Policy Brief

A SYMPHONY, NOT A SOLO

HOW COLLECTIVE MANAGEMENT ORGANISATIONS CAN EMBRACE INNOVATION AND DRIVE DATA SHARING IN THE MUSIC INDUSTRY

By *David Osimo, Laia Pujol Priego, Turo Pekari* and *Ano Sirppiniemi*

The views expressed in this policy brief are those of the authors alone and may not represent the views of the Lisbon Council, ESADE Business School, Teosto or any of their associates.

The paper was commissioned by Teosto and co-funded by the Lisbon Council and ESADE Business School. It is the outcome of a year-long collaboration intended to stimulate discussion among key players in the music industry with the aim of accelerating data-driven innovation in music.

ABOUT THE AUTHORS

David Osimo is director of research at the Lisbon Council.

Laia Pujol Priego is researcher at ESADE Business School.

Turo Pekari is senior advisor on innovation and discovery at Teosto.

Ano Sirppiniemi is head of research and development at Teosto.

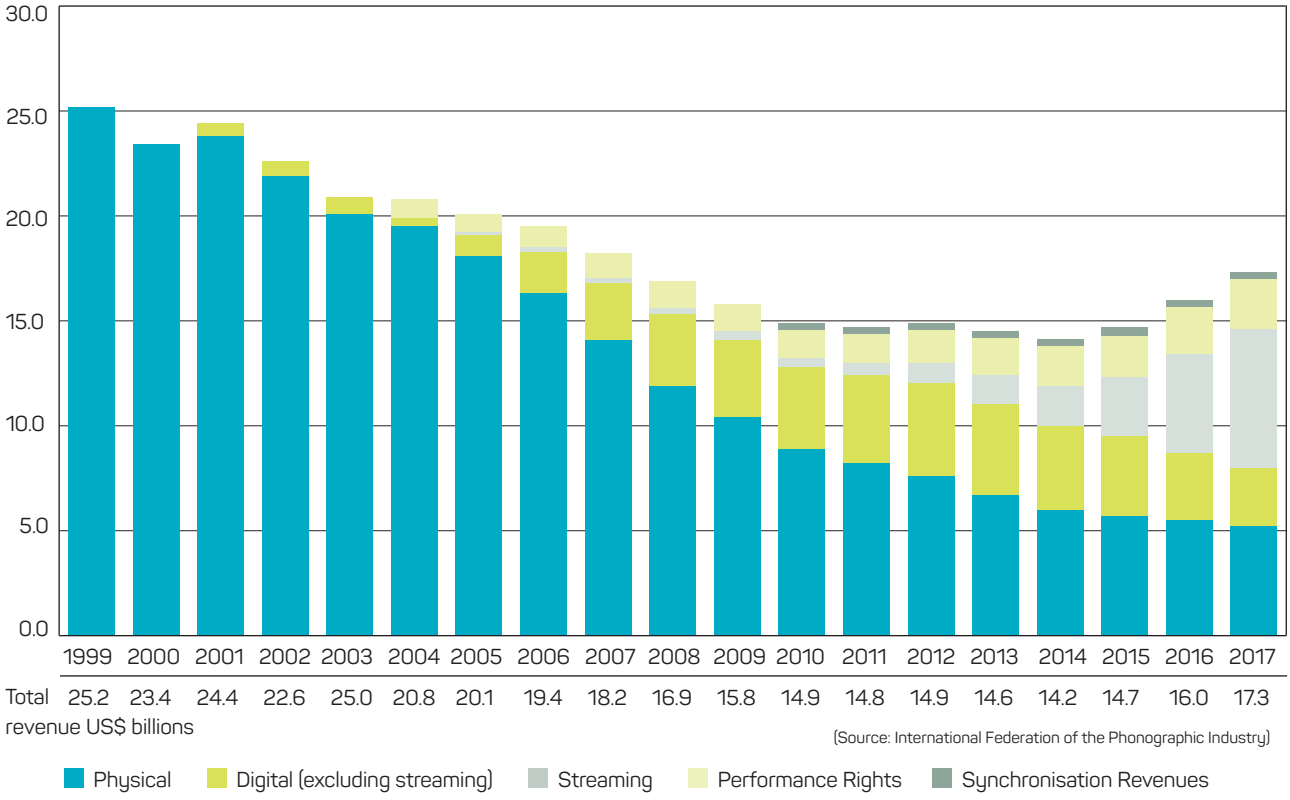
INTRODUCTION

Calum Scott was one of the sensations of *Britain's Got Talent*, the television show where amateur performers compete for a cash prize – along with fame and glory.[1] The episode – where Mr. Scott covered Swedish pop icon Robyn's "Dancing On My Own" – had 200 million views on YouTube, and Mr. Scott reached the finals of the show. But this fame proved short-lived, and ultimately after the end of the show he didn't manage to sign with any label. One year later, Mr. Scott decided to publish his performance as a solo song on YouTube. It didn't become an immediate viral phenomenon with millions of reproductions, but it showed a subtle and rapid growth of subscribers.

Alerted by this growth, Instrumental, a London-based startup that uses artificial intelligence (AI) to analyze platform data and identify talent, reached out to him to publicly release the song across digital platforms. Using the data from these platforms, Instrumental fine-tuned the marketing campaign and Mr. Scott's performance eventually reached No.14 in the UK charts. Later, Capitol Records signed him and the song became the top selling single in 2016 by a British artist. And Instrumental released their AI platform as a service to any label to identify new talent. Thanks to platforms, data availability and AI, a small gem was uncovered, and a market created.

The music industry has been a pioneer of digitization – from CD ROM to Mp3 – and has probably experienced earlier than any other sector the impact of the Internet, both positively and negatively. The good news is that after years of decline, music revenues have grown since 2015, mainly thanks to streaming solutions such as Spotify and Apple Music (see Chart 1 below). Nordic countries are at the forefront of the market transformation. In Finland, streaming music services account for more than 80% of total music sales and 90% of the population use streaming services.[2]

CHART 1: GLOBAL RECORDED MUSIC INDUSTRY REVENUE (1999-2017)



(Source: International Federation of the Phonographic Industry)

1) The authors would like to thank Kevin Bacon, Jake Beaumont-Nesbitt, Paul Hofheinz, George Howard, Katarzyna Jakimowicz, Alex Loscos, Shain Shapiro, Tjerk Timan, Ken Umezaki and Stefaan Verhulst. As always, any errors of fact, judgment or omission are the authors' sole responsibility.

