European Journal of Operational Research 312 (2024) 324–337

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Hit or miss: A decision support system framework for signing new musical talent

Aritad Choicharoon[a,*], Richard Hodgett[a], Barbara Summers[a], Sajid Siraj[a,b]

[a] Centre for Decision Research, Leeds University Business School, University of Leeds, Leeds LS2 9JT, United Kingdom
[b] COMSATS University Islamabad, Wah Campus, Pakistan

A R T I C L E I N F O

Article history:
Received 28 November 2021
Accepted 8 June 2023
Available online 14 June 2023

Keywords:
Decision support systems
OR in entertainment
Decision analysis
Problem structuring

A B S T R A C T

In the music industry, the process of signing new musical talent is one of the most complex decision-making problems. The decision, which is generally made by an artist and repertoire (A&R) team, involves consideration of various quantitative and qualitative criteria, and usually results in a low success rate. We conducted a series of mental model interviews with the aim of developing a decision support framework for A&R teams. This framework was validated by creating a decision support system that utilises multi-criteria decision analysis to support decision-making. Our framework and subsequent implementation of the decision support system involving decision rule and weighted sum methods show an improvement in the ability to analyse and decide on greater amounts of talent. This paper serves as a building block for developing systems to aid in this complex decision-making problem.

1. Introduction

In the music industry, an artist and repertoire (A&R) department is responsible for the process of finding and signing new talent to the label roster so that the label can then market and sell them to the public. This complex and highly volatile decision-making problem of whether a performance should be signed involves several conflicting criteria and is a decision of critical importance to a music label's success. There has also been exponential growth in the supply of music, which has led to a significant increase in the content that A&R teams need to consider. For example, YouTube received 500 hours of new videos uploaded every minute in 2020 (Statista, 2021). There has been a change in the pattern of how music is released and consumed, resulting in a need for labels to adjust the process of finding content. This is illustrated by the introduction of music streaming, as found by Cooke (2020), who stated that 142.9 million albums were consumed in the UK in 2018 with an approximate retail revenue of £1.33 billion, 40 per cent of which came from paid streaming services. The process of finding and signing new talents is also inefficient. Industry statistics have shown that the process is lengthy and expensive; according to the IFPI (2021), signing costs are between £300,000 and £500,000, but can surpass £1 million in a competitive scenario. Of those artists signed, only 5.7 per cent managed to sell 100,000 albums or more, with less than 2 per cent achieving more than 300,000 albums sold.

Overall, music labels, a large provider for this consumption, must continue to release new musical talent that best fits the public's interest (Morris & Powers, 2015) while adapting to the increasing number of musical talents available, changes in the consumption of music, and the inefficiency of their own processes. This paper shows that a decision support system can improve the efficiency of A&R work, and provides a framework for the creation of a decision support system to aid in the decision-making of A&R experts when signing new musical talent. We introduce the use of the mental model approach to understand experts' decisions, propose a framework that encapsulates the criteria of the experts' model, and test this framework in a real use case. Our contributions are threefold: we show how to elicit experts' model of the existing A&R decision-making process for signing new musical talent, we propose a decision support system for the signing of new musical talent, and we demonstrate the use of our proposed framework practically within a UK-based music label.

The rest of the paper is divided into eight sections. After discussing the background literature in Sections 2 and 3 describes our approach to this research. We next investigate the talent identification process through the creation of the mental model of experts in Section 4, based on the mental model interviews. In Section 5, we propose a framework for decision support based on the find-

* Corresponding author.
E-mail addresses: a.choicharoon@leeds.ac.uk (A. Choicharoon), R.E.Hodgett@leeds.ac.uk (R. Hodgett), bs@lubs.leeds.ac.uk (B. Summers).

https://doi.org/10.1016/j.ejor.2023.06.014
0377-2217/

ings from the expert models. Sections 6 and 7 provide the implementation and evaluation of the proposed framework through an application with a small A&R team. The paper closes with Sections 8 and 9, which provide discussion and conclusions.

2. Literature review

2.1. The role and responsibility of the artist and repertoire team

Music labels are businesses that distribute musical talent through various channels including physical, live, streaming, and online. Their goal is to maximise profit by controlling access to both performers and performances that are popular with consumers at large (Janssen & Verboord, 2015). As shown by IFPI (2021), the process of identifying and signing talent that aligns with public interests is a difficult problem, with only 5.7 per cent of signings by major music labels selling at least 100,000 albums and less than 2 per cent selling more than 300,000 albums. To support their business model, music labels have dedicated A&R teams to identify musical talent that aligns with public preferences and to whom they can offer a contract for the right to specific or future performance (referred to in music industry language as signing). This contract allows the label to represent the artist's interests, develop their brand, and sell their music to the public in exchange for financial benefits and other support. Despite an increasing number of performers who are self-publishing their music (Bender et al., 2021), music labels still retain a large presence, in terms of both revenue size and audience reach.

Looking at the research that studies sources and criteria for the signing of new musical talents, there are limited investigations into sources of new talent. Primarily, the work by Negus (2011) identifies radio channels, expert connections, and live performances as sources that are used regularly by A&R in the US and the UK. Other comparable work looking at A&R in the Netherlands, by Zwaan & Ter Bogt (2009), and in Poland, by Galuszka & Wyrzykowska (2017), did not provide any insight into sources for new musical talent, focusing on how A&R evaluates musical talent. As such, we considered sources of new talent to be an area of significant research interest. Looking at the criteria in A&R decision-making. Negus (2011) identified criteria consisting of the musical skill of the performer, music quality of performances, musician's appearance, motivation, and existing size of audiences for the performance. This has been replicated in studies by both Zwaan & Ter Bogt (2009) and Galuszka & Wyrzykowska (2017), although these differ from Negus by discussing the growing importance of the internet medium in measuring the aforementioned criteria, while Negus's work relied mostly on the older physical media of radio, record store sales, and television. Overall, past research found that criteria for A&R can be summarised into music performance and meta-information about both the performance and the performer, but there has been a lack of clear insight into how this can lead to approaches that improve A&R's job.

While there has been limited study of A&R work, there has been more research into the classification of successful talent based on finance. There is a clear link between the A&R job description and the goal of getting music that can be successful. Of particular interest are works that have looked into the use of music information retrieval (MIR), which is a representation of music with quantitative and qualitative features, to identify successful performance. One of the first was by Dhanaraj & Logan (2005), which proved statistically that MIR's features of lyric and mel frequency cepstral coefficients (MFCC), a representation of sound as a power spectrum, can classify financially successful performance with an AUC-ROC curve, a summary evaluation metrics for binary classification that plots the area of true positive rate against false positive rate, of 0.68 on evaluation data. Subsequent work by Herremans et al. (2014) identified beats and timbre as features in popular dance music classification, with a cross-validation AUC-ROC of 0.65, and Yang et al. (2017) used mel-spectrograms to demonstrate an improvement in recall accuracy of 30 per cent in the classification of both Mandarin and Western music. In short, these works show that there is an opportunity to augment A&R preference on musical performance through the use of MIR features.

These researches into a classification based on financial success also look at meta-information around talent. They demonstrate the relationship between previous success, release date, and marketing support with financial success. In particular, Strobl & Tucker (2000) investigated a hundred best-selling albums in the UK. They found album type, seasonal factors, the public perception of performers, and the existing popularity of the performers to be significant factors in success. This is supported by Steininger & Gatzemeier (2019), who showed that public opinion aligns with financial success. On the other hand, Dewan & Ramaprasad (2014) found that marketing investment was also a significant connection to financial success. However, the marketing aspect is beyond the scope of this research as A&R decision-making is done pre-signing and the marketing decision is taken after. Overall, meta-information on talent is highly relevant to the success of the talent, and A&R would be keen to look into it as a set of criteria. However, not all of it would be appropriate given that the decision is made pre-signing.

Overall, this section has looked into literature that has investigated the job of A&R and tasks that are similar to it. The review found a limited quantity of work looking at A&R's processes and criteria for the signing of new musical talent. Studies that have looked into this have focused on the qualitative study of the process rather than trying to encapsulate the decision-making problem to solve the challenges that A&R faces. A&R looks for potentially successful music and we identified more research that has identified information and meta-information on talent that is related to financial success. This research focuses on the classification of talent from others using MIR and information about performers. The work has allowed us to identify a gap in the literature in terms of decision support for A&R that goes beyond qualitative work into the A&R decision-making process. This paper will focus on this and identify factors that are used in the process and how a decision support system (DSS) can be used to support A&R decisions.

2.2. Decision support systems

The current research aims to create a framework for a comprehensive DSS involving talent identification, information management, and the MCDA on criteria and alternatives. This is based on the success of an end-to-end system in aiding complex multi-criteria decision problems similar to those that A&R faces. Studies such as urban transportation policies selection (Arampatzis et al., 2004), waste management (Haastrup et al., 1998), and ore blending cost optimisation (Zhang et al., 2011) demonstrate how comprehensive guidelines for DSS can provide an implementation plan for real-life decision-making problems.

An A&R task, as defined by Negus (2011), is a repetitive decision-making process with the ever-changing importance of criteria, multiple options, and developing alternatives. One of the areas of interest for our research is a decision-rule-based approach towards decision-making. This approach is commonly used to generate a subset of alternatives (Slowinski et al., 2009) from a large number of alternatives, such as the identification of a supplier for a green development (Bai & Sarkis, 2010), fraud alerts in consumer credit (Leonard, 1995), and customer churn predictions (De Caigny et al., 2018). These examples of decision rules demonstrate the use of strict, easily applied boundaries to filter alternatives by eliminating those that do not fit. In addition to filtering via the decision rule, A&R teams also need to prioritise musical performances for signing. There are many families of MCDA methods suitable for this task, the more popular ones being WSM, AHP, ELECTRE, PROMETHEE, MACBETH, MAVT, and MAUT (Velasquez & Hester, 2013). In particular, we were influenced in our work by the use of weighted average methods using satisfaction thresholds in MACBETH, which allows decision makers to establish the importance and a minimum and a maximum expected value for all criteria considered (Bana e Costa & Vansnick, 1994). Given the large number of alternatives and the need to identify several appropriate alternatives, this work has elected to focus on the weighted sum approach. These methods are accessible to the tool's user but can be sufficiently complex to handle the complex problems that A&R faces. It has also been shown that prototypes are a useful step in ensuring that the framework can efficiently support the decision-making process (Quintero et al., 2005), so our framework will be tested to ensure it represents experts' explicit beliefs. In this section, this review identifies various MCDA techniques that could be used in the creation of DSS for A&R. It also discusses the potential strengths and weaknesses that A&R would face with different MCDA approaches to the DSS and this leads to the need to clearly identify A&R's mental model to choose an appropriate MCDA and DSS framework.

