

# NAUTILUS: A CASE STUDY IN HOW A DIGITAL SCORE CAN TRANSFORM CREATIVITY

Craig Vear, Carla Rees and Adam Stephenson

---

**Abstract:** This article discusses *Nautilus* (2022), a composition for solo bass flute created using machine-learning techniques and a Unity game engine. We consider the approaches we adopted and how they enhanced creativity and musicianship for those involved. We reflect on Unity's potential as a novel and flexible driver for the creation of a musical score in which traditional elements of compositional design are presented to a performer as a co-creator for interpretation and communication inside the act of musicking. Through this we offer insights into performer agency and how a performer decodes media, sound, images and AI through their instrument, personal skills and musical aesthetic. We describe how the notion of a music score was re-conceptualised, transforming our understanding of the activities of composition, collaboration and performance.

---

## Introduction

This article discusses a practice-based case study undertaken as part of a European Research Council (ERC)-funded project, The Digital Score: Technological Transformations of the Music Score (DigiScore).<sup>1</sup> The core aims of this project are to determine scientific knowledge of how digital scores stimulate new creative opportunities and experiences within a range of music practices, to develop a theoretical framework for digital scores as an important transdisciplinary area of research and to build a scientific study of inclusive digital musicianship through the transformative potential of the digital score. A series of practice-based case studies places experts at the centre of their practice, offering them meaningful experiences about which they can report back to the researchers of DigiScore who will in turn synthesise the results within a developing framework.

A key part of the framework that supports DigiScore is to understand meaning-making in music from inside the creative act. We have adopted Christopher Small's notion of *musicking*,<sup>2</sup> that 'to music is to take part... in any capacity, in a musical performance, whether by performing, by listening, by rehearsing or practicing, by

<sup>1</sup> Details are available online at <https://cordis.europa.eu/project/id/101002086> (accessed 25 May 2022).

<sup>2</sup> Christopher Small, *Musicking* (Middleton: Wesleyan University Press, 1989).

providing material for performance (what we call composing)'.<sup>3</sup> Small stresses that 'the act of musicking establishes in the place where it is happening a set of relationships, and it is in those relationships that the meaning of the act lies'.<sup>4</sup> Simon Emmerson clarified Small's use of the term 'meaning' to 'what you mean to me',<sup>5</sup> a subtle shift that circumvents the significant issues of value and of who is doing the evaluation of meaning. Meaning, or 'what you mean to me', is to be found in the relationships formed between the new creative acts of musicking and the technologies and media of the digital score.

### The research process

*Nautilus* (2019–22) was the first DigiScore case study. The music was inspired by the idea of a deep-sea journey, as a nautilus mollusc navigates deep ocean trenches, and *Nautilus* describes this journey, the bass flute and generative sound-design highlighting the topography of the oceans and vast openness of the depths. The work was initiated as a collaboration between the composer Craig Vear and low-flutes expert Carla Rees in 2019 but was delayed by the COVID-19 pandemic and then resumed as the beginning of the DigiScore project. The project involved three practitioners: composer, digital score researcher and project principal investigator Craig Vear, bass flute player Carla Rees and Unity programmer Adam Stephenson. Each brought their own experience, perspective and creative practice to the work, driving both the aesthetic and practical considerations of the work, but with the shared goal of creating a digital score that supported and enhanced Carla's sense of musicking to a point where it felt as if it was operating *with her* in the making of the music.

Five research questions were identified:

1. How can a games engine such as Unity be integrated into a digital score without detracting from the flow of musicking experience?
2. How can a neural network be trained with an aesthetic design to generate a series of digital scores?
3. How do we publish such a digital score so that others may engage with this composition?
4. How can narrative structures from game design be used to enhance the experience of a musician with a digital score?
5. How can this approach develop or enhance (or restrict) performer agency in the interpretation of the score, and what are the challenges that arise from this?

To address these questions, the research was conducted in three phases:

#### AI behavioural design and development

*Nautilus* uses a Unity game engine as the main platform for the visual elements of the digital score, which includes sea-bed imagery and 'sinking notation' (musical notation written on staves tied to anchors sinking to the bottom of the sea; see [Figure 1](#)). The Unity engine also listens to Carla and makes judgements about what and when to

<sup>3</sup> *Ibid.*, p. 9.

<sup>4</sup> *Ibid.* p. 13.

<sup>5</sup> Simon Emmerson, *Living Electronic Music* (London: Routledge, 2007), p. 29.