2) Simon Bugge Jensen and Marie Christiansen Kroyer, *Polaris Nordic: Digital Music in the Nordics* (YouGov, October 2018) <https://www.koda.dk/media/184140/polaris-nordic_digital-music-in-the-nordics-2018_yougov.pdf>.

Streaming proved to be a game changer not only in terms of new sources of revenues, but also by making available data of unprecedented granularity about music consumption – a tremendous opportunity for creators and producers to better understand how their products are used. Spotify reports to have 200 petabytes of data, and every day its users generate 150 billion data points such as "playing songs, sharing, selecting recommended music, skipping, following, and active participation through the upvote and downvote buttons."[3]

This staggering amount of data – combined with additional data such as from social media - is changing the music industry across the whole value chain, from discovering talent to marketing to managing rights.

Platforms such as Spotify and Pandora Music have long provided recommendation engines for customers. Now they have started delivering analytics services for artists, producers and publishers, too, and have increased consolidation in the industry by purchasing music data companies (See Table 1 for more).

TABLE 1: OVERVIEW OF RECENT ACQUISITIONS OF MUSIC DATA COMPANIES

ACQUIRED COMPANY	DATA SERVICE	ACQUIRER	YEAR
Echo Nest	Consumption analytics	Spotify	2014
Next Big Thing	Consumption analytics	Pandora	2015
Semetric	Consumption analytics	Apple	2015
Audiam	Rights management	Socan	2016
Gracenote	Consumption analytics	Nielsen	2017
Mediachain	Rights management	Spotify	2017
Loudr	Rights management	Spotify	2018
Shazam	Consumption analytics	Apple	2018
Sodatone	Consumption analytics	Warner Music	2018

As shown in Table 1 above, the two main applications of big data in music are related to consumption analytics and rights management.

What's more, as is, record companies spend a massive 17% of their revenues on discovering and signing new talents – a process known as "artists and repertoire," or A&R. As a percentage of the overall business, it is more than what pharmaceutical or software companies spend on research and development (13 and 11%, respectively).[4] These days, the music industry is complementing the traditional "gut instinct" with a data driven approach to help this scouting activity. Startups such as Instrumental dig into platforms' consumption data (Spotify, YouTube, Facebook, Instagram and others) to identify emerging talent and trends, in order to help the work of A&R through artificial intelligence.[5] And Spotify is trying to position itself as the "R&D department of the entire music industry."[6]

Secondly, a new wave of new music startups emerges that challenges the established ways of distributing music and dealing with rights management in the industry. For instance, UnitedMasters, a startup that aims to give musicians an alternative to traditional record labels, received $70 million [€62 million] in funding from Alphabet, the parent company of Google. Artists remain the owners of master records but they pay the startup a fee for distributing their music on the digital music platforms as well as they split the royalties with the UnitedMasters. Music rights collection startup, Kobalt Music Group, which was the first investment of GV (formerly Google Ventures), have gathered $200 million [€175 million] in funding, with an estimated valuation of $789 million [€698 million]. Kobalt is basically aiming to substitute collective societies in dealing with Digital Service Providers (DSPs), providing big-data tools and dashboard to collect royalties from all the internet plays. Other companies provide blockchain-based services, such as Ujo Music Group to manage smart contracts and make payments in real-time.[7]

In short, data is becoming a fundamental source of competitive advantage in music, just as in other sectors, and streaming services in particular are generating large volume of new data offering unique insight around customer taste and behavior. As *Financial Times* recently put it, the music industry is having its "moneyball" moment.[8]

But how are the different players getting ready for this change? This policy brief aims to look at the question from the perspective of CMOs, the organisations charged with redistributing royalties from music users to music rightsholders (such as musical authors and publishers).

The paper is divided in three sections. Part I will look at the current positioning of CMOs in this new data-intensive ecosystem. Part II will discuss how greater data sharing and reuse can maximize innovation, comparing the music industries with other industries. Part III will make policy and business-model reform recommendations for CMOs to stimulate data driven innovation, internally and in the industry as a whole.

3) As reported in the SEC filing of Spotify of 28 February 2018. To put this in perspective, Netflix generated 60 petabytes as of November 2016.

4) Héctor Hernández and others, *EU R&D Scoreboard the 2017 EU Industrial R&D Investment Scoreboard.* (Seville: JRC IPTS, 2017); IFPI, *Investing in Music* (London: IFPI, 2016).

5) For more, visit https://www.weareinstrumental.com/.

6) As announced by Gustav Söderström, Spotify's chief research and development office, at a press conference in New York on 24 April 2018. See Cherie Hu, 'Spotify Wants to Be "The R&D Department for the Entire Music Industry" -- What Does That Actually Mean?', *Billboard,* 2018 <https://www.billboard.com/articles/business/8376561/spotify-rd-department-entire-music-industry-netflix> [accessed 7 November 2018].