2.3. Mental model

To understand how A&R considers talent, this research uses a mental model approach. A mental model is an internal construct of reality that is unique to individuals and the situations that they face, including their assumptions, beliefs, experiences, and biases. It may not truly reflect reality and is only explicitly known to the individual but it has an impact on their decision-making process (Chermack, 2003). Research has utilised a mental model to understand a variety of choices taken by human decision makers, such as human-computer interaction (Carroll & Olson, 1988), water-management (Kolkman et al., 2005), human resource management (Chermack, 2003), and risk communication (Bravo-Lillo et al., 2010; Jungermann et al., 1988; Otto-Banaszak et al., 2011). The model itself allows non-decision makers to understand the criteria that influence experts' decision-making, which is central to MCDA. To elicit a mental model, Jones et al. (2011) separated the process into two categories: direct and indirect elicitation. Direct elicitation uses a diagrammatic interview based on the assumptions that the model can be represented as a network of beliefs and that the individual can define it. This allows for direct control to cover areas of interest but has the limitation of potentially ignoring information that does not fit into preconceptions. In contrast, indirect methods require researchers to infer the network from open-ended oral interviews, allowing for exploration during the interview but requiring a skilful interviewer to maintain control. As A&R decisions are a multi-criteria decision-making process, indirect methods are preferred for their flexibility.

In this research, the construction of expert processes to identify and evaluate musical talents is fundamental to creating a framework for a DSS. This review has identified an appropriate method to do so as well as the importance of such a process.

3. Methodology

This research utilises indirect elicitation to extract a mental model of how A&R signs new musical talent. The mental model is used to create a framework for a decision support system to assist with this process. The framework is then evaluated, both qualitatively and quantitatively, through a real-life implementation in a music label.

3.1. Mental model interviews and framework elicitation

The interviews were based on the indirect elicitation of a mental model with two interviewers. The lead interviewer guided the conversation in accordance with the interview protocol, while the second interviewer focused on clarifying the questions and ensuring that an appropriate response was collected for each question. The open-ended interviews started with a broad questioning that permitted free responses, allowing experts to guide the conversation to a specific area of interest, such as the interviewee's definition of success or personality traits of the artist, that came up during the interview (Morgan et al., 2002). Probing questions were then used based on the responses, focusing on *"how"* and *"why"*, for example: *"Can you explain to me what the role is of an artist and repertoire expert?"*; and then probing questions like *"How do you differentiate between different types of music?"* or *"Why do you focus on an artist's live performance ability?"*. A detailed interview protocol is provided in the supplementary material.

We initially reached out by email to 15 experts who worked for music labels that are based in the UK and had signed new musical talent within six months before that email if they would be interested in participating in an interview that studies the working of A&R. Subsequently, a total of ten A&R experts, from various backgrounds, nationalities, and responsibilities, as listed in Table 1, participated in this research. The final number of interviews was determined based on data saturation, where a new expert is interviewed until no additional beliefs and goals are identified. Content analysis was then used to determine patterns of the shared mental model. Interviews were analysed with a coding process to identify new and existing concepts (Abel et al., 1998). Once all interviews had been coded, shared concepts across interviews were identified and tagged accordingly. The final list showed all identified concepts, how many interviewees discussed them, and synonyms for each concept. These shared concepts form an understanding of the process and identified criteria that experts used to evaluate new musical talent and become the basis for our framework.

3.2. Testing by implementing the framework

To evaluate the framework for the DSS, it was implemented with a small A&R team in the UK. The team consisted of five people, with a mix of experience, focusing on dance music for UK audiences. This music label focuses on talent as individual performances rather than performers owing to their size, financial considerations, and focus on dance music. The DSS was implemented to aid in the whole process, but the A&R staff retained the right to make the final decision on a signing. A&R was asked to utilise the DSS for three months before both qualitative and quantitative feedback was gathered. The quantitative numerical indicators included the ability to manage the large quantities of performances that the experts are unable to effectively evaluate at present, the ability to analyse the performances in a timely manner, and the ability to evaluate performance in equal or better quality than human evaluations of the musical talent. We used two aspects of quantitative and qualitative analysis to achieve this evaluation. Quantitatively, the framework was measured and compared to human evaluations on the following quantitative numerical indicators:

1. Quantity of performances that were analysed, in comparison to the current manual processing by experts in the same period.
2. Quantity of performances that were evaluated and considered of suitable quality for detailed evaluation, in comparison to current manual processing by experts in the same period.
3. Number of successful recommendations of a musical performance that resulted in a signing within a month.

Table 1
Detailed information of participants in this study.

ID	Job title	Experience	Initial role	Education background
1	Director of A&R	20–30 years	Intern	Bachelor's degree (non-music)
2	Senior A&R manager	10–20 years	Intern	Bachelor's degree (non-music)
3	Senior A&R manager	5–10 years	Intern/Scout	Bachelor's degree (non-music)
4	Director of A&R	30+ years	Intern/Scout	Secondary Level
5	President of music label	30+ years	Intern	Bachelor's degree (non-music)
6	Head of A&R	20–30 years	DJ	Bachelor's degree (non-music)
7	A&R manager	10–20 years	Scout	Bachelor's degree (non-music)
8	A&R manager	5–10 years	DJ	Bachelor's degree (music)
9	Director of music label	20–30 years	Scout	Secondary level
10	Director of music label	30+ years	Scout	Secondary level

Table 2
Expert's (horizontal axis) alignment in the mental model of aspects of talent discovery and signing (vertical axis) with shaded square indicating topic is discussed.

List of Topic and A&R Staff			1	2	3	4	5	6	7	8	9	10
Scouting		Traditional Scouting	X	X		X	X	X		X	X	
		Digital Scouting	X	X	X	X	X	X	X	X	X	X
		Recommendation from Network	X	X	X	X	X	X	X	X	X	X
Initial/Intuitive Evaluation		Quantity of Audience	X	X	X	X	X	X	X	X	X	X
		Quality of Performance	X	X	X	X	X	X	X	X	X	X
Detailed Evaluation	Music	Audience Reaction to Music (Quantity)	X	X	X	X	X	X	X	X	X	X
		Audience Reaction to Music (Qualitative)				X	X	X	X	X	X	
		External Influential Opinion on Music	X	X		X		X		X	X	X
		Internal Influential Opinion on Music		X			X		X	X	X	X
		Music Qualitative	X	X	X	X	X	X	X	X	X	X
		Music Memorability				X	X	X	X	X	X	X
		Uniqueness/Fit With time	X	X	X	X	X	X	X	X	X	X
	Artist	Artist's Visual	X		X	X	X	X	X	X	X	X
		Artist's Background and Story	X	X	X	X	X	X	X	X	X	X
		Artist's Follower	X	X	X	X		X	X	X	X	X
		Artist's Past Work		X			X			X		X
		Audience Reaction	X		X	X	X		X			
		Working Mentality	X		X	X	X	X	X		X	X
		External Influential Opinion on Artist	X					X			X	X
		Internal Influential Opinion on Artist		X			X	X	X	X	X	X

To qualitatively assess it, the researcher asked the expert for their opinion on how the system fitted into their workflow and how they assessed the framework's ability to evaluate talent compared to how A&R had done it in the past. The assessment was on both initial evaluations and detailed evaluations.

4. Mental model of A&R in the process of finding and signing new musical talents

4.1. Mental model of A&R's role and responsibility

Based on mental model interviews, Table 2 lists how respondents in the research (horizontal axis) discussed each aspect of talent discovery and signing (vertical axis). This showed that A&R worked in two stages: the scouting process to discover new musical talent and the evaluation process to decide whether to sign them. While they have slightly different scouting and evaluation criteria, it can be seen that the majority share similar considerations. Experts describe their job as being responsible for finding new musical talent that can connect with the public audience, or essentially identifying "hit records". This is supported by this excerpt from Expert #9: *"A&R is about finding a hit. Finding success in music... we have made a living for many years based on finding music that will connect with a bigger audience than a small one."* It must be noted that musical talent, as discussed by experts, includes performance, an album collection of multiple performances, and musical performers. This can be seen in the following response from Expert #9: *"It is about finding tracks and artists that we think will connect with the public."* Given the definition, it is important to note that interest in performances versus performers varied based on the music labels' goals. For instance, dance or electronic music labels are more interested in individual performances.

4.2. Mental model of finding talent

Experts used a variety of sources and methods to find new musical talent. Expert #8 described this as: *"For me, A&R scouting is about utilising the platform and the place that has new artist and new* *talent that these people were able to put their record up... There are just more platforms now."* They discussed various combinations of locations, situations, individuals, and methods that have led to the discovery of new talent. This research established shared themes from these interviews on methods of identifying new musical talent and categorised them into three main categories: traditional scouting, digital scouting, and recommendations from their network. These three categories are discussed in detail in the subsections below. In the scouting process, experts emphasised that the growth of the internet has increased the number of products that are available for digital scouting and indicated the growing importance of this over other methods.

