proposal does not aim to replace existing databases or institutions. Instead, it introduces a coordinated framework designed to make them interoperable, accountable and fit for the realities of digital markets.

## WHAT EMMI IS

EMMI is conceived as:

- **A federated infrastructure** that links existing music data systems
- **A shared governance and standards framework** rather than a new database
- **A public-interest solution** that balances transparency with legitimate commercial confidentiality

At its core, EMMI is meant to enable reliable connections between musical works, sound recordings and performances, so that digital uses can be translated into remuneration more accurately and transparently.

EMMI draws upon the most relevant and promising effort made by industry actors in this field: the 2023 UK Industry Agreement on Music Streaming Metadata, a government-facilitated but voluntary compact that sets out a coordinated framework for improving metadata quality across the recorded music value chain. The UK Agreement is a voluntary but formal initiative that effectively defines a Core Data Set to be shared across the music and tech sectors.

On the grounds of Task Force discussions, D5.1 takes the view that the UK Agreement provides a credible model given the UK's proximity to the EU in terms of regulatory framework and the fact that the stakeholders actively involved in this initiative are the same as (or very close to) those who could reach an equivalent agreement for EMMI.

## **WHAT EMMI IS NOT**

To avoid repeating past failures, the Task Force was explicit about what EMMI should not be:

- It is not a single, centralised repertoire database
- It does not require rightholders to surrender ownership of proprietary data
- It is not a purely technical tool detached from governance and incentives

EMMI is designed as a coordination layer that respects Europe's pluralistic rights management environment.

## **WHY A EUROPEAN SOLUTION IS NEEDED**

The European music market is characterised by:

- Multiple legal traditions and collective management models
- Strong competition policy and resistance to monopolistic infrastructures
- High sensitivity around data ownership, trade secrets and confidentiality

A European-specific solution is therefore required. EMMI is designed to work within existing EU law, including copyright, data protection and trade secrecy rules, while addressing information asymmetries that undermine fair remuneration.

# Why EMMI is Different from Past Initiatives

## Coordination without control

Earlier initiatives failed because they attempted to centralise data or resolve ownership conflicts through a single authority. EMMI takes a different path.

It focuses on linking, not pooling. On verification, not ownership. On governance, not dominance. By operating as a federated, supervised infrastructure, EMMI aims to deliver coordination without concentration and transparency without expropriation.

## Regulation and EU data policy: why EMMI is not only an industry issue

D5.1 stresses that improving music metadata is not just a technical upgrade. It sits at the intersection of rights, market power, and access to information. This is why the deliverable links infrastructure reform to the broader EU pol-

icy environment: how transparency can be required, how access can be governed, and how responsibilities can be distributed more fairly across platforms, rightsholders, and intermediaries.

In practical terms, EMMI is easier to implement when rules and incentives support data sharing, interoperability, and verifiable reporting (especially where dominant platforms control key visibility points). The deliverable also connects this discussion to EU-level efforts to strengthen evidence and observability in the music ecosystem, so that fairness debates can rely on verifiable facts rather than un claims.

# What Data Should Be Mandatory: Core Data Set 2.0 and the “Golden Standard”

A central contribution of D5.1 is the identification of **what data must exist, and when**, for any music data infrastructure to function effectively.

## THE CORE DATA SET 2.0

The Task Force proposes a **Core Data Set 2.0** that should accompany every new recording from the moment of release. This dataset would include, at a minimum:

- Recording identifier (ISRC)
- Musical work identifier (ISWC), where available
- Names and identifiers of featured performers
- Relevant party identifiers (i.e., IPI, ISNI and IPN)

The objective is to ensure that key links between works, recordings and performances are established as early as possible.