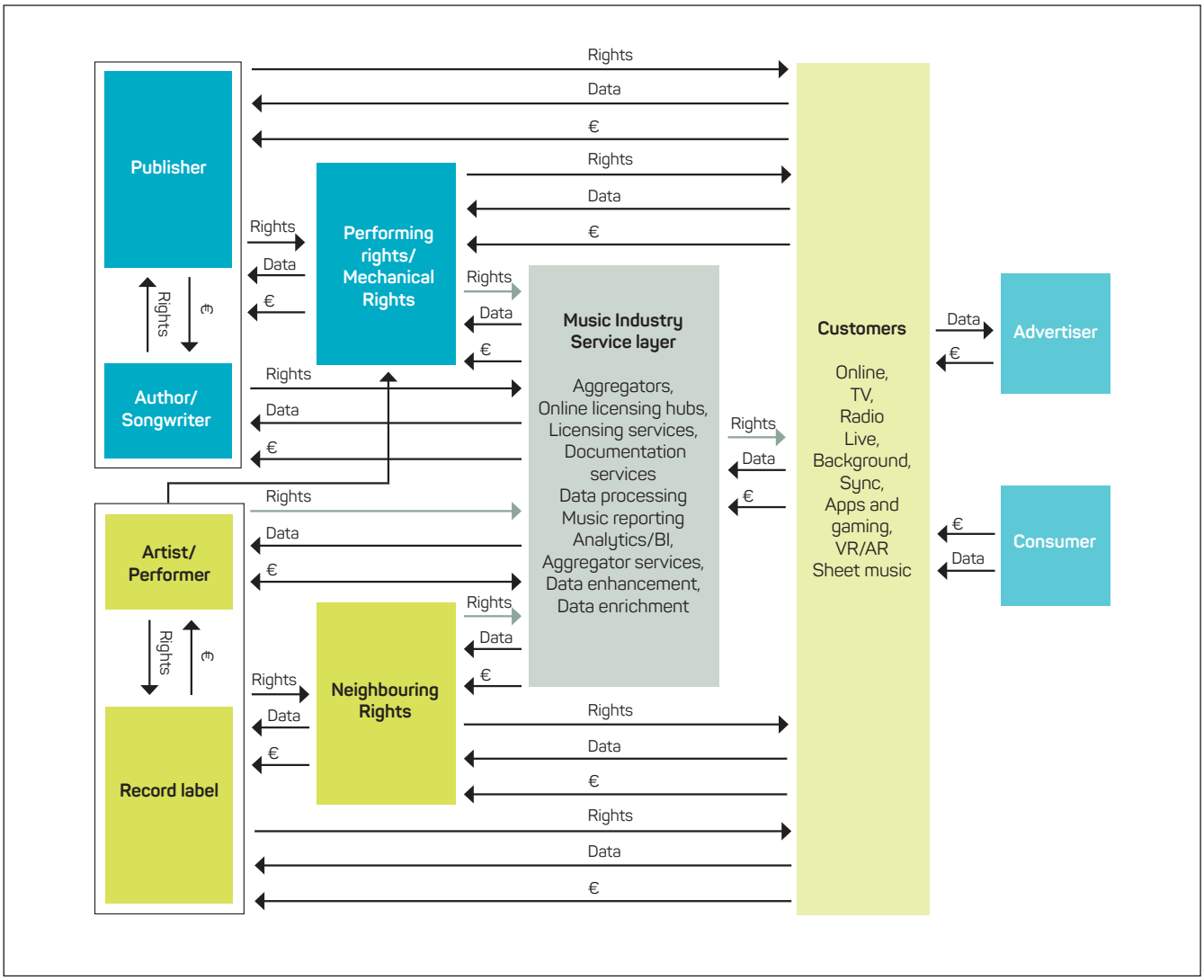
7) For more, visit https://ujomusic.com/.

8) Michael Hann, 'Music's "Moneyball" Moment: Why Data Is the New Talent Scout', *Financial Times*, 5 July 2018 <https://www.ft.com/content/474ae18a-7f1b-11e8-bc55-50daf11b720d> [accessed 09 November 2018].

1. THE OPPORTUNITY FOR COLLECTIVE MANAGEMENT ORGANIZATIONS

Collective management organizations, e.g. collecting societies, are charged with licensing copyrighted works and collecting royalties negotiated on behalf of their members. Collecting societies collect royalty payments from users of copyrighted works and distribute royalties to copyright owners. They have been established for managing and licensing of performing and mechanical rights of authors and publishers as well as performers and labels. CMOs are often considered a legacy of the past, and in many countries they are a particularly conservative form of organization which benefited from a monopolistic position. But at the same time, some CMOs are embracing change and launching innovative ventures, building on their unique competitive advantage such as the combination of data and specialized knowledge.

MUSIC INDUSTRY DATA AND VALUE CHAIN



Managing collective rights generates a complex data flow about rights and payments between different players. Basically, there are three layers. The **authoring /publishing /performing** layer (left part) includes record labels and music publishers as well as authors (songwriters, composers, lyricists, arrangers of music in any form) and artists (musicians and singers who create recorded music and/or perform music for live audiences). The **music industry service layer** (middle part) includes intermediaries such as CMOs, aggregators, licensing hubs and metadata platforms. The **customer layer** (right side) includes music users (except final consumers) such as music streaming services (e.g. Spotify, Apple Music) and video on-demand services (VOD) (e.g. Netflix, HBO, Amazon.); broadcasters such as TV, radio and telecommunication companies; Live music event organizers; companies using background music for films, ads and games.

1.1 THE OPENING OF THE MARKET FOR COLLECTIVE RIGHTS MANAGEMENT

The market of collective rights management has recently become highly competitive. National CMOs can now expand their activities in other EU countries. Technology platforms such as Spotify and Pandora increasingly manage licences directly. And a new wave of startups for music rights management has emerged, such as Kobalt and Auddly.

This increased competition is the result of the opening up of the market through a European Union regulatory intervention. In February 2014, the EU adopted *Directive on Collective Management of Copyright and Related Rights and Multi-Territorial Licensing of Rights in Musical Works for Online Uses in the Internal Market* (CRM Directive).[9] The aim is to ensure that rights holders have a say in the management of their rights, and at improving the functioning and accountability of CMOs and facilitate the multi-territorial licensing by collective management organizations of authors' rights in musical works for online use. The directive has built ground for multi-territorial licensing, which is an important part of the European Digital Single Market strategy, created transparency and financial standards for controlling the way European CMOs work. At the same time, the directive has enabled also new market participants, also for-profit players, entering into the domain, which has already happened in Italy and Spain.[10]

At the same time, the growing importance of direct licensing is shrinking the economic importance of collective management intermediaries. When music consumption went digital in the early 2000s, at first author rights were nationally licensed mostly through CMOs, like in many other areas of music licensing. However, in an online world both publishers and CMOs have started to license their repertoire directly to digital platforms for pan-European or global music licensing deals, withdrawing their repertoire from collection societies.[11] Some large publishers (Sony/ATV Music Publishing, Universal Music Publishing, Warner/Chappell, Bertelsmann Music Group) have also made deals with big CMOs and formed joint ventures for direct licensing of music, bypassing national players.[12] What this all means, is that CMOs are suffering from shrinking online repertoires and a loss of negotiating power, when they are able to license only a part of the repertoire used in their territory. Especially for smaller CMOs it means that in digital music, "you are worth what your local repertoire is worth."