4.2.1. Traditional scouting

Traditional scouting, according to experts in mental model interviews, encompasses in-person activities that A&R uses to discover new talent and includes activities such as attending live performances, listening to music on the radio, and reviewing busking performances. The differentiation of traditional scouting from other methods is that it has been used consistently throughout the history of music labels and requires no additional tools or individuals beyond the expert themselves. For example, Expert #9 stated: *"The conventional method of going to gigs is a huge part of talent scouting... I think there is always something different about going and seeing somewhere live because you can browse reaction and see what is actually happening with the artist with their music"*. However, the actual day-to-day process can be influenced by experience, location, and the expert's understanding of the music industry and by knowledge of the environment such as knowing which venues to attend, which radio channels to monitor, and which cities have new performances. This method has become less important because of limitations in terms of geographical and timing constraints as well as the growth in the importance of digital scouting and the efficiency of network recommendations. Additionally, this form of scouting was more restricted during the Covid-19 pandemic when live performances were either banned or poorly attended.

4.2.2. Digital scouting

The majority of experts stated that digital scouting was central to their day-to-day work. It was described by Expert #5 as the process of finding talent using various sources on the internet: *"Searching through new music that has been released online to SoundCloud, YouTube, whatever that may be. There are a lot of acts out there that can be heard by doing that type of A&R. There is also a lot of talent out there to be captured."* It should be noted that, in most experts' descriptions, digital scouting shares a lot of similarities with traditional scouting. Tasks that were once accomplished in traditional scouting are replicated online, with the internet as a new medium allowing access to talent beyond the geographical and time restrictions of traditional scouting. However, sources in digital scouting are not limited to replication of traditional scouting such as consumption platforms like Spotify and SoundCloud; there are also new ways of finding new talents such as social media like Facebook or TikTok, internet TV on YouTube, and blogs like Tumblr. Also, as the internet retains talent and performances for longer, digital scouting has greater availability of content, although this greater quantity also provides a challenge in dealing with the volume of information during the initial evaluation of content when compared to other methods with smaller pools of content, which will be discussed in Section 4.3.1.

4.2.3. Recommendation from a network

Recommendations from network connections within the music industry are a longstanding source of new musical talent. Expert #10 said: *"Because we have been around for a long time, we get approached by most people and most people know that we are around to send a track to."* Connections such as music professionals, artists' managers, lawyers, DJs, and radio hosts are exposed to potential music talent as part of their job and can bring these to A&R's attention. They can recognise A&R's preferences and suggest talent for consideration that closely aligns with their interests and expertise. However, talent identified by this network is heavily dependent on A&R staff having a good network within the music industry resulting from their length of time working in the industry and social skills, as they need to build a network before they can make use of them.

4.3. Mental model of evaluation on music talent

The expert model of A&R's evaluation, as shown in Fig. 1, shows that their signing decision is based on an assessment of their potential to become a hit. There is, however, no universal consensus on the definition of a hit. Various quantitative and qualitative definitions have been proposed. The quantity has included measures such as audience size, streaming size, and physical sales. Aside from the measurement, the actual quantity that determined whether talent would be a hit was also subjective to the individual expert. For example, Expert #2 stated: *"I would look at stream and said that is a hit song and then I would look outside at the single chart and I think that make me slightly different from other people since I think the chart is still important. Some people nowadays go for stream only... because there are less song in chart and they stay for much longer now"*. Other definitions are based on qualitative criteria such as star quality and mega personality. These traits and descriptions come from individual A&R staff members' preferences and cannot be easily quantified but they are integral to the decision to sign new musical talent, as described by Expert #2: *"The world is full of talent, but not full of stars. So many times, they can definitely come in and sit with you and they sell you on the idea of them as a star. We are all here because of the star. That is what we are looking for."*

Beyond the definition of a hit, there are also other business factors that affect the evaluation of musical talent including personal preferences, budget considerations, the genre of focus for the music label, the choice of performance or performers, and perceived risks. Expert #7 described this in terms of *"balancing the roster"* referring to the idea that the expert may put the music label's interest in talent over consideration of future success: *"I think as [part of] a label, we are responsible for balancing their roster, but at the same time, I think individuals taste normally bring in artists that are more successful, we want to work with what we are passionate about."* This aspect of music labels plays a role in their consideration, but the focus of their evaluation, as shown in the mental model, is the likelihood that talent will become a hit.

Overall, the resulting expert model of evaluation criteria is summarised in Fig. 1. The two stages of evaluation coincide with a system of thinking as described by (Kahneman, 2011): fast, intuitive thinking (System 1) used for sub-conscious decision-making, and detailed logical thinking (System 2) used to solve a complex problem. The decision process starts with fast decision-making that filters talent with a low likelihood of success, followed by a more deliberative process that estimates the future success of the remaining musical talent considering information about the performance and performers. The performance is central, but information about performers also has an important role as described by Expert #3: *"We sign the artist, not just the track. First, you would listen to music, but then you would want to meet the person to understand who they are and get the measure of them."*

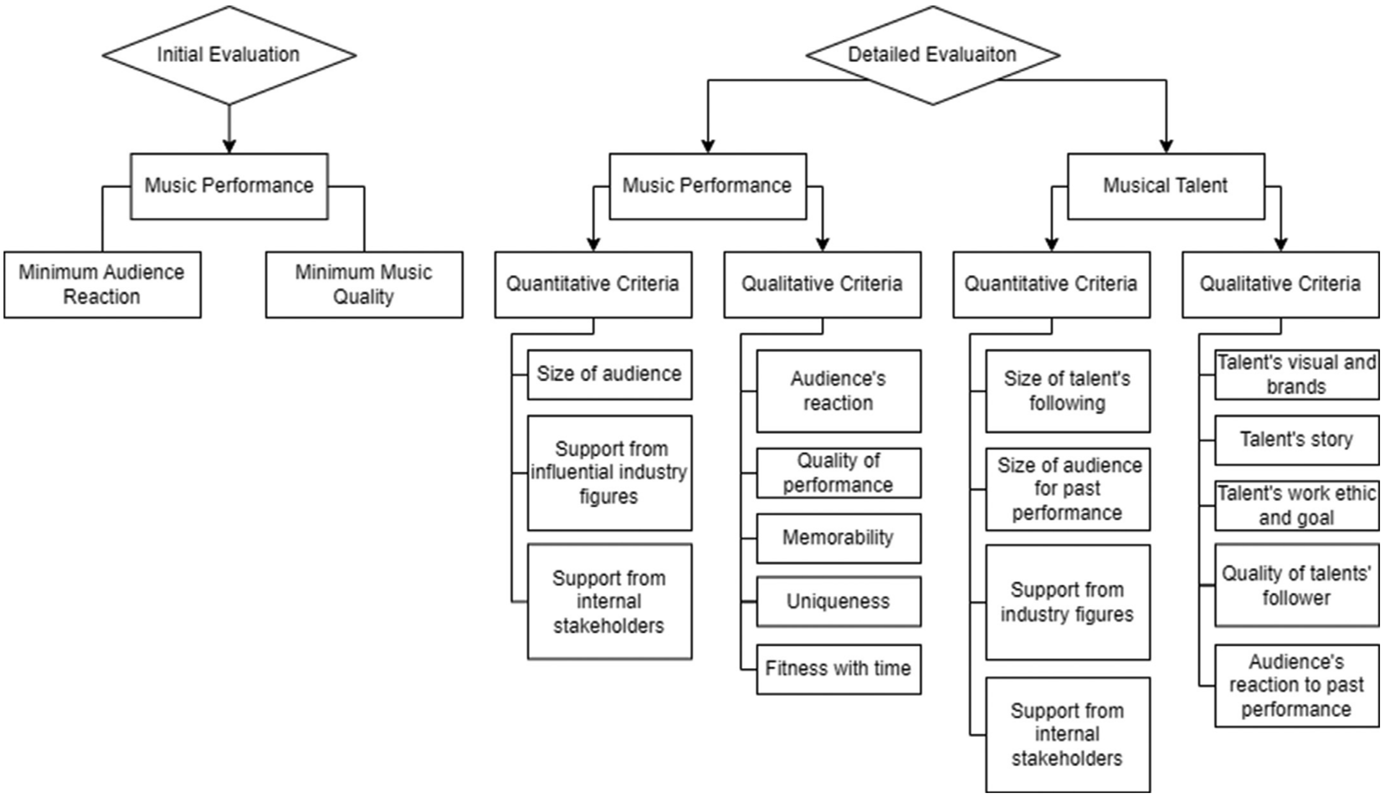

Fig. 1. Discovered mental model of A&R evaluation of musical talent likelihood of success.

Table 3
Criteria during the initial evaluation of the likelihood of success.

Trait of interest	Description	Variables that are used to measured and define trait	Description of variable
The minimum quantity of audience reaction	Performance that generates enough support from a public audience is considered for further evaluation	Quantity of audience, discussion, comments, etc	Quantity of consumption and discussion on online and offline platforms about musical performance
Minimum quality of musical performance	Music that has inappropriate quality in terms of production sounds are likely to be unsuccessful	A&R's judgement	A&R evaluate musical performance quality as sufficient for further evaluation

4.3.1. Initial evaluation of musical talents

The initial evaluation focuses on the talent. Experts want to filter out those that are unlikely to be successful. The process focuses on issues such as the minimum quantity of audience and minimum quality in the performance, as evidenced in the description by Expert #6: *"Usually, most of the stuff is not of interest, I can easily reject eighty per cent of the content. It is just too much, but the rest, unless there is a clear indicator, then it is about the gut, it just has to be gut."* It is important to note that, while the expert above may have described the next stage of decision-making as based on gut, the same expert testimony in a later part of the interview suggests otherwise, as can be seen from their responses in Table 2. Overall, as described in Table 3, experts use the heuristic of a minimum musical quality in performance or a minimum quantity of audience, such as the number of audiences on a streaming platform as the decision rule for whether to consider talent in detailed evaluation.