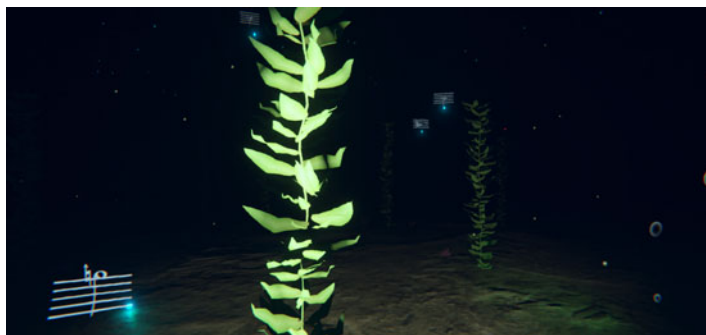


Figure 1:  
Screenshot of the visual part of the digital score for *Nautilus*.

generate a sound design. The behaviour of the AI and generative processes were designed and tweaked so that they felt part of the composition, not an external process. The neural networks were trained and tuned within a defined compositional aesthetic so that Carla was able to feel the presence of their behaviour and evaluate their effect upon her musicianship. Relatively short examples were used that did not progress compositionally but allowed the creative team to develop the dynamic behaviours of the technology.

The compositional process started with an improvisation by Carla on the idea of the nautilus's journey. This improvisation then became the source material for AI processes and sound-design manipulation that are heard during the performance. Machine-learning processes and a neural net make in-the-flow decisions about how the music is to be shaped. During the development stage of the project a neural network<sup>6</sup> was trained to predict and output a type of information based on input data. For *Nautilus* the neural network was trained on transcribed jazz improvisations: given a note from Carla's original improvisation as an input the network calculated what the next note would be as if it was a jazz improviser, suggesting new note possibilities that expanded the scope of her original improvisation.

#### Compositional development

Once some of the parameters of the aesthetics and behaviours of the piece were understood the composition was allowed to develop. The machine-learning process was the basis of a development environment that was eventually fixed and migrated into Unity behaviour. At the start of each iteration of the development phase, random notes from Carla's original improvisation were passed through the neural network and it in turn output a notated improvisation based on the input note choices. This notation formed part of the digital score for live interpretation and was designed to offer Carla both a sense of familiarity, from her original improvisation, and suggestions of other materials predicted by the neural network, so that she felt engaged in a co-creative musicking between AI and human musician. The team came to understand these elements to such an extent that

<sup>6</sup> A neural network is a complex statistical analyser used to output useful information from a given set of input data – for example, to predict the house prices of a particular suburb based on previous sale, size, demographic, socio-economic and location data. The training code for *Nautilus* used this repository, with some minor changes: [https://github.com/haryoa/note\\_music\\_generator/blob/master/Music%20Generator.ipynb](https://github.com/haryoa/note_music_generator/blob/master/Music%20Generator.ipynb) (accessed 28 November 2022).

the relationships between them became the working materials of the composition.

Another element of the digital score is generative sound design, taking the audio recording of the original improvisation as its source material and responding to the live sound as a stimulus. It also manipulates the playback speed, again offering Carla a sense of familiarity and suggestion as well as in-music relationships and a sense of musicking involvement, a way of drawing her into the complex programming of the AI, which might otherwise feel like a black-box devoid of musicking soul.

The final version of the digital score migrated all the learnt behaviour of the AI and neural networks into a single Unity environment that creates an immersive world for the audience and musician to inhabit through the piece. This Unity engine version still generates random notation from a fixed library, developed using the development processes; it also listens to Carla and generates a backing track using procedural algorithms designed to mimic the behaviour of the developmental AI. As the Unity engine listens to the performer and its own backing track it moves the camera through the ocean using the amplitude of each source: left for live sound, right for generated sound. The aim throughout the developmental phase was that Carla's regular engagement with and incorporation of the AI into her sense of musicking would become part of the way she responded to the computational elements as a cooperative other.

It is important to emphasise that the digital score consists of all these elements: Unity design, 'sinking notation', generative backing track, behaviour of the Unity AI, presences of thinking processes and audio files. To consider the score as only consisting of notation ignores the serious influence of all the other elements on meaning-making, on the ideas communicated through the digital score and the ways in which musicking is shaped by these elements. This, however, presents a significant challenge to existing notions of musicianship: how to create a musicking experience that binds these elements together, rather than presenting something that is created from individual elements like a Frankenstein score. The least successful version of this mixed-media approach would have been to construct an experience for the performer that feels like 'a bit of this', stuck together with 'a bit of that'.