The industry has also experienced fast technological innovation. Four key drivers deriving primarily from technological change have an effect on CMOs:

1) Decentralization of rights management technologies. Although online music distribution and consumption have been decentralized – from a technology perspective – for almost a decade, the current wave of digitalization is promising to decentralize core aspects of the music rights management business. Distributed ledger technologies, or blockchains, are based on the idea of decentralizing data collection and quality control, identity management, system governance and trust.

2) Business and process automation based on machine learning and AI. Financial industries (banking, insurance and fintech companies) are already routinely using machine learning for tasks such as credit scoring and fraud detection, as are also music streaming services for automated content recognition and music recommendations. Companies are more an more relying on learning algorithms for handling complicated business processes, and the music rights management will not be an exception. And the more and better the data, the faster machines learn.

3) Requirements for interoperability across the value network. Emerging common protocols for music industry interoperability, such as the minimum viable data specifications of the Open Music Initiative project, and common protocols and application programming interface (API) specifications are, together with existing music industry standards, identifiers and protocols (such as ISWC, ISRC, CWR, CRD, DDEX and others), important for enabling linkages across "data islands."

4) Real-time processing of data. Current processing times for data at CMOs, from music use to payout to rightsholders, are typically very long. This type of delay in payouts could in the future prove to be a risk for traditional CMOs facing new types of competition, especially if the delay cannot be justified by it resulting in more accurate payouts or lower total costs of processing. Speed can be a differentiating factor for new competitors entering the rights management business.

9) Directive 2014/26/EU

10) In Italy, from 2018 the incumbent CMO SIAE is now in competition with a non-profit entity related to Soundreef, a UK startup.

11) Tim Ingham, 'Kobalt's AMRA Signs Its First Global Collection Deal… with Apple', *Music Business Worldwide*, 2015 <https://www.musicbusinessworldwide.com/kobalts-amra-signs-first-global-collection-deal-apple/> [accessed 14 December 2018].

12) Andre Paine, 'Solar Activity: Sony/ATV, GEMA, PRS and Solar Extend Pan-European Licensing Deal', 2018 <http://www.musicweek.com/publishing/read/solar-activity-sony-atv-gema-prs-and-solar-extend-pan-european-licensing-deal/074278> [accessed 14 December 2018]; Chris Cooke, 'ICE and SACEM Sign up More Indie Publishers for Digital Licensing | Complete Music Update', 2018 <http://www.completemusicupdate.com/article/ice-and-sacem-sign-up-more-indie-publishers-for-digital-licensing/> [accessed 14 December 2018].

13) Paul Hofheinz and David Osimo, *'Making Europe A Data Economy: A New Framework for Free Movement of Data in the Digital Age'*, Lisbon Council Policy Brief, 2017.

14) http://open-music.org/our-api/

Blockchain in music rights management

As an in-house concept, Teosto, a Finnish CMO, has designed a concept for the use of blockchain technologies in improving data exchange practices between CMOs. As a result, a Teosto-owned spin-off company, Mind Your Rights, was established in 2017. It focuses on creating a platform for managing live music data and payments across CMOs. Mind Your Rights has received funding from The Finnish Innovation Agency Business Finland to carry out a market study and plan for an alpha version release of the platform.

Dotblockchain Media is a US-based company offering music industry solutions metadata management, minimum viable dataset data sharing, synchronizing songwriting and performing rights, metadata enhancement with machine learning and building a collective truth of ownership. Dotblockchain uses Intel hyperledger sawtooth as its core blockchain technology and has published industry partnerships with Warner Music Group, CD Baby, FUGA, Socan, Medianet and Unison Rights.[15]

JAAK is a UK-based tech-company that builds blockchain solutions for music industry. The company announced the KORD data sharing initiative in fall 2017 and published first results of data sharing pilot with partners including Warner Music Group, Warner/Chappell Music, Global Music Rights and Bertelsmann Music Group.[16] Jaak has been open in public about both difficulties of data sharing in the music industry's competitive and siloed data landscape but also about the possibilities of value-added innovation that the whole industry could gain from.[17]

Blokur is a UK-based company developing music industry solutions built on blockchain technology and was founded in 2016 by Phil Barry (previously founder of Ujo Music and involved in Imogen Heap's Tiny Human blockchain project in 2015). The company is working on a solution for bridging the siloes between different sources of rights data. The Blokur platform uses algorithms that automatically reconcile different claims on rights to a single state, captured on the blockchain. Although not publicly named, Blokur representatives have spoken in public about their collaboration with both big publishers and CMOs.

15) For more information, visit http://dotblockchainmedia.com/main/#about-section and http://dotblockchainmedia.com/warner/.

16) Stuart Dredge, 'Jaak Hails Its First Blockchain Music-Rights Pilot with Rightsholders', 2018 <https://musically.com/2018/05/02/jaak-hails-successful-blockchain-music-rights-pilot-with-rightsholders/> [accessed 14 December 2018].

17) Dan Fowler, "The KORD Network: Lowering the Barriers for Developers in the Music Industry," *JAAK blog*, 05 October 2017. https://blog.jaak.io/the-meta-network-lowering-the-barriers-for-developers-in-the-music-industry-b0a0431c3528

1.2 THE VALUE OF DATA HELD BY CMOS

Faced with increased competition, CMOs have to look for ways to remain competitive on the market. And one of their assets is certainly data. But across all industries, it has proved difficult, not to say impossible, to assign a value to data. The OECD suggests the notion of data as an "experience good": data are considered "a good that consumers must experience in order to value" and the value of which is not intrinsic but depends on the context of its use. The implication of this definition is that market mechanisms are not effective as the value is context dependent, and markets do not always converge towards a price accepted by supply and demand.[18] The value lies therefore not in the data themselves, but in the analytical insight they enable. As O'Reilly's Strata Conference's Alistair Croll has pointed out: "The digital divide isn't about who owns data – it's about who can put that data to work."[19]

While the value is not intrinsic to the data, it is evident that datasets are not equal in value. There are "superstar" datasets, the low hanging fruit that are easy to exploit, and then the rest. In most industries, the high value datasets are about individual consumption habits: granular data about "who consumes what": consumer spending pattern in retail, car driving patterns, banking transactions, individual patient or genome data, company and house registers in government. And the value does not lie in individual data points, but in the aggregation of massive, granular datasets that cover almost 100% of the relevant market in great detail. Or that can be extrapolated to do that as in the case of mobile data traffic, where operators can adjust the traffic data from their own clients to produce representative analytics.