4.3.2. Detailed evaluation of musical performance

A detailed evaluation of talents is akin to System 2 thinking and a critical part of decision-making. The evaluation of the talent's likelihood of becoming successful involves both qualitative and quantitative criteria. It includes information surrounding performances that must be acquired to support the evaluation process, through either direct measurement or secondary information. Looking at quantitative information, the reaction from audiences is one such criterion that can be assessed through measurements such as the reaction to live performances, views on streaming platforms, and sales of music. Audience reception is an important factor that can be directly associated with potential revenue. Experts compare this with past performances, using these as benchmarks of both past success and failure. Benchmarks are chosen based on the similarity of genre, composition, and origins of the performance. These quantitative comparisons of audience response are crucial as described by Expert #3: *"Definitely, I do rely on my experience... maybe the one that did not work and said okay, I learned my lesson with that, but everything is always changing."*

A&R also looks at the quantity of support from programme schedulers at radio stations, social media influencers with large followings, content managers on streaming platforms, etc. These people have experience and an understanding of public audiences' tastes as well as the ability to utilise their platform to influence the public perception of musical talent. These support figures are used as indicators of success, as described by Expert #8: *"I always check a couple of playlists on YouTube and SoundCloud on Monday morning. Just to keep tapped of what is in-trend right now. Not a lot of playlists, mind you, but there are a couple that I know I can rely on to understand what is going on."* Experts also quantified support from

Table 4
Quantitative criteria in detailed evaluation of musical performance.

Trait of interest	Description of trait	Variables	Description of variable
Size of audience	Performance with a large audience and greater likelihood of public discussion	The quantity of audience and discussion	The quantity of audience and discussion can be measured and assessed through social media, physical sales, and streaming
Support from influential industry figures	Performance with a large quantity of support from influential figures such as DJs and radio hosts are more likely to become successful	Quantity of influential figures' supporting the talent	Level of articles and discussion from influential figures on social and news media
Support from internal stakeholders	Music that is favourably evaluated by a large number of internal experts is more likely to be successful	Quantity of internal stakeholder support	Other expert viewpoints are gathered through direct discussion or through interest from competitors.

Table 5
Qualitative criteria in the evaluation of musical performance.

Trait of interest	Description of trait	Variables	Description of variable
Audience's reaction	Performance that generates a positive reaction from the public audience is more likely to be successful	Subjective assessment of audience reaction	Quality of audience reaction can be directly measured or through social media, radio, physical stores, and streaming
Quality of performance	Performance with appropriate quality of production is more likely to be successful	Subjective judgement	Experts evaluate musical performances for quality of production focusing on vocals, instruments, and tone
Memorability	Music performance that is memorable is more likely to be successful	Subjective judgement.	Experts evaluate musical performances for memorability and evaluate digital/real-life repetitive use of performances in video, audio, and other forms
Uniqueness	Performances that have a different sound from other performances are more likely to be successful	Subjective judgement.	Subjectively evaluate performance for the uniqueness of the sounds.
Fitness-with-time	Performances that fit in with other popular musical performances are more likely to become successful	Subjective judgement	Experts evaluate musical performances and compare them with other works that are currently popular with the public

within the music label, such as their superiors, peers, and subordinates. This reliance on selective influential figures internally could potentially result in confirmation bias in that A&R could seek opinion that is aligned with their thoughts and ignore others. However, experts believe in the potential validation and improved accuracy of this information. The majority of experts relied on partnerships with trusted colleagues in their assessment. Overall, Table 4 summarises these aforementioned quantitative criteria to evaluate musical performances in detailed evaluation.

As quantitative criteria may not be available, such as if the music has not been publicly released, experts also consider qualitative information. There is a competitive and monetary advantage in the ability to evaluate talent directly before other music labels that rely only on the audience and external validation. The overview of these criteria is listed in Table 5. Experts compare these qualities with similar performances to assess the likelihood of becoming a hit. The first of the criteria is the quality of music. This can be best described by a statement from Expert #6: *"If I can say that it sounds like a hit, then that is just the quality of the song, it is just as simple as that."* The quality is believed to be easy to ascertain as part of the initial evaluation in the mental model interviews, but the interviews also show that the descriptions of preferred quality were highly subjective and varied depending on the expert and genre of interest. Examples included lyrics, the balance of the audio, the skill set shown on various musical instruments, and the vocal ability. Another qualitative criterion of performance is memorability. This was described in terms of whether the performance can be recalled in a person's mind and embedded in their memory after their initial listening. This ability to capture audience attention beyond the listening period is a trait of successful performance according to Expert #4: *"This music would have a hook that is memorable... If I can hum to a song after a few listens through, that is an interesting starting point."* This is also known as the *"stuck song"* syndrome and ear-worm effect. There are methods of assessing this memorability such as measuring the repeatability of the chorus of the performance by an audience and subjective observation of the public mentioning it through media, social media, and real-life conversations.

The last two qualitative criteria are uniqueness and fitness with time. Uniqueness is a feature that indicates dissimilarity to other performances. The ability to stand out in a crowded market is of interest, as described by Expert #7: *"There are normally three things that I look for, one of them is a unique vocal. It does not have to be the greatest most technical vocal ever, but it should be recognisable by people on streaming in 30 seconds."* However, experts considered this in balance to fitness with time, which describes how the performance must not be entirely out-of-place with current trends. Hit songs should be in alignment with the current sound and trend in music, adopting instruments, tempo, beats, or vocals that are widely popular with the audience. Overall, this creates a conundrum in which experts adjust their assessment of the likelihood of success based on information about the uniqueness of the performance to take into account the likelihood of it being out-of-touch with the public taste.

4.3.3. Detailed evaluation of musical performers

At this stage experts also consider quantitative and qualitative criteria on performers such as public perception, their story, exist-

Table 6
Quantitative and qualitative criteria in the detailed evaluation of musical talent.

Trait of interest	Description of trait	Variables	Description of variable
Size of talent's following	Quantity of followers on mediums such as social media or streaming platforms.	Quantity of Followers	Quantity of followers on mediums such as social media or streaming platforms.
The size of the audience for past performances	Performer's past work, both successful and unsuccessful projects, will influence the assessment of future success	Quantity of audience reaction	Quantity of the audience on the performer's past work.
Support from industry figures	Artists that are favourably rated by influential figures in the industry are more likely to be successful.	Quantity of support	Number of discussions from influential figures, such as radio and playlist organisers is important
Support from internal stakeholders	Artists that are favourably rated by the decision maker in the company will have more support if signed.	Quantity of support	Number of support from the decision maker in management can allow for greater monetary support
Talent's visual and brand	Performers that aligned with public perception are more likely to be successful.	Subjective judgement	Performers dressed or present themselves in alignment with public expectation. This can also be measured through market research
Talent's story	Performers with an appropriate story increase their probability of becoming successful	Subjective judgement	Performer's background and fitness for use in a musical performance. This can also be assessed through comments on social media and market research
Talent's work ethic and goal	Working mentality and desire to be successful.	Subjective judgement	Subjective evaluation of artists through discussion and observation.
Quality of talents' followers	Reactions from followers to performer's communication on social media and traditional media.	Subjective judgement	Reaction from followers to performer's communication on social media and traditional media.
Audience's reaction to past performance	Performers with a positive energetic following is more likely to be successful	Subjective judgement	Audience reaction to performer's past work on, quality, and similarity with new performances

ing following, past performances, ability to perform live, and working mentality. These are detailed in Table 6. This information, indicating how the public perceives talent, includes important factors in evaluating the likelihood of talents becoming a hit as this is described by Expert #6: *"I think about the media, who would be my champion in the media, who would support it from online press to radio, to get into radio eventually".* The first criterion considered around performers is public perception which has been a long-standing factor and has grown in importance due to the expansion of the internet as described by Expert #8: *"A track might not be enough... There were a lot of projects that I would not be surprised to say that we like what you are doing, but we need to see more of you as an artist [before making a decision]."* Experts considered how this perception could enhance experience and fondness for the talent by looking at a performer's characteristics such as presentation, personality, and style. They also look at the background story which could shape an emotional connection between the audience and the performer. These are evaluated through the connection between public perception of the story that the performer has and their musical background. Examples included a hip-hop artist aligning in their visual style and story with audience expectations of hip-hop artists, an exhibition of unique standout content in social media, and a story that resonates with their key audience.

The next criteria are both quantitative and qualitative. The size and sentiment of existing followers are a measurement of how successful the performer has been in the past. Performers with a large active passionate existing following are more likely to be successful as they have demonstrated evidence in the past and could benefit from existing sources of support to generate revenue on a new project. Similarities and differences between performances under evaluation and those that were released in the past are also looked at to identify how past performances can potentially influence a new release. Experts believe that performance that is similar to past success is more likely to be successful. Experts also assessed the ability of the artist to deliver in an engaging manner with an audience during a live show, which is of great importance even during Covid with real in-person performance being replaced by live streaming with video.

Work ethic is another important consideration as described by Expert #5: *"I have worked with* [redacted] *. Great musician, but he is happy to just stay at home and enjoy the money that he got from his first release rather than travel around on a tour. You can't force them to change who they are."* A performer that is passionate and driven is more likely to be successful, either by quality or quantity of performance released. Finally, opinions of other internal and external experts on performers also impact the evaluation, similar to the assessment of the performance itself. Experts rely on figures in the industry to tune their judgements and fill in the missing information. Overall, qualitative and quantitative information about the performers, have an impact on the assessment of future success.

5. Framework for the decision support system in the musical talent signing process

Based on the experts' mental model, we propose a framework for a DSS to support A&R experts in signing new musical talent as shown in Fig. 2. It includes a process to discover new musical talent and the evaluation of potential talent to inform the decision on whether to sign them. The framework itself can be used to help design a future DSS for any music label.