Such integration is very difficult to achieve as the creative musicians (coder, composer, performer) must work together to build something that has a unified aesthetic and a singularity of message,<sup>7</sup> draws together and enhances the communicative value of a digital score. This required us to embrace a transdisciplinary approach where we sought to find new, common principles and factors that contribute to a wholeness of all our musicking experiences, and this will normally go beyond/distort/transform/enhance/transcend our own training and ways of thinking.

#### Performance in a real-world environment

The final test of *Nautilus* was to present it to a critical audience and to gather qualitative data from the creative team. This was designed using a two-way perspective of *encoding–decoding* through which the

<sup>7</sup> This is discussed in more detail at <https://digiscore.dmu.ac.uk/2022/01/17/the-digital-score-through-the-medium-and-its-message/> (accessed 25 May 2022).

experiences of those encoding communicative elements into a digital score (composer, coder, designer) were compared with the experiences of those realising the digital score (performer, audience). Furthermore, the reflections of the composer, coder and designer were captured and included in this dataset, as was the legacy of the experience for the performer, through questionnaires many weeks after the event. Carla's performance of *Nautilus* can be heard at [www.youtube.com/watch?v=XK-9eXCJxCg](https://www.youtube.com/watch?v=XK-9eXCJxCg) (accessed 22 November 2022).

### Some reflections

We reflected on the meaning-making through musicking that occurred during the research process, each of us separately completing an online form to create a qualitative account of the legacy of working on the project.

#### The performer's perspective (Carla Rees)

The Unity immersive score gives the performer considerable agency in the interpretation of meaning through musicking. In a conventional score, pitches and rhythms are notated in detail and space is left open for the nuances of interpretation, taking into account relevant performance practices. *Nautilus*, as well as communicating these nuances, offers scope for an individual performance practice centred upon digital musicianship and musicking creativity to develop. Visual objects propose relationships which are translated into sound, with decisions made according to personal aesthetics, visual awareness and the performance experience, because the material played changes the direction of travel through the virtual space through interaction with the audio file. The performer must notice and interpret a range of visual and aural cues in order to create a meaningful musical experience.

Choices can be made about which of the visual elements to play (since there are too many to play them all), as well as the duration of each sonic event. For example, a pitched note can last for as long as it can be seen on screen, but it may be musically more appropriate to move one's attention to another visual object long before it leaves the field of view. Everything seen and heard within the score provides potential for sonic material, and it is inevitable that different performers (and on different instruments) would decode these according to their own musicianship, technical skills, instrumental resources and musical aesthetics.

The game environment, combining visual and audio elements, however, provides information relating to the mood, atmosphere and general ambience which define the overall character of the piece. The changing scenes can be interpreted in different ways, but the pace and energy of each scene define the overarching structure of the composition and the underlying narrative. The piece can therefore be interpreted and decoded individually by each performer, with a certain level of creative freedom, while the piece itself maintains its overall identity. This is a very different way of interpreting and working with a score, and one that can shift the musicking experience for composers, performers and audiences.

### The composer/creative director's perspective (Craig Vear)

I felt that the choice of materials and behavioural presences was well judged for this composition. The iterative process helped immensely in identifying the value of each of them, and the open exploration process enabled radical thinking and novel experimentation. Although it would have been useful to explore the machine-learning for longer, we arrived at a point where what we had was working, so further investigation was not necessary. However, from a research perspective this core research question remains unanswered.

We created a fully determined but flexible 14–16-minute-long piece with a linear sequence of form. Subsequently we adapted this to a single-movement six-minute version. This works well and guides (perhaps even collaborates with) the performer in the construction of the music. We could have developed the code further to listen for the performer's cues at transition points such as 'wait for long held low', but the migration to a Unity-only system didn't easily support this option, especially within the time constraints of the project. An open question remains about the open-world potentials of this piece, by removing the fixed form, and entering a more emergent open-world format and how that might change the nature of the composition. Another open question is about remote multi-player involvement in a digital score, like the experience of gamers in *Fortnite*. Bringing this together as a new case study with the same team is an exciting proposition.

I felt my role become more that of creative director than composer. The term 'composer' was too limited and did not embrace the different roles and conversations through which I drove forward the vision of this project and accommodated the team's considerable input. I felt like an auteur who was also a parent, welcoming ownership of the project from the team members, evaluating and incorporating the team's ideas. The enhancement of the performer's role offered by this approach was especially rewarding; I got the impression that so much more was being communicated through the relationships and presences that were 'alive' inside the musicking experience. This presented many new ways of communicating with performers, expanding the types of musical ideas that can be contained within a score paradigm.