## THE “GOLDEN STANDARD”

Beyond the minimum dataset, the Task Force introduces the concept of a **“Golden Standard”**. This refers to a high-quality, fully linked metadata set that includes:

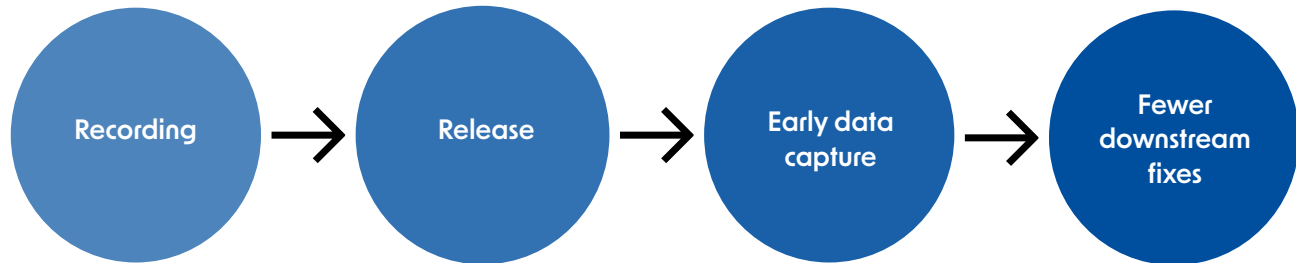
- Complete authorship and ownership splits
- Comprehensive performer information, including non-featured musicians
- Stable links between ISRCs, ISWCs and IPNs
- Versioning and update mechanisms

The Golden Standard is not expected to be achieved instantly. Instead, it serves as a **benchmark** against which data quality improvements can be measured over time.

## **WHY TIMING MATTERS**

A key insight from the Task Force and D5.1 is that music data quality problems are often created by when the information is provided or captured, not just by whether and how it is formatted.

If musical work and performer data are missing at the point of release, they are significantly harder and more expensive to reconstruct later. By setting expectations around early data capture, EMMI aims to shift the system from retrospective correction to preventive accuracy, the same way as in the UK Industry Agreement on Music Streaming Metadata.



# Core Data Set 2.0: What Changes Compared to Today

From partial metadata to structured accountability

| Today's common practice      | Core Data Set 2.0               | Golden Standard                    |
|------------------------------|---------------------------------|------------------------------------|
| ISRC assigned pre-release    | ISRC + ISWC expected at release | Fully linked ISRC-ISWC-IPN         |
| Performer data often missing | Featured performers identified  | Featured + non-featured performers |
| Author data added post-use   | Authors identified early        | Complete ownership splits          |
| Fixes done retrospectively   | Preventive data capture         | Continuous quality benchmarking    |

The Core Data Set 2.0 does not replace existing standards. It establishes a minimum expectation about when essential data should exist, reducing downstream correction costs and improving transparency.

# How EMMI Would Work: Architecture and Governance

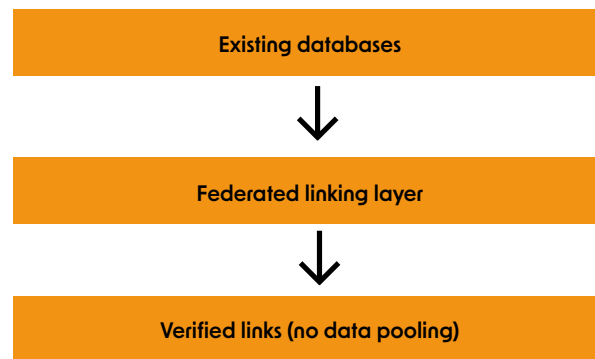
For the European Music Metadata Infrastructure to function in practice, technical interoperability must be combined with credible governance. The Task Force therefore proposes an architecture that is **federated, layer-aware** and **supervised**, rather than centralised.

## A FEDERATED LINKING ARCHITECTURE

EMMI would operate through a **federated linking layer** that connects existing databases without extracting or duplicating their contents. This layer would:

- Enable secure queries across databases
- Verify and match identifiers across rights layers
- Support the creation of stable ISRC-ISWC-IPN links

Data would remain with its original holders. What changes is the ability to connect information reliably across systems.



## LAYER-SENSITIVE DESIGN

The architecture explicitly recognises that:

- Works, recordings and performances are governed differently
- Data quality and availability vary across layers

- Responsibilities for data creation and maintenance are not uniform

EMMI therefore avoids imposing a single data model. Instead, it enables interoperability while preserving the legal and institutional specificity of each rights layer.