5.1. Framework on the scouting of new musical performance

The framework for scouting, shown in the first half of Fig. 2, distinguishes between digital scouting, where the DSS should be directly involved in the process, and manual addition by experts of talent from network recommendations and traditional scouting. Digital scouting presents a significant opportunity to achieve benefits from a decision support system because of (1) the interest in this from experts, (2) the ease of interaction between a DSS and these sources in comparison to the time-consuming nature of direct engagement by the expert, and (3) the fact that the current A&R process can be easily replicated. A DSS will simply automate and scale the job that experts are doing to a larger pool of talent. For example, monitoring for new talent on streaming platforms

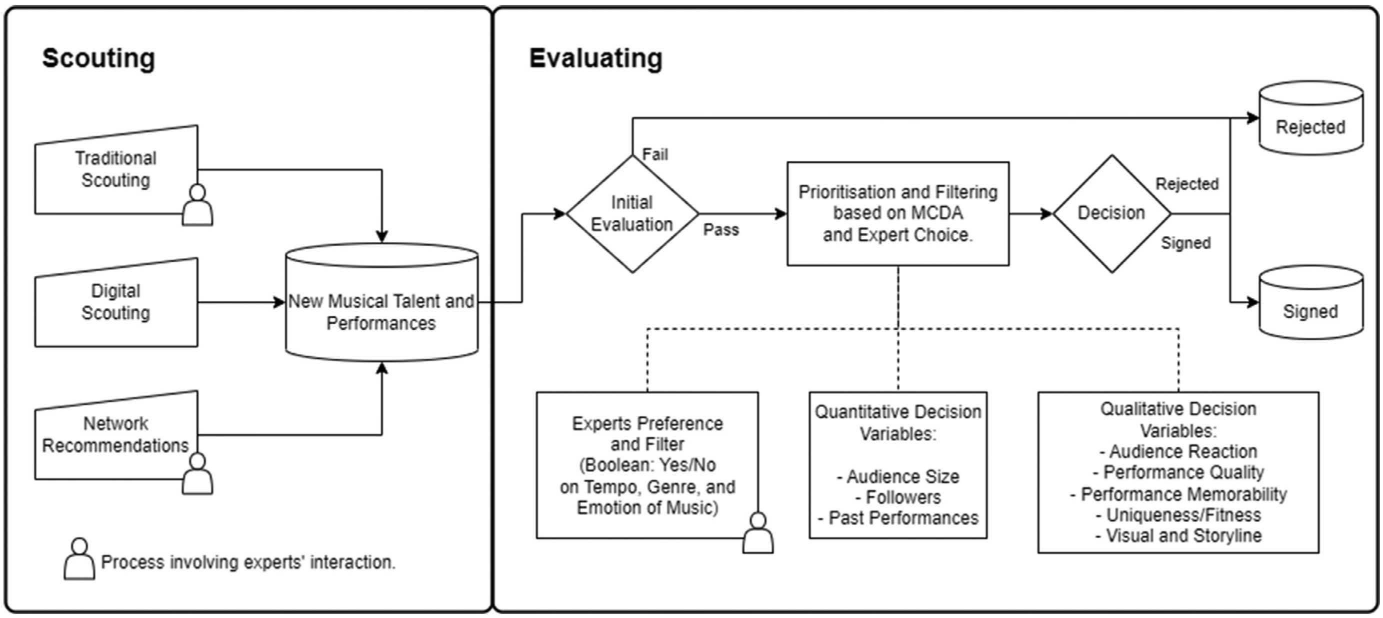

Fig. 2. Proposed DSS framework on A&R signing of musical performance.

Table 7
Type of source in digital scouting and implementation framework for DSS.

Source	Description of the source of interest	DSS implementation policy in framework
Streaming platform	Platforms to temporarily rent music to audiences such as Spotify, SoundCloud and Deezer	Good source, new talent can be discovered by referencing the date that they joined the platform. Should be cross-referenced with social media and distribution platform
Distribution platform	Platforms to digitally sell performances such as iTunes and Google music store	Good source, new talent can be discovered by referencing the date that they joined the platform. Should be cross-referenced with social media and streaming platform
Social media	Platforms for sharing opinions and social media such as Facebook, TikTok, Instagram and Twitter	Focus on popular discussions of interest to experts. Limited source so DSS should focus on gathering names, distribution, and streaming platforms to expand upon
Other platforms	Any potential source of information that may introduce new talent such as online stores for tickets to live performances.	Very limited source. DSS should focus on gathering names, social media, distribution, and streaming platform locations.

based on release dates can be automated with a DSS. Table 7 summarise various digital scouting sources and how a DSS would deal with them.

5.2. Framework on initial evaluation of new musical talent

The DSS, based on the expert model, should adopt a decision-rule approach in the initial evaluation to filter performances based on Slowinski et al. (2009). This initial evaluation ensures that the talent has a minimum quantity of existing support or a minimum quality of audio according to criteria set by the expert. First looking at existing support for talent, the DSS extracts audience qualitative and quantitative information such as the number of streams and sentiment on the talent from sources such as social media, streaming platforms, and distribution platforms. In terms of the quality of music performance, MIR can be used to extract qualitative features of the performance as a quantitative representation such as a vocal note, bpm, and tone. These can then be used as features for machine learning models to classify performances in terms of fitness based on experts' past decisions.

It is important that these decision rules are grounded in the context of time and performance type, such that the value of minimum aggregate streams in the decision rule is in comparison to the days that the talents were available to the public with adjustment required for a longer or shorter period of evaluation. Equation (1) represents a decision rule, with 1 for $Pass_{IE}$ representing talent that has the potential to be successful and 0 representing talent that does not. The criteria of audience reaction (for example minimum quantity of streams on Spotify at time x) and minimum quality of talents (for example alignment in chords) are Ar_n and Mq_m, respectively. This is compared to an expert's threshold of τ_{Mq} and τ_{Ar}. W_{Ar_n} and W_{Mq_m} are the importance of audience and music quality as decided by the decision maker.

$$Pass_{IE} = \begin{cases} 1 & \text{if } \sum_{n=i}^{j} W_{Ar_n} Ar_n + \sum_{m=k}^{l} W_{Mq_m} Mq_m > \sum_{n=i}^{j} W_{Ar_n} \tau_{Ar} \\ & + \sum_{m=k}^{l} W_{Mq_m} \tau_{Mq} \\ 0 & else \end{cases} \quad (1)$$

Given two criteria are included, the importance of W_{Ar_n} and W_{Mq_m} can be adjusted so that one or the other criteria could be sufficient for the talent to pass this initial evaluation. Given that audience reaction can change over a period of time, this initial evaluation process is done repeatedly on each talent as seen in Fig. 2. These initial tasks, as currently done by a human expert, are time-consuming and prone to errors due to time limitations and the volume of material to evaluate. The DSS will allow A&R to focus their energy on detailed evaluation assisted by subsequent prioritisation and information from the DSS.

5.3. Framework on detailed evaluation of new musical talent

In detailed evaluation, the role of the DSS as shown in Fig. 2, is to prioritise talent based on quantitative and qualitative crite-

Table 8
Quantitative criteria for musical talent's success and the role of DSS as suggested in this framework.

Trait of Interest	Description of trait of interest	DSS Application
Audiences' Quantitative Reaction	Audience quantity measured online and offline such as Spotify stream, YouTube play and ticket sales	Data collection. Generate statistical comparison and benchmark from a specific time and groups using historical data
Followers	Followers or supporter base can be measured such as social media support and membership of Patreon	Data collection. Generate statistical comparison to benchmark from specific time and group using historical data
Past reaction	Previous support based on other performances	Data collection of quantity and established benchmark for comparison based on historical data

Table 9
Qualitative criteria for future musical talent's success and the role of DSS as suggested in this framework.

Trait of Interest	Description of trait of interest	DSS Application
Audience qualitative reaction	Sentiment relating to talents/performances on social media and reaction during live performances	Data collection and sentimental extraction using natural language (NLP) processing and computer vision
Music quality	Subjective quality of performance in terms of vocals and instruments, individual and in collaboration with each other	Data collection and music information retrieval to imitate expert judgement of the quality of performance
Music memorability	Subjective memorability of performances, particularly on a hook/chorus portion of the vocal and rhythm throughout.	Using acoustic analysis of MIR and NLP to identify memorability through similarity with benchmarks.
Uniqueness/Fitness with time	Subjective uniqueness contrasting with alignment to the currently popular practice	Data collection and summarising of the current musical trends based on the popular charts and social media along with MIR and NLP which can be used for comparison
Performer's brand and back story	Subjective assessment of talent's background and aesthetics that align with public taste and expectation	Data collection and summarisation of current and performer characteristics through natural language processing and computer vision

ria that the expert considers crucial for future success before presenting these along with necessary information for A&R consideration of signing. Starting with quantitative variables based on the expert model, the DSS should consider audience reaction, followers, and past performances. These quantities are described in more detail in Table 8 and must be considered in the context of time such as aggregate, recent, and future values. They must also be normalised using a comparable product or target as a benchmark. An example would be the aggregate, recent, and future number of streams for performances on Spotify, normalised to values from a benchmark of similar music performance. For qualitative variables, the DSS should transform unstructured data into quantitative variables measuring audience reactions, quality, memorability, uniqueness/fitness with time, and the back story of the talent. Examples would be social media text representing audience reaction, vocal suitability to the music performance, and talent's representation in live performances. This can be done through natural language processing, computer vision, and MIR. Table 9 shows these traits and methods to extract and quantify such information.