Hence, in the context of music, consumption data held by intermediaries such as Spotify, YouTube and Pandora Media Inc., providing real time granular data about millions of user choices, is the superstar dataset of the music industry. These are the data that could enable Spotify to become the R&D department of the music industry.

This is not to say that other datasets are not valuable – by definition, the data economy is built on reusing data for a previously unplanned purpose. But they are much more difficult to extract value from. Consumption datasets are the low hanging fruit of the data economy, because they provide new data previously unavailable: the granular patterns of individual song consumption.

This is not the case for data held by CMOs. CMOs do not own granular data on consumption – they actually use the data provided by intermediaries. CMOs, because of their positioning in the value chain, own different types of data:

1) Basic data to identify works. These contain the basic information to identify the song, author, publisher, year of publication and title.

2) Data about the ownership of the works. These are complex datasets about how revenues should be distributed among the many different rightsholders. This is probably the most important and unique dataset owned and has been the object of several failed data sharing initiatives.

3) Data about the works performed at live events, including the works performed at the event that are necessary to calculate revenue distribution. This is particularly important in view of the growing share of industry revenue due to live events. According to PwC, the value of the international concert business is projected to increase at a compound annual growth rate of 3.3% over the next five years.[20]

4) Aggregate data about music consumption in online services, broadcast television and radio, and background music uses, that are used to calculate royalties.

In other words, CMOs do not own "superstar datasets" and their existing data coverage is shrinking. One of the common mistakes in the data economy is that organizations and people – and CMOs are no exception – tend to overestimate the value of their data. Holding on to their data will not provide the sustainable competitive advantage they need to survive. And monetizing those data is difficult because of the lack of granularity. The solution is to move towards greater data openness and sharing. It's not just about opening up data held by CMOs, but to favor greater data sharing across the industry value chain.

18) OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being* (Paris: OECD, 2015), CCXV <https://doi.org/10.1787/9789264229358-en>.

19) Alistair Croll, "Who Owns Your Data?," 12 January 2011. http://news.yahoo.com/owns-data- 20110112-030058-029.html

20) Ennel Van Eeden and Wilson Chow, *Global Entertainment & Media Outlook 2018–2022* (London: PricewaterhouseCoopers, 2018) <https://www.pwc.com/gx/en/industries/tmt/media/outlook.html> [accessed 05 November 2018].

2. TOWARDS GREATER DATA SHARING IN THE MUSIC INDUSTRY

2.1 THE CURRENT STATUS

One of the principles of the data economy is that to spur innovation, data should be reusable for a different purpose than the one it was originally gathered for. As Rufus Pollock, founder of the Open Knowledge Foundation, puts it: "the best thing to do with your data will be thought of by someone else." This has become even more important in recent years, as it became clear that machine learning and artificial intelligence require large scale data access and reuse. Yet for all the opportunities, the reality is that companies remain very protective about their data: the vast majority of big data business projects do not entail any form of data sharing.[21]

This is true in the music industry too. Just as in other sectors, accessing and reusing third party data spurs competition and innovation in the industry, and it ensures greater trust and transparency. But the music industry is very competition driven and data are typically in closed silos: as one interviewed expert put it, "The music industry is predominantly bad at sharing data and not collaborative in any way. ... the hardest difficulty that we have is accessing anonymized, aggregated consumption data from music."

To be clear, there have been several initiatives to promote data collaboration and sharing in the music industry, namely on the music repertoire and ownership.

- The Open Music Initiative (OMI), a non-profit initiative of Berklee College of Music in partnership with Context Labs, IDEO, and the MIT Media Lab, composed of over 150 members, has developed standards for interoperability for participants in the music ecosystem to share information about recordings, works, creators/contributors, and their relationships, embodied in the OMI API.

- In Europe, the Global Repertoire Database (GRD) Initiative (2008-2014) was launched in 2008 by the European Commission to create a global open database providing access to "authoritative, comprehensive, multi-territory information about the ownership and control of the global repertoire of musical works" to increase the transparency and effectiveness in royalty collection and distribution and lower the associated administrative costs. The project was discontinued in 2014 due to loss of funding and lack of cooperation of CMOs.[22]

On the market side, digital platforms today provide access to their data by third parties. All the main platforms today provide artists and publishers with access to their analytics. In addition, platforms such as Spotify provides access to the overall analytics to third parties through its public API, and Pandora Media has recently opened its artists' data to the public. Digital platforms have undoubtedly had a net positive effect on data availability and innovation: thanks to the availability of these data, startups such as Instrumental offer innovative solutions to uncover new talent.[23] At the same time, the data made available represent a tiny portion, at a very aggregate level, of the millions of data points created every day.

In other words, data sharing in the music industry happens either through the voluntary initiative of the platforms, or through collaborative initiatives still limited in scope and in progress, focusing mainly on the music repertoire and its ownership. This piecemeal approach is sufficient to grasp the low hanging fruits of innovation, but not the new, disruptive products and services that we cannot yet imagine. For instance, real-time aggregate streaming statistics are certainly changing the way industry does A&R, but ultimately, it's an incremental innovation with respect to the charts of the 80s: more data, available earlier. Yet music is a beacon of innovation and experimentation within and beyond the music industry, and beyond the traditional datasets of consumption and rights. Just look at how Sound Diplomacy is helping local government to use music to shape the city of tomorrow or to promote tourism.[24] Or at how the University of Helsinki enhances the live music experience through emotion aware biosensors and affective computing. Or at how Amper AI music composer delivers in seconds original, machine-generated background music for video content. Music is increasingly digitized, and the future innovation of the music industry will largely rely on accessing and reusing the widest range of data and purposes. Data silos will slow down innovation, make it costlier and ultimately stifle it.[25]

21) Laia Pujol Priego and others, *'Data-Driven Business Models in the Digital Age: The Impact of Data in Traditional Businesses'* (presented at the World Open Innovation Conference, Barcelona, 2016).

22) Klementina Milosic, 'The Failure Of The Global Repertoire Database', Hypebot, 2015 <https://www.hypebot.com/hypebot/2015/08/the-failure-of-the-global-repertoire-database-effort-draft.html> [accessed 5 November 2018].