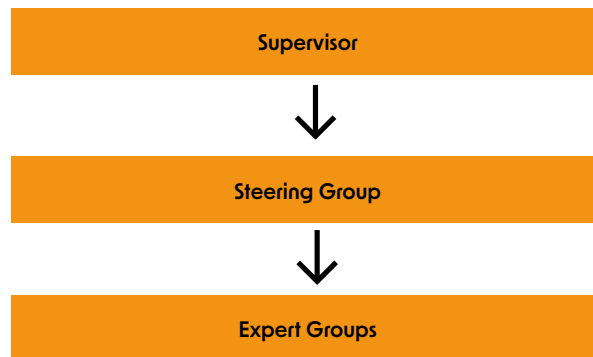
### **GOVERNANCE STRUCTURE**

Trust in EMMI depends on governance that is transparent, inclusive and credible.

D5.1 proposes:

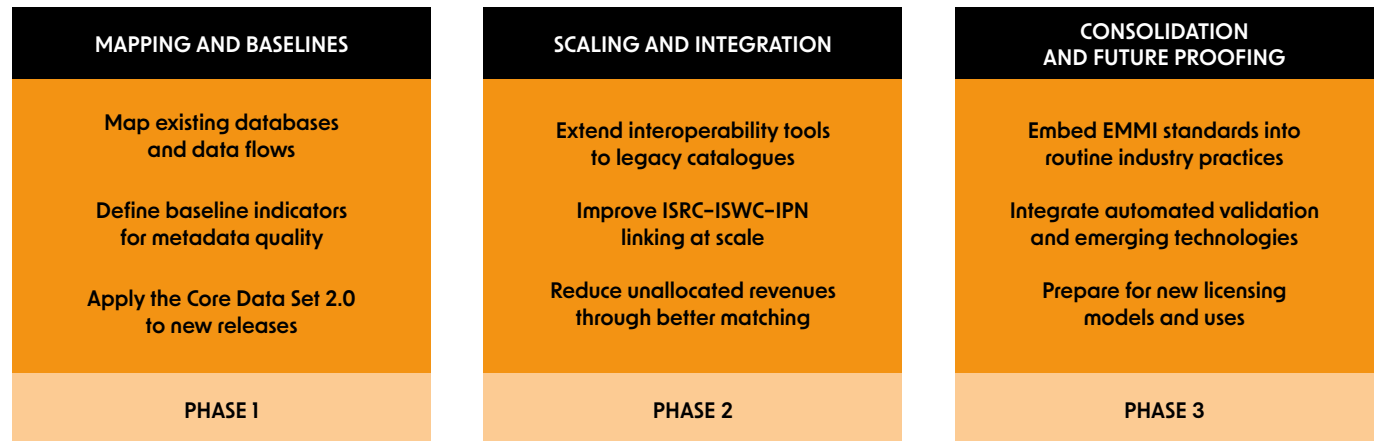
- **A light steering body**, with representation from creators, performers, CMOs, IMEs, music publishers, labels, platforms and public authorities
- **Expert working groups** on technical standards, education and data quality
- **A public supervisory role** for an EU body or agency, ensuring alignment with EU law and the pursuit of digital and cultural policy objectives

Governance would focus on coordination, monitoring and continuous improvement rather than operational control.



# Implementation Roadmap: a Phased Approach

To ensure feasibility, the Task Force proposes a gradual implementation:



# Incentives, Compliance and a Regulatory Safety Net

Voluntary cooperation alone is unlikely to be sufficient. The Task Force therefore emphasises the importance of incentives and accountability.

## Incentives for participation

EMMI participation can be encouraged through:

- Reduced administrative costs
- Improved auditability and legal certainty
- Access to higher-quality linked data

## SOFT-LAW INSTRUMENTS AND MONITORING

Rather than immediate legislative mandates, the Task Force favours:

- Codes of practice
- Reporting and benchmarking on data quality
- Public visibility of compliance efforts

These tools can create reputational and operational incentives without disrupting the market.

## A REGULATORY BACKSTOP

If voluntary uptake proves insufficient, existing EU legislation already provides leverage:

Transparency obligations under EU copyright directives  
Data access and accountability rules for platforms  
Competition and market regulation instruments

EMMI is designed to work with the existing legal framework, not to replace it.