### The Unity developer's perspective (Adam Stephenson)

The interaction design in *Nautilus* was intended to give the performer a feeling of influence on the environment's behaviours but not one of exact control. The interaction system is inherently ambiguous, as is the backing-track generation: it can decide to take inputs and react to the performer or ignore them and make its own decisions. These design choices were intended to avoid the performer feeling that they were in a game and could predictably control the outcome with certain clear behaviours. Instead, the performer may begin to recognise patterns and learn to adapt to the score, living alongside it and either playing with or against it. There is no hierarchy of control between the performer and the score; instead, the two coexist in the same space.

The presentation of the music notes in the scene as objects, influenced by physics and instantiated with random velocities, meant that sometimes the performer was unable to read and interpret a note in time. Notes would also sometimes spawn too far away, so that they could not be read. This might frustrate the performer but it feeds

into the idea that this is a living world that makes no efforts to accommodate them. The performer was also encouraged to make decisions and only play notes when and how it felt right for them in their journey through the environment. If I were to make any changes, I would increase the influence the performance had on the environment.

For me it was a new challenge to develop software for a specific user who interacts with the work in a way that I cannot. I could not test the effect a live flute performance would have but could only emulate the performer's input with taps on the microphone and simple vocalisations. This led me to develop systems that could be easily, quickly and deeply customised, maximising the opportunity for live feedback and implementing changes during the few test sessions with the performer. The project taught me a lot about creating an effective project architecture in which incremental, experimental development can be achieved more quickly.

The experience of working with experts in fields entirely different to mine has given me a greater perspective on collaboration. Each test session with the performer teased out so much more potential because they had a very different perspective on the score. My perspective was more technical, and after seeing the visual presentation hundreds of times during development I was accustomed to it; but when the performer tested it, I was mesmerised and forgot about the hours spent tweaking code.

## Conclusions

In this section we return to the research questions introduced earlier.

1. How can a games engine such as Unity be integrated into a digital score without detracting from the flow of musicking experience?

This case study suggests that Unity is an excellent environment for encoding and decoding musical ideas. The environment (visuals and sound design) provided a sense of mood/atmosphere that could be interpreted by both the individual performer and the audience. The environment is immersive and, if carefully managed, should not get in the way of the musicking experience. By that we mean that we felt it should remain musically focused and avoid turning the experience into a game, thereby shifting the role and position of the performer. But how much information can be incorporated in conventional notation and how much can be left free for improvisation? If one considers the environment to be similar to performance directions, offering tempo, mood, expression marks, etc., then there is potential for pitch/rhythm information to be incorporated at many different levels of detail. For example, we used flashing lights to indicate rhythmic information and the speed and density of the 'sinking notation' to invite interpretations about melodic line construction and gestures (see [Figure 2](#)). The format might also increase accessibility, as it could be formed in a way that does not require an ability to read complex musical notation, making the music approachable for performers from different musical cultures/backgrounds in which European notation systems are unfamiliar.

2. How can a neural network be trained with an aesthetic design to generate a series of digital scores?

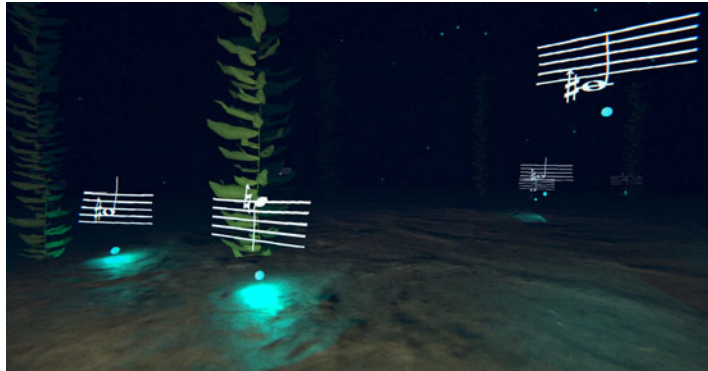


Figure 2:  
Detail of flashing lights anchoring  
the notation to the seabed.

In *Nautilus* neural networks were used as a rapid-prototyping tool. Their role was to generate materials and provoke responses that aided in the development of the team's understanding of the potentials of this piece. In a sense, they helped us find the boundaries of the composition and aesthetic and test out behaviours from inside musicking. The use of transcribed jazz improvisations as training data was problematic, however, because jazz transcriptions are not part of Carla's improvisation aesthetic. She felt that the jazz dataset was too limited harmonically and that she was stuck in a language that was neither her own nor designed specifically for the bass flute. For her, it highlighted the need for the musical material and the visual environment to match in terms of potential for exploration. It would be interesting to use multiple versions of her own improvisation to create the dataset, producing material that was more idiomatic for the Kingma System bass flute.