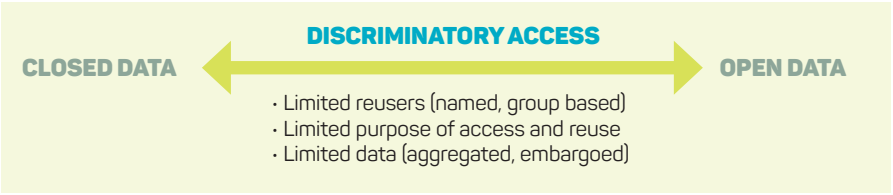
23) Daniel Sanchez and 2018, 'Pandora Opens All of Its Artist Data to the Public Through Next Big Sound', *Digital Music News*, 1 October 2018 <https://www.digitalmusicnews.com/2018/10/01/latest-pandora-next-big-sound-metrics/> [accessed 7 November 2018]; Stuart Summer, 'Instrumental: Using AI and Spotify Data to Find the next Big Thing in Music | Computing', Http://Www.Computing.Co.Uk, 24 July 2018 <https://www.computing.co.uk/ctg/feature/3036430/instrumental-using-ai-and-spotify-data-to-find-the-next-big-thing-in-music/page/3> [accessed 7 November 2018].

24) https://www.sounddiplomacy.com/

25) Thomas H. Davenport, 'Analytics 3.0', Harvard Business Review, 91.12 (2013) <https://hbr.org/2013/12/analytics-30> [accessed 14 December 2018].

2.2 THE OPPORTUNITIES FOR DATA SHARING

CHART 2:THE DEGREES OF DATA SHARING (adapted from OECD and Open Data Institute)



The answer to the challenge is not a "radical openness" agenda. The choice on data sharing is not simply between "open" and "closed" data. There is a wide variety of options in between, based on limiting those who can access, the data availability and the purpose. Each option can be best in a specific context – in particular, depending on the value and sensitivity of the data.

The most widely recognized way to share data is simply to publish them as open data. And this is an important opportunity for innovation, even more for CMOs facing regulatory scrutiny over transparency. But today, opening up data has reached a mature stage beyond the initial "PR exercises" such as publishing a data portal, or dumping ".csv" files from time to time in a repository. It is now a strategic operation which includes:

1. The identification, in dialogue with potential re-users, of the most valuable datasets to be released.
2. The curation of data and metadata to ensure their reusability.
3. The publication of data through real-time interactive services such as APIs.
4. The stimulation of reuse through hackathon and prizes.

Because opening up data in a useful way can be a substantial cost, it has to pursue clear strategic goals beyond the "feel good" effect. For instance, it has to aim at creating a community of developers, at purchasing innovative services built on open data, and at enriching the data.

An organic approach to open data: Teosto Open API Initiative

Teosto, the Finnish CMO, has released its live music data as open data via an API (http://api.teosto.fi). The data includes information from over 250 000 gigs in Finland from 2014 to 2017, including geolocation data, venue information, dates, performers, complete setlists and author and publisher information for the performed songs.

Teosto Open API data has been actively promoted in 10 hackathons, most notably in Teosto's own Teosto Hack Day in 2015, and Music Tech Fest in Stockholm in September 2018, with the API used and tested by approximately 145 developers in these events. Hackathons have produced 10 prototypes or proof-of-concepts, where Teosto's data has been enriched with e.g. artist biographies and genre information, and the data has been visualized in different ways. An example of these is http://keikkahistoria.fi , a demo where Teosto live music data is put on a map, and is searchable by artist name.

Moreover, Polaris Works API, containing a static dataset of Teosto's work repertoire with 260 million lines of data, was developed in partnership with Revelator in 2018 as a part of Polaris Future Lab (collaboration of three Nordic CMOs Teosto, Koda and Tono), with the goal to offer Polaris Works API as a data sandbox for R&D projects for customers. The objective is to build new efficiency in operational processes and to improve data sharing practices with external partners.

Open data is a flagship initiative for public sector organization, but it is increasingly taken up by private companies, too. Banco Bilbao Vizcaya Argentaria (BBVA) shares openly aggregated data about bank transactions. Vodafone Group plc has set up open data initiatives for researchers to analyze customers' mobility data for purposes of addressing societal challenges, such as malaria. And the EU policy debate has recently focused precisely on extending the open data provision from government data towards data held by private business that could be useful for public purposes.[26]

But as we have seen, open data is only one of the options, and there are other opportunities for sharing through discriminatory type of access. In particular, as shown in Table 2 below, over the last years different solutions have emerged across many industries which can be categorized in three groups: open data initiatives, discriminatory data sharing initiatives with competitors, and discriminatory initiatives with other players in the value chain, such as suppliers and customers.

TABLE 2: ILLUSTRATIVE EXAMPLES OF DATA SHARING INITIATIVES ACROSS SECTORS

SHARING MODEL SECTOR	Open data	Discriminatory data sharing with competitors	Discriminatory data sharing with other stakeholders
Pharmaceutical		OpenTargets[27]	Yoda[28]
Water utilities		California Data Collaborative[29]	
Energy	ODRE[30]		
Agri-food			A.T. Kearney fast food supply chain data sharing
Aerospace			Skywise[31]
Telecom	Data Pop Alliance and Open Algorithm project[32]		Telefonica Insights, Vodafone Analytics
Finance	BBVA open data portal[33]	Data sharing with third party providers based on PSD2 directive	

Surprisingly, these initiatives (falling under the different names of "data commons," "data trusts," "data collaboratives," "cleanrooms" and "industrial data platform") are emerging particularly in intellectual-property-intensive industries such as pharmaceutical. Faced with increasing difficulties in developing new successful drugs, increasing failure rates, and the formidable opportunities provided by merging different datasets, companies are sharing data with competitors to accelerate the rate of discovery.

For example, the Yoda project allows pharmaceutical companies to share clinical trial data in a protected environment with researchers. The data requests and sharing process is managed by an independent third party at Yale University. The OpenTargets project is a public private partnership involving some of the biggest pharmaceutical companies (Sanofi, GSK) sharing data among themselves and with the European Bioinformatics Institute in order to identify good drug targets (the proof that a molecule is related to a disease). Initiatives such as Yoda show that companies are willing to share clinical trial data, previously considered "unshareable" in order to accelerate drug discovery, reduce costs or drive innovation. But the answer is not to simply publish data openly: rather, it is to set up carefully designed structures that enable data sharing while ensuring clear rules to avoid free riding effects. In many cases, such as Yoda and OpenTargets, companies recur to trusted intermediaries, typically research centers, to ensure proper data management.