Aggregated methods such as weighted sum are preferred over other methods owing to their ability to deal with various criteria, their interpretability, and their computational speed, which allows for re-evaluation over various criteria choices. Other methods that were considered, such as outranking, are discussed in detail in the discussion section, Section 8. Equation (2) represents such evaluation, with variables normalised to benchmarks and weighted for importance by the expert. Quantitatively, Ar_n represents audiences' reactions to the performance, which may be further split according to time comparison and the platform of interest, Fo_n is performers' followers, and Pr_n is the audience reaction to previous performances. Qualitatively, Qr_n represents audience reactions, Mq_n represents the judgement of music quality, Mm_n represents memorability of music, $UnFt_n$ is the variable that represents uniqueness and fitness with time in relation to each other, and Ps_n represents the talent's story. W represents the importance of each criterion for the expert.

$$Success_{Likelihood} = \sum_{n=i}^{j} W_{Ar_n} Ar_n + \sum_{n=i}^{j} W_{Fo_n} Fo_n + \sum_{n=i}^{j} W_{Pr_n} Pr_n + \sum_{n=i}^{j} W_{Qr_n} Qr_n + \sum_{n=i}^{j} W_{Mq_n} Mq_n + \sum_{n=i}^{j} W_{Mm_n} Mm_n + \sum_{n=i}^{j} W_{UnFt_n} UnFt_n + \sum_{n=i}^{j} W_{Ps_n} Ps_n \quad (2)$$

In presenting this score, representing the likelihood of success, the DSS prioritises talents for consideration to A&R based on this likelihood with higher scores needing more urgent evaluation. The system, in doing this, must ensure that information used in the prioritisation is also provided in an intuitive and interpretable manner. It must provide suitable values and visualisations to allow the expert to understand the recommendation in a transparent manner.

6. Testing the framework with an implementation

To evaluate the framework, a DSS was implemented with an A&R team of five people in the UK focusing on dance music. Based on the finding about digital scouting, this is central to A&R's DSS implementation in this evaluation. The system automates the identification of new talent by replicating and expanding on manual work done by A&R on social media and streaming platforms. The exact platform and process are not detailed, at the request of the company, but a total of four platforms are used to identify new talents $socialmedia_1$, $streaming_1$, $streaming_2$, and $streaming_3$. In this context, discovered contents are performances on one of the streaming platforms, which are the focus of the music label, rather than the performers. In $socialmedia_1$, which is a text and image social platform, talent is discovered through a search of keywords that the expert stated are used by the public to discuss musical talents. This is then matched with talent on streaming platforms so that the output from social media is in terms of performance on either $streaming_1$ or $streaming_2$. On both $streaming_1$ and $streaming_2$, talent is also discovered by monitoring for new talent being added to specific groupings that the expert is interested in.

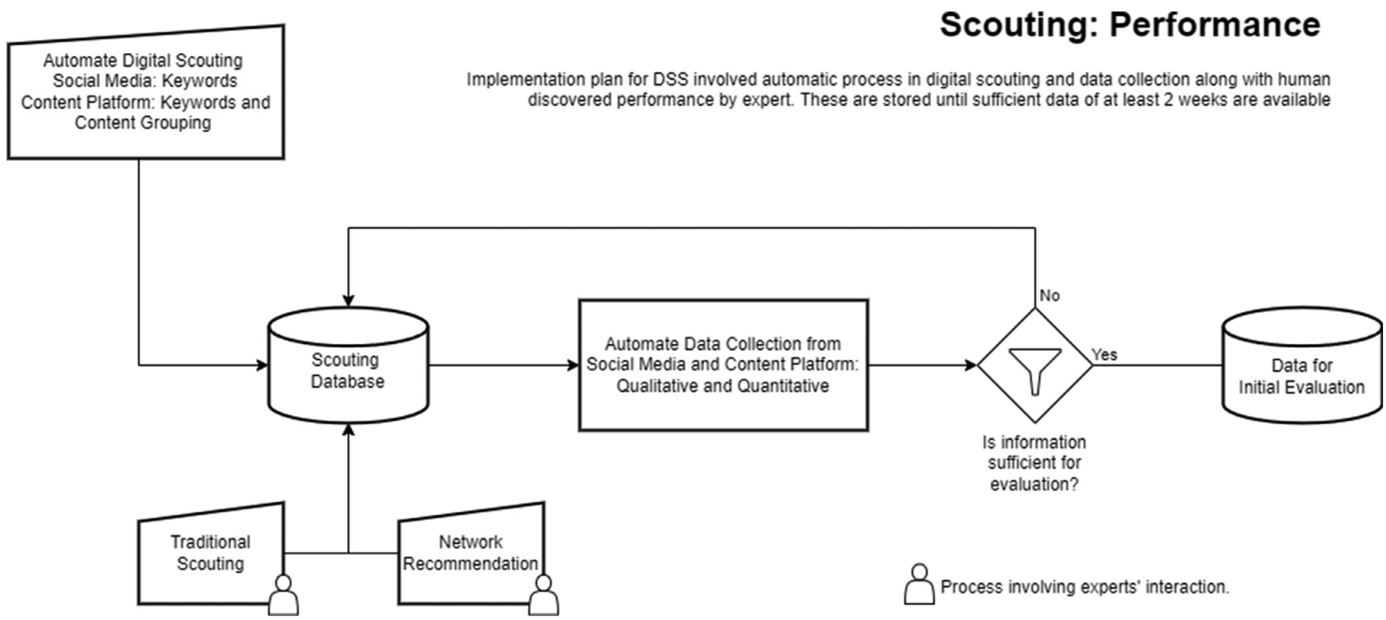

Fig. 3. Implementation of DSS for scouting of musical talents.

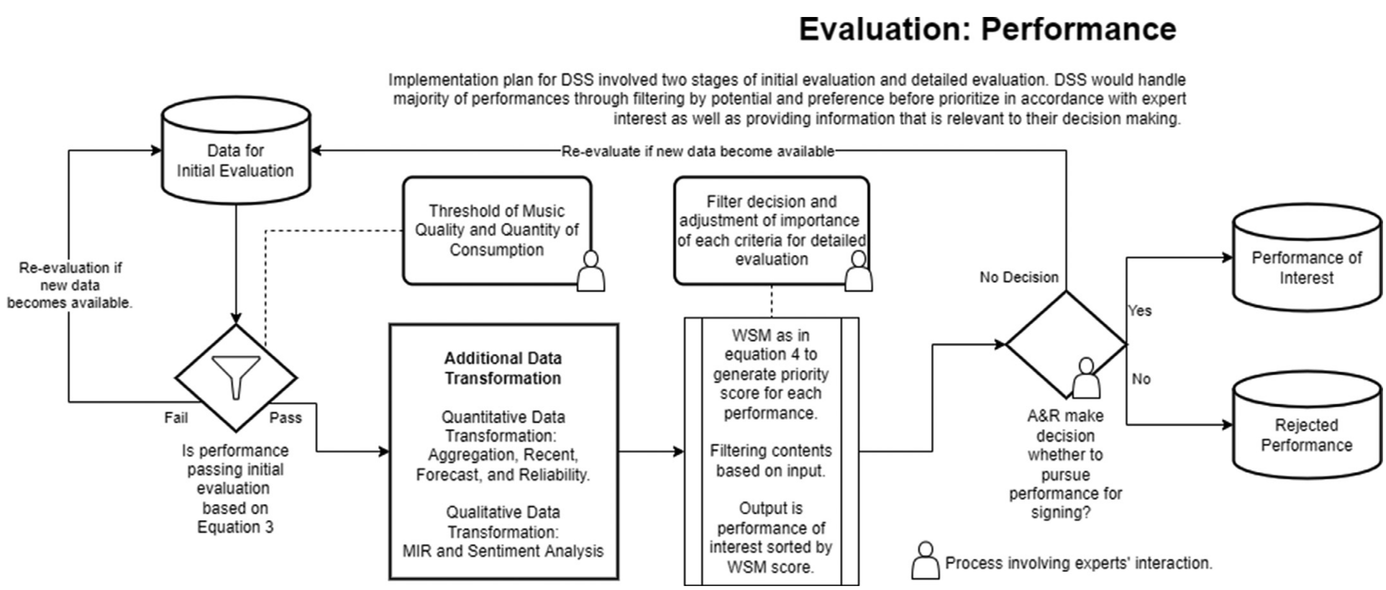

Fig. 4. Implementation of DSS for evaluation of musical talents.

Finally, $streaming_3$ is monitored for all new talent. An overview of this is shown in Fig. 3.

In the next stage of evaluation, performances with at least 14 days of historical data are initially evaluated through a decision-rule filter based on the quantity of audience and quality of performance. On all three streaming platforms, the quantity of audience is based on aggregate consumption of views and listening. Quality is based on the similarity of music quality to a group of performances in the same genre based on MIR. Rejected music is retained for re-evaluation as new information on the quantity of audience can change over time. Equation (3), which is based on Eq. (1) in the framework, describes this process based on aggregate consumption ($Consumption$) and quality of music ($Music_{Quality}$) compared to a threshold ($Threshold_X$). The overview of this stage can be seen on the left side of Fig. 4.

$$Pass_{IE} = \begin{cases} 1 & \text{if } Consumption > Threshold_{Cons} \text{ and } Music_{Quality} \\ & > Threshold_{Quality} \\ 0 & else \end{cases} \tag{3}$$

The second part of the detailed evaluation is based on quantitative and qualitative criteria that the expert uses in their decision as listed in the framework. The likelihood for the performance to become successful with the public ($success_q$) is based on the audience size, the qualitative response from the followers, and the quality of the musical performance itself. Each performance is evaluated using only data from the same platform and is normalised in relation to performance from the same genre and time frame on the same platform. For example, the aggregate audience number from $streaming_1$ is normalised against the highest aggregate audience number of a performance from $streaming_1$ of the same genre within the three previous months. It must be noted that not all criteria from the framework are used due to limitations of the implementation and the requirement that is set by the user for the trial.

Starting with audience size, each of the $streaming_1$, $streaming_2$, and $streaming_3$ platform has a similar measurement for online audience number. This audience is broken down into four variables of aggregate consumption (Agg), recent consumption (Rec), forecast of consumption in the future (Pro), and reliability of data (Rel). On the qualitative response from the followers (Fol), comment on performance on $streaming_1$ and $streaming_2$ are collected and natural language processing is used to extract audience sentiment on the performer and performance on a scale from negative to positive ranging from -1 to 1. If no comments are found or the platform does not have a comment function then the weight is adjusted to zero. Lastly, the quality of musical performance (Qua) is measured using beat, tempo, and mel-spectrogram, based on similarity to desired performances in each genre selected by the expert. These quantitative and qualitative variables are scaled between 0 to 1 and weighted according to individual experts' preferences. W is the weight of importance for the criteria as assigned by an expert, which must total 100 per cent across all criteria.