3. How do we publish such a digital score so that others may engage with this composition?

We reached a collective decision that this piece needs to be published in a format that is relatively inexpensive to produce and distribute and easy for performers to set up without needing access to specialist software. Unity provided an interesting solution that allows app exporting in both Windows and Mac formats as well as through WebGL, so it can be hosted online and accessible through a webpage, key factors in the decision to unify all the processing into the Unity engine, rather than having separate systems for audio production, music-notation generation and visual design.<sup>8</sup>

4. How can narrative structures from game design be used to enhance the experience of a musician with a digital score?

Our approach allows for quick communication of ideas through different worlds/structures/universes in which one can break away from conventions such as gravity or place the player in an upside-down world (see Figure 3). This promotes creativity: the performer can choose how to interpret the various elements seen on screen, perhaps also breaking free of the musical restraints of a particular aesthetic

<sup>8</sup> Mac and Windows versions of the digital score can be downloaded at <https://digiscore.dmu.ac.uk/2022/01/27/nautilus/> (accessed 22 November 2022).



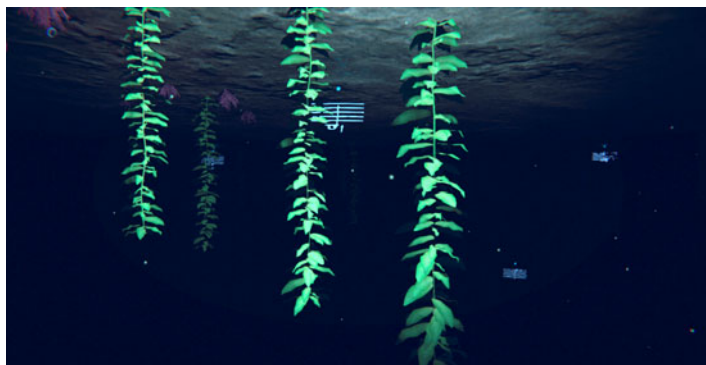


Figure 3:  
Detail of the *Nautilus* digital score  
showing the 'upside-down world'  
of section 3.

approach. We found that the visuals could create emotions which were then reflected in the music. The absence or reduction of notated material invites the performer to develop memory skills, such as remembering that particular symbols are performed in a particular way, to provide some structural coherence, although this is not significantly different from the requirements of improvising or playing graphic or text scores.

5. How can this approach develop or enhance (or restrict) performer agency in the interpretation of the score, and what are the challenges that arise from this?

We found that the performer can assign meaning to different objects, most of which are not traditional or typical music notations. Other players might have a similar approach to the instrument and yet still have opportunities to create an individual performance because the materials within the score afford different modes of musical meaning. Performances might sound quite different from one another but maintain a sense of aesthetic identity with the piece. For example, following Carla's performance, the singer Franziska Baumann expressed an interest in realising *Nautilus*. Her performance can be heard here [www.youtube.com/watch?v=SV6TqzJkiX4](http://www.youtube.com/watch?v=SV6TqzJkiX4) (accessed 22 November 2022); her interpretation varies from Carla's but the core aesthetic of the composition is retained.

The experience of working on *Nautilus* demonstrates that digital scores can allow performers to co-create material and bring their musical personality fully into the performance process. Traditional musicianship skills, such as listening, responding and communicating ideas through one's instrument, come to the fore, enabling the performer to be in the moment, experiencing a state of flow through the immersive nature of the materials. It is important, however, that the music produced is in keeping with the ambience and mood created by the visual materials and soundtrack; without this the inherent logic the narrative would be lost and communication with an audience might break down.

Through this case study we have begun to explore Unity's potential as a novel and flexible driver for the creation of a musical score. Traditional elements of compositional design, such as structure, narrative, mood and atmosphere, are presented to a performer for interpretation and communication. They engage with the materials using musical, technical and interpretative skills and have agency in the development of their decoding of the visual cues, depending on

their instrument, personal skills and musical aesthetic. An accompanying audio soundtrack also contributes to the musical direction of the work, enabling the performer to interact with sonic as well as visual cues to help them develop their individual approach to the performance. This project has provided us with a springboard for potential future development, opening up the possibility of interactivity with other performers.