26) See in particular the proposal (COM/2018/234) for a revision of the Directive on Public Sector Information (2003/98/EC).

27) https://www.opentargets.org/

28) http://yoda.yale.edu/

29) http://californiadatacollaborative.org/

30) https://opendata.reseaux-energies.fr/pages/accueil/

31) https://www.airbus.com/aircraft/support-services/skywise.html

32) https://www.opalproject.org/

33) https://bbvaopen4u.com/es

Companies are also developing new solutions merging their data value chain and their supply chain.[34] For instance, A.T. Kearney recently built a cleanroom for a global fast-food chain and its key capital equipment vendors. The fast food chain shared forecasted future demand and outlet expansion data and capital equipment vendors provided data about topics such as their global manufacturing footprint and total landed cost for products. As a result of data integration, both sides were provided with a more accurate view of future optimized supply chain and potential gains, while at the same time all sensitive information was protected. Another example is a Skywise, an open digital platform for aviation, launched by Airbus, where companies make data available in a close, secure environment in return for data analytics services.

From the perspective of the music industry, there seems to be an "experimentation gap" with respect to other sectors to make better use of the full range of options for greater data sharing and reuse. If intellectual property intensive sectors where competition is fierce and R&D expenditure is high, such as pharmaceuticals, are sharing data in precompetitive platforms, why shouldn't this be possible in the music industry?

a. How high value data can be shared

These data sharing initiatives are not the results of a sudden idealism of the involved companies, deciding to put collective good in front of private interest. It is instead the result of a mature culture of data management that is able to carefully design system to ensure cooperation without compromising each participant's competitive advantage.

In particular, these solutions have different mechanism in place to ensure the benefits of data sharing without the drawbacks:

1) **Level of Aggregation**.
Superstar datasets are highly granular, but can be shared openly at different levels of aggregation, as shown by the BBVA open data portal or by the streaming statistics provided by digital platforms.

2) **Timeliness.**
Superstar datasets are typically real time, but they can be shared more openly with a delay. For instance, in the pharmaceutical industry, some data are published openly after a one-year embargo period.

3) **Completeness.**
The full dataset covers millions of data points, but partial views of the datasets (or with limited metadata) can be offered for free or under specific conditions.

4) **Purpose of Use.**
Data can be shared with nonprofit organizations for research purposes, as in the YODA project. But they can be also shared with other actors in the value chain, with a contained risk for competitiveness (see examples of industry data platforms above).

5) **Access-On Demand Without Data Sharing.**
The MIT Open Algorithm project provides remote access and data analysis without actually sharing the underlying data.

In summary. there are ways to make "superstar datasets" more accessible. They can help the ecosystem, but required careful governance - and often the engagement of an independent, trusted third party. Data sharing is a growing phenomenon across industry, and requires some level of infrastructure provided by the dedicated intermediary and it is based on reciprocity. Practice shows that when clear rules are established on which data are shared, what is the minimum entry to the data pool, what organization needs to contribute and who gets the access to the data, organizations are more eager to share the data.

34) Michael Hu and Sean T. Monahan, 'Sharing Supply Chain Data in the Digital Era', *MIT Sloan Management Review*, 57.1 (2015) <https://sloanreview.mit.edu/article/sharing-supply-chain-data-in-the-digital-era/> [accessed 09 November 2018].

3. RECOMMENDATIONS

The music industry has pioneered digitization and generates massive amounts of data, but lags behind other sectors when it comes to ensuring the maximum reuse of these data beyond the data silos: despite the potential opportunities, there is a scarcity of initiatives for data sharing. This piecemeal approach to data innovation is not a problem for delivering incremental innovation today, but in the medium term it stands in the way of disruptive innovation.

There is a need for increased experimentation with data sharing, and CMOs have the opportunity to play a key role. Facing increasing regulatory scrutiny and competition, they can't defend their competitive position simply by holding on to their data. And opening up their data is unlikely to have per se a substantial impact because of the limited value of their data. Their future existence depends on the capacity they have to generate valuable data-driven services to their members and to their customers: by managing their data and by enriching them and by maximizing data flows in the industry.

To remain relevant, CMOs – and especially the innovative ones - should place data at the center of their strategy and aim to play a catalyst role in stimulating greater data sharing across the music value chain: between different CMOs, and with other players in the value chain. Greater data sharing is both a normative goal to increase the footprint of the industry as a whole and a strategic step to re-position the organization in the competitive data economy.

As it should be clear from this policy brief, we are not arguing for indiscriminate opening up of music data. Data are a source of competitive advantage and should be managed carefully. No business can survive by opening up commercially sensitive data indiscriminately. But there are many ways to stimulate innovation through data access and reuse without compromising the competitive position of the various actors.

In this context, CMOs should work on a set of priorities:

1) **Place Data at the Center of Business Strategy.**
CMOs should develop a strategic approach to data management, aligned with the strategic goal of the organization. This includes a "data audit": an analysis of data held by the organization, of its internal and external value, and of the steps necessary to maximize the value of these data. It also implies clear responsibilities for data management, and ensuring data analytics is represented at the most senior positions of the organizations.

2) **Develop Internal Capacity.**
CMOs need in-house competences for data management and analytics. In a data intensive sectors such as the music industry of today, no company can survive without data skills in-house. This is not about creating large scale data science unit, but ensuring the capacity to manage data flows to help navigating the data opportunities with a broad, ecosystem-wide perspective. CMOs are not poised to become big data startups, but should be capable of innovating their services by providing a mix of internal and external data services to its members. To remain relevant, CMOs need to improve and innovate their service provision, in terms of speed, transparency and accuracy. Both members and clients of CMOs are expecting levels of automated, proactive, immediate service to those provided by the leading online platforms.

3) **Share Data.**
CMOs should clarify the goals and the modalities of opening up data, including the identification of which data to open, the curation and publication of data, and the stimulation of reuse. Opening data should be a mean, not a goal. It should help positioning the organization for the future, by enriching their data and providing a better positioning in the data value chain. And it should ultimately benefit the artists. Where open data is not appropriate, discriminatory data sharing solutions should be adopted.