Finally, the variable $Priority$ is used to rank performances based on the likelihood of success ($Prob_{Success}$) and preference by the expert. This is based on Eq. (2) from before, with success adjusted by a multiplier based on decision maker preference for genre, tempo, and emotion ($Pref_G, Pref_T, Pref_E$). If the selection of the performance matches either of these three, an extra multiplier between 1 and 1.1 is used resulting in a final priority that is more aligned with their preference. On the contrary, the priority will immediately be set to 0 and filtered out (Fil_G, Fil_T, Fil_E) if the performance is on the list of characteristics not desired by the decision maker. The formula is seen in Eq. (4).

$$Prob_{Success} = \frac{W_{Agg}Agg + W_{Rec}Rec + W_{Pro}Pro}{Rel} + Fol * W_{Fol} + Qua * W_Q$$
$$Priority = Success_{Quantitative} * ((1 + Pref_G) * Fil_G) * ((1 + Pref_E) * Fil_E) * ((1 + Pref_T) * Fil_T) \quad (4)$$

The final output is a list of performances ranked according to expert criteria, filters, and preferences. The DSS supports experts in their subsequent consideration of the performances with information deemed to be crucial to consider, such as actual quantitative and qualitative variables used in prioritisation, information on the performer, samples of the performance audio, and artwork used for the performance.

7. Evaluation of the implementation of the decision support system

In evaluating the implementation, five experts were asked to utilise the DSS for three months between December 2019 and February 2020. The quantitative assessment compared manual evaluation that was done by experts without DSS support with evaluation by those with DSS support, based on three numerical indicators: the quantity of performances discovered, the quantity of performances analysed, and the quantity of performances found to be of interest to the decision maker. As seen in Fig. 5, in two of these criteria, the quantity of discovered and analysed performances, the DSS showed a clear increase in the quantity performances processed compared to previous human evaluation. Previously, the process averaged 1000 initial evaluations per day, with none being looked at during the weekend. In contrast, the DSS discovered around 16,000 performances per day, which were then filtered and prioritised to 580 per day for decision maker consideration. However, the third evaluation criterion of successful recommendations that resulted in the decision to pursue signing still averaged two musical performances per month, which is approximately equal to the current A&R process. Overall, we were able to show clear quantitative improvement in two areas of evaluation.

In qualitative evaluation, we want to ensure that the quantity increased corresponds to an improvement or parity in the evaluation by DSS. To ensure this, five experts were asked to provide qualitative feedback on their experience of using the DSS to find, evaluate, and sign new musical talents. They were asked to discuss the overall quality of content in an initial and detailed evaluation, particularly using specific examples from the three months of implementation in their explanation. They were also asked to contrast the quality of performance discovered, filtered, and recommended working with and without the DSS. Overall, we summarised the finding into three key areas of interest that were mentioned. These are listed in Table 10. We found that the quality was similar between the evaluation before and with DSS. This achieved our main goal of ensuring that A&R is able to reduce their responsibility on initial discovery and filtering and focused on detailed evaluation and preference choice.

8. Discussion

This section discusses the limitations and future research avenues. Starting with limitations, as the experts interviewed were solely working in the United Kingdom, that country's unique status as a major source and distributor of musical performers and performances may have an impact on the expert process (Cloonan, 2016), skewing the importance on criteria. For example, the ability to perform live may be less important. It is important to conduct similar work in other regions to validate the expert model. It is important, though, to note that we found a strong alignment between our work and other studies into A&R in Poland (Galuszka & Wyrzykowska, 2017) and the Netherlands (Zwaan & Ter Bogt, 2009).

In our framework, the initial evaluation is a decision rule. This can result in a border effect, with a sharp threshold (Slowinski et al., 2009) between those accepted and rejected. However, the initial evaluation is redone regularly based on new information on audience number and the border is updated based on the expert's input. This led us to consider this limitation to have a minimal impact since each performance has a chance to show its merit over time. In detailed evaluation, there are a number of criteria, discovered in the mental model, that are difficult to assess quickly at scale by both human experts and subsequently also by the DSS, such as uniqueness of performances and fitness with time (Kuroyanagi et al., 2019). These require in-depth study to measure and translate them into quantifiable variables for MCDA methodology. In this work, the research used MIR features and similarity between the performance of interest and benchmark performances (McFee et al., 2012), as an alternative to directly assigning scores to these variables.

There is also a question over which MCDA method should be used for detailed evaluations by A&R. There is a need for methods that are computationally fast, are accessible to users, and incorporate intuitive adjustments based on different priorities for different genres. An example of the need for flexible criteria is the difference between dance which put high importance on criteria of vocal quality, memorability, and tempo in contrast to slower tunes such as Jazz (Tzanetakis & Cook, 2002). This framework proposes utilising MCDA methods of the aggregation family to conduct the assessment of alternatives. It has a time complexity that is $O(n)$, with n as the number of alternatives, more interpretable for non-experts with an easy-to-understand output of higher as better alternatives and has relatively simple granularity for the im-

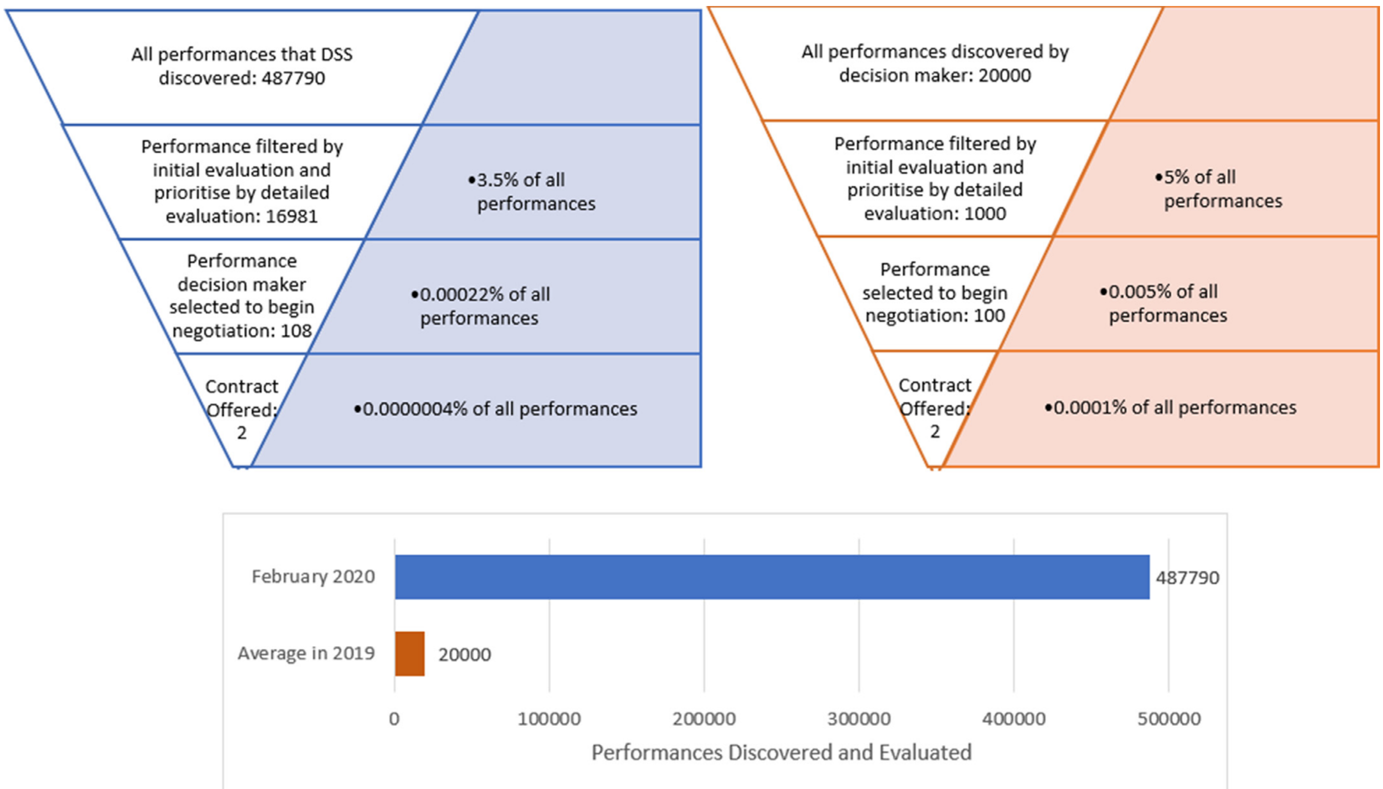

Fig. 5. Comparison between contents evaluated by DSS in February 2020 on the left in blue with approximate contents evaluated per month in 2019 as done solely by experts on the right in orange. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 10
Qualitative feedback from DSS's user.

Topic of Interest	Feedback from Expert
Quality of the content that was in the initial evaluation	Experts found no significant change in comparison to a manual review by a human expert. Experts reviewed performances based on the evaluation process and were largely in agreement with those that passed the initial evaluation.
Quality of the performance that was ranked in detailed evaluation	Experts felt the system was able to replicate what a human would be looking at and prioritise performances according, but that the qualitative information such as quality of performance and sentiment behind them needs to be more transparent and may need more iteration.
Integration of DSS into the workflow	Experts feel that there is a clear benefit from DSS in the expansion of new performances discovery coverage, removal of contents that are not of interest, improved access to useful information, improved comparisons between performance, a useful ranking that align with their interest, and intuitive visualisations.
Other comments	There were questions about whether qualitative assessments could truly represent judgments on music quality as done by humans and that more qualitative sources were not included. Also, the content recommended by the DSS was not always original, but more granular preference filters could solve this.

portance of each criterion. However, several outranking methods are arguably better in handling fuzziness in the decision-making process (Vaidya & Kumar, 2006). For example, the ELECTRE family (Figueira et al., 2016) offers the preference threshold, indifference threshold, and veto threshold to better capture the contextual information for a given decision problem. However, these methods require more input and cognitive effort from the A&R expert than the aggregation methods. Therefore, the selection of the MCDM method turns out to be a trade-off between reducing the complexities in tool design against the usability of tool operation. This is an area of investigation that could result in significant improvement in the future. Finally, expanding future work into different weight elicitation procedures (Riabacke et al., 2012) would be beneficial, particularly the ranking criteria with the SIMOS method (Figueira & Roy, 2002). The elicitation of criteria weight was a challenge as seen by the request for simplification by users.