# Incentives, Soft Law and a Regulatory Backstop

## Voluntary does not mean optional

D5.1 acknowledges that EMMI must begin as a voluntary, collaborative framework. However, voluntary participation is meaningful only if supported by incentives and accountability.

Incentives include lower administrative costs, improved auditability, and access to higher-quality linked data. Soft law instruments such as codes of practice, benchmarking and public reporting can reinforce these incentives through reputational effects.

Importantly, EMMI is not isolated from EU law. Existing transparency obligations under EU copyright directive provisions, platform accountability rules, and competition regulation already provide a regulatory safety net. If voluntary coordination fails, these instruments offer legal pathways to enforce greater data access and accountability.

# Benefits for the Music Community

A functioning European Music Metadata Infrastructure would generate concrete benefits.

## For creators and performers

More accurate and timely remuneration  
Clearer attribution of works and performances  
Greater ability to verify platform uses

## For CMOs, publishers and labels

Lower costs linked to data correction  
Improved licensing efficiency  
Stronger compliance with reporting and transparency obligations

## For platforms

Reduced legal uncertainty  
More predictable reporting and licensing processes  
Increased trust from rightholders and regulators

## For policymakers and the public

Greater transparency in digital music markets  
Better evidence for policy evaluation  
Stronger foundations for fair remuneration

# Looking Ahead: Metadata for AI and Synthetic Uses

Emerging technologies, including generative AI, will further increase the importance of reliable music metadata.

Training datasets, synthetic music generation and new forms of reuse raise complex questions about attribution, licensing, derivative works, and creators' remuneration. Without interoperable and trustworthy data infrastructures, these challenges risk amplifying existing asymmetries.

By strengthening metadata quality and governance now, EMMI can help prepare Europe's music ecosystem for future technological developments.

## CONCLUSIONS

# Building a Fairer Digital Music Ecosystem

The Fair MusE Task Force's work shows that the challenges of music data in Europe are not merely technical. They are structural, legal and economic.

Fragmented infrastructures, uneven access to metadata and weak incentives for accuracy have created a system in which fairness depends too often on market power rather than rights. A European Music Metadata Infrastructure offers a realistic and balanced path forward.

By connecting existing systems through shared standards, federated architecture and transparent governance, EMMI can reduce information asymmetries, improve accountability and support the principle of appropriate and proportionate remuneration.

Progress will require collaboration across the entire value chain, including meaningful engagement from powerful platforms. The reward is a digital music ecosystem that is not only more efficient, but also more equitable and sustainable for the creators and performers who underpin it.

# References

Mazziotti, Giuseppe, 'Music Copyright Infrastructure' (Fair MusE, D5.1, 2025).

Directive 2014/26/EU on collective management of copyright and multi-territorial licensing (CRM Directive). Referenced conceptually in D5.1's licensing and governance framing.

Directive (EU) 2019/790 on copyright and related rights in the Digital Single Market (DSM Directive). Used as background to EU fairness and transparency debates reflected in D5.1.

UK Industry Agreement on Music Streaming Metadata, Intellectual Property Office, UK Government (2023) available at <https://www.gov.uk/government/publications/uk-industry-agreement-on-music-streaming-metadata>

U.S. Digital Millennium Copyright Act (1998); more information at <https://www.copyright.gov/dmca/>

U.S. Music Modernization Act (2018); more information available at <https://www.copyright.gov/music-modernization/>

# List of acronyms

**CDM** — Common Data Model

**CMO** — Collective Management Organisation

**CRM Directive** — EU Directive 2014/26/EU on collective rights management and multi-territorial licensing

**DDEX** — Digital Data Exchange (standard-setting organisation for music data exchange)

**EMMI** — European Music Metadata Infrastructure

**GRD** — Global Repertoire Database

**IME** — Independent Management Entity

**IPN** — International Performer Number

**IPI** — Interested Party Information (identifier system for rightsholders/parties)

**ISRC** — International Standard Recording Code

**ISWC** — International Standard Musical Work Code

**MLC** — Mechanical Licensing Collective (US)

**MSS** — Music Streaming Service

**UGC** — User-Generated Content



Find us on



[fairmuse.eu](https://fairmuse.eu)



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.