4) **Promote Innovation Across the Value Chain.**
CMOs should stimulate and take part in efforts to catalyze greater data sharing in the industry, in collaboration with the widest range of partners across the value chain: with clients and members; with other CMOs; and with other actors in the value chain (from labels to digital platforms to data analytics companies). This implies analyzing the different modalities for data sharing presented here and experimenting with them: from open data to discriminatory data to "data commons".

4. BIBLIOGRAPHY

Bugge Jensen, Simon, and Marie Christiansen Kroyer.
Polaris Nordic: Digital Music in the Nordics (London: YouGov, October 2018)
https://www.koda.dk/media/184140/polaris-nordic_digital-music-in-the-nordics-2018_yougov.pdf

Cooke, Chris.
"ICE and SACEM Sign up More Indie Publishers for Digital Licensing | Complete Music Update," 2018
http://www.completemusicupdate.com/article/ice-and-sacem-sign-up-more-indie-publishers-for-digital-licensing/

Davenport, Thomas H.
"Analytics 3.0,", *Harvard Business Review*, 91 (2013)
https://hbr.org/2013/12/analytics-30

Dredge, Stuart.
"Jaak Hails Its First Blockchain Music-Rights Pilot with Rightsholders.'" 2018
https://musically.com/2018/05/02/jaak-hails-successful-blockchain-music-rights-pilot-with-rightsholders/

Hann, Michael.
"Music's 'Moneyball' Moment: Why Data Is the New Talent Scout," *Financial Times*, 05 July 2018
https://www.ft.com/content/474ae18a-7f1b-11e8-bc55-50daf11b720d

Hernández, Héctor, Antonio Vezzani, Petros Gkotsis, Mafini Dosso, Sara Amoroso, Lesley Potters et. al.
EU R&D Scoreboard the 2017 EU Industrial R&D Investment Scoreboard (Seville: JRC IPTS, 2017)

Hofheinz, Paul, and David Osimo.
'Making Europe A Data Economy: A New Framework for Free Movement of Data in the Digital Age', *Lisbon Council Policy Brief*, 2017

Hu, Cherie.
"Spotify Wants to Be 'The R&D Department for the Entire Music Industry' -- What Does That Actually Mean?,' *Billboard, 2018*
https://www.billboard.com/articles/business/8376561/spotify-rd-department-entire-music-industry-netflix
International Federation of the Phonographic Industry. Investing in Music (London: IFPI, 2016)

Ingham, Tim.
"Kobalt's AMRA Signs Its First Global Collection Deal... with Apple," *Music Business Worldwide*, 2015
<https://www.musicbusinessworldwide.com/kobalts-amra-signs-first-global-collection-deal-apple/

Milosic, Klementina.
"The Failure Of The Global Repertoire Database," *Hypebot*, 2015
https://www.hypebot.com/hypebot/2015/08/the-failure-of-the-global-repertoire-database-effort-draft.html

Monahan, Michael Hu and Sean T.,
"Sharing Supply Chain Data in the Digital Era." *MIT Sloan Management Review*, 57 (2015)
https://sloanreview.mit.edu/article/sharing-supply-chain-data-in-the-digital-era/ OECD.
Data-Driven Innovation: *Big Data for Growth and Well-Being* (Paris: OECD, 2015)
https://doi.org/10.1787/9789264229358-en

Paine, Andre.
"Solar Activity: Sony/ATV, GEMA, PRS and Solar Extend Pan-European Licensing Deal" 2018
http://www.musicweek.com/publishing/read/solar-activity-sony-atv-gema-prs-and-solar-extend-pan-european-licensing-deal/074278

Pujol, Laia, David Osimo, Jonathan Wareham, and Federica Porcu.
'"Data-Driven Business Models in the Digital Age: The Impact of Data in Traditional Businesses"
(presented at the World Open Innovation Conference, Barcelona, 2016)

Sanchez, Daniel.
"Pandora Opens All of Its Artist Data to the Public Through Next Big Sound," *Digital Music News*, 01 October 2018
https://www.digitalmusicnews.com/2018/10/01/latest-pandora-next-big-sound-metrics/

Summer, Stuart.
"Instrumental: Using AI and Spotify Data to Find the next Big Thing in Music Computing,"
http://*Www.Computing.Co.Uk*, 24 July 2018
https://www.computing.co.uk/ctg/feature/3036430/instrumental-using-ai-and-spotify-data-to-find-the-next-big-thing-in-music/page/3

Van Eeden, Ennel, and Wilson Chow.
Global Entertainment and Media Outlook 2018–2022 (London: PricewaterhouseCoopers, 2018)
https://www.pwc.com/gx/en/industries/tmt/media/outlook.html

TEOSTO

About Teosto
Teosto is a non-profit organisation founded in 1928 by composers and music publishers to administer and protect their rights. The group represents more than 33,000 Finnish and almost three million foreign composers, lyricists, arrangers and music publishers. Teosto collects and distributes royalties to the music authors it represents, for the public performance and mechanical reproduction of their music in Finland. Its website is www.teosto.fi.

theLisboncouncil
Lisbon Council Research

About the Lisbon Council Research
Lisbon Council Research is the scientific, academic arm of the Lisbon Council for Economic Competitiveness and Social Renewal asbl, a Brussels-based non-profit association committed to making Europe "the most competitive and dynamic knowledge-based economy in the world capable of sustainable economic growth with more and better jobs and greater social cohesion," as European leaders vowed to do in Lisbon, Portugal at a landmark 2000 summit. Lisbon Council Research conducts on-going research into Europe's economic and social challenges.
Its website is www.lisboncouncil.net.

ESADE
Ramon Llull University

About ESADE Business School
ESADE Business School is a global academic institution known for the quality of its education, its international outlook and its clear focus on holistic personal development. With strong ties to the business world, ESADE stands out for its innovation and contribution to social transformation. ESADE's greatest asset is its people: faculty, professionals, and more than 60,000 alumni all over the world.
Its website is www.esade.edu.

Copyright © Teosto 2018

CC BY SA

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 licence.

Published in Helsinki, Finland by Teosto

TEOSTO
Finnish Composers' Copyright Society Teosto
Urho Kekkosen Katu 2C
00100 Helsinki, Finland