Looking at the evaluation of the implementation, there is an equally low number of performances worthy of the contract for both the DSS and manual A&R which presents a limitation on the validation of DSS effectiveness. We believe that this is due to the nature of A&R in a small firm in which a signing is a significant investment. Further research focusing on the difference in the signing would need to be longer or on a larger label. Nevertheless, the results provide a good indicator that the task can be replicated and the integration of DSS into the decision-making process can be done. It also reduces the A&R's concern over the fear of missing out by enhancing the search coverage.

9. Conclusion

This research proposed and evaluated a framework for a DSS for the signing of new musical talent. It is based on experts' mental model discovered through open-ended structured interviews showing two stages of A&R process: discovery and evaluation. The discovery stage is scouting for new musical talent from three different sources: traditional scouting, such as attending live performances at various venues; network recommendations, such as suggestions from peers and colleagues in relevant industries; and digital scouting through social media and streaming platforms. Experts also described growing reliance on networks and digital sources compared to traditional sources. The evaluation process is split into two stages, starting with an initial evaluation that aims to limit the work in the next stages by filtering out content that is not appropriate in terms of quality or has limited potential to be successful. Talent must achieve a minimum quantity of interest from the public and possess sufficient quality (with the definition of quality varying among experts). Talent which has passed the initial stage is then evaluated in detail to assess their potential to be signed with criteria used for both performer and performance described in detail in Tables 8 and 9. The framework also involves extracting and utilising quantitative criteria and qualitative criteria for initial and detailed evaluation, the use of rule-based methods for initial evaluation, and the adoption of MCDA methods such as WSM for detailed evaluations by the DSS. The framework can be easily adapted for implementation in music labels of different sizes and focuses. Finally, the researchers tested the framework by implementing the DSS for a small dance music label. The results showed a major benefit of expanding the quantity of talents discovered and evaluated over three months, while maintaining an evaluation process of a similar quality as manual evaluation, according to five experts in the trial.

Acknowledgement

This research is supported by the Knowledge Transfer Partnership between the University of Leeds and All Around the World Productions (AATW). The authors would like to thank AATW for their support.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2023.06.014.

References

Abel, N., Ross, H., & Walker, P. (1998). Mental models in rangeland research, communication and management. *The Rangeland Journal, 20*(1), 77–91.
Arampatzis, G., Kiranoudis, C. T., Scaloubacas, P., & Assimacopoulos, D. (2004). A GIS-based decision support system for planning urban transportation policies. *European Journal of Operational Research, 152*(2), 465–475.
Bai, C., & Sarkis, J. (2010). Green supplier development: Analytical evaluation using rough set theory. *Journal of Cleaner Production, 18*(12), 1200–1210.
Bana e Costa, C. A., & Vansnick, J.-C. (1994). Macbeth - an interactive path towards the construction of cardinal value functions. *International Transactions in Operational Research, 1*(4), 489–500.
Bender, M., Gal-Or, E., & Geylani, T. (2021). Attracting artists to music streaming platforms. *European Journal of Operational Research, 290*(3), 1083–1097.
Bravo-Lillo, C., Cranor, L. F., Downs, J., & Komanduri, S. (2010). Bridging the gap in computer security warnings: A mental model approach. *IEEE Security and Privacy, 9*(2), 18–26.
Carroll, J. M., & Olson, J. R. (1988). Mental models in human-computer interaction. In *Handbook of human-computer interaction* (pp. 45–65).
Chermack, T. J. (2003). Mental models in decision making and implications for human resource development. *Advances in Developing Human Resources, 5*(4), 408–422.
Cloonan, M. (2016). *Popular music and the state in the UK: Culture, trade or industry?*. Routledge.
Cooke, C. (2020). Uk record industry figures grew 7.3% in 2019, says bpi. https://completemusicupdate.com/article/uk-record-industry-figures-grew-7-3-in-2019-says-bpi.
De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research, 269*(2), 760–772.
Dewan, S., & Ramaprasad, J. (2014). Social media, traditional media, and music sales. *MIS Quarterly, 38*(1), 101–122.
Dhanaraj, R., & Logan, B. (2005). Automatic prediction of hit songs. In *Ismir* (pp. 488–491). Citeseer.
Figueira, J., & Roy, B. (2002). Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European Journal of Operational Research, 139*(2), 317–326.
Figueira, J. R., Mousseau, V., & Roy, B. (2016). ELECTRE methods. In *Multiple criteria decision analysis* (pp. 155–185). Springer.
Galuszka, P., & Wyrzykowska, K. M. (2017). Artist and repertoire goes online: Evidence from poland. *Media Industries Journal, 4*, 1.
Haastrup, P., Maniezzo, V., Mattarelli, M., Rinaldi, F. M., Mendes, I., & Paruccini, M. (1998). A decision support system for urban waste management. *European Journal of Operational Research, 109*(2), 330–341.
Herremans, D., Martens, D., & Sörensen, K. (2014). Dance hit song prediction. *Journal of New Music Research, 43*(3), 291–302.
IFPI (2021). Industry data - engagement with streaming drives growth of the global music market. https://www.ifpi.org/our-industry/industry-data/.
Janssen, M. S. S. E., & Verboord, M. N. M. (2015). Cultural mediators and gatekeepers. In *International encyclopedia of the social sciences and behavioral sciences* (pp. 440–446).
Jones, N., Ross, H., Lynam, T., Perez, P., & Leitch, A. (2011). Mental models: An interdisciplinary synthesis of theory and methods. *Ecology and Society, 16*, 1.
Jungermann, H., Schütz, H., & Thüring, M. (1988). Mental models in risk assessment: Informing people about drugs. *Risk Analysis, 8*(1), 147–155.
Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
Kolkman, M. J., Kok, M., & Van der Veen, A. (2005). Mental model mapping as a new tool to analyse the use of information in decision-making in integrated water management. *Physics and Chemistry of the Earth, Parts A/B/C, 30*(4–5), 317–332.
Kuroyanagi, J., Sato, S., Ho, M.-J., Chiba, G., Six, J., Pfordresher, P., … Savage, P. (2019). Automatic comparison of human music, speech, and bird song suggests uniqueness of human scales. In *9th folk music analysis conference* (pp. 35–40).
Leonard, K. J. (1995). The development of a rule based expert system model for fraud alert in consumer credit. *European Journal of Operational Research, 80*(2), 350–356.
McFee, B., Barrington, L., & Lanckriet, G. (2012). Learning content similarity for music recommendation. *IEEE Transactions on Audio, Speech, and Language Processing, 20*(8), 2207–2218.
Morgan, M. G., Fischhoff, B., Bostrom, A., & Atman, C. J. (2002). *Risk communication: A mental models approach*. Cambridge University Press.
Morris, J. W., & Powers, D. (2015). Control, curation and musical experience in streaming music services. *Creative Industries Journal, 8*(2), 106–122.
Negus, K. (2011). *Producing pop: Culture and conflict in the popular music industry*. Edward Arnold.
Otto-Banaszak, I., Matczak, P., Wesseler, J., & Wechsung, F. (2011). Different perceptions of adaptation to climate change: A mental model approach applied to the evidence from expert interviews. *Regional Environmental Change, 11*(2), 217–228.
Quintero, A., Konaré, D., & Pierre, S. (2005). Prototyping an intelligent decision support system for improving urban infrastructures management. *European Journal of Operational Research, 162*(3), 654–672.
Riabacke, M., Danielson, M., & Ekenberg, L. (2012). State-of-the-art prescriptive criteria weight elicitation. *Advances in Decision Sciences, 2012*, 24 276584. https://doi.org/10.1155/2012/276584.
Slowinski, R., Greco, S., & Matarazzo, B. (2009). Rough sets in decision making. In *Encyclopedia of complexity and systems science* (pp. 7753–7787). Springer.
Statista (2021). Hours of video uploaded to YouTube every minute as of May 2019. https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/.
Steininger, D. M., & Gatzemeier, S. (2019). Digitally forecasting new music product success via active crowdsourcing. *Technological Forecasting and Social Change, 146*, 167–180.
Strobl, E. A., & Tucker, C. (2000). The dynamics of chart success in the UK pre-recorded popular music industry. *Journal of Cultural Economics, 24*(2), 113–134.
Tzanetakis, G., & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing, 10*(5), 293–302.
Vaidya, O. S., & Kumar, S. (2006). Analytic hierarchy process: An overview of applications. *European Journal of Operational Research, 169*(1), 1–29.
Velasquez, M., & Hester, P. T. (2013). An analysis of multi-criteria decision making methods. *International Journal of Operations Research, 10*(2), 56–66.
Yang, L.-C., Chou, S.-Y., Liu, J.-Y., Yang, Y.-H., & Chen, Y.-A. (2017). Revisiting the problem of audio-based hit song prediction using convolutional neural networks. In *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 621–625). IEEE.
Zhang, R., Lu, J., & Zhang, G. (2011). A knowledge-based multi-role decision support system for ore blending cost optimization of blast furnaces. *European Journal of Operational Research, 215*(1), 194–203.
Zwaan, K., & Ter Bogt, T. F. (2009). Research note: Breaking into the popular record industry: An insider's view on the career entry of pop musicians. *European Journal of Communication, 24*(1), 89–101.