

96 POSTCARDS IN REAL COLOR

David Kim-Boyle

The University of Sydney
Sydney, Australia
david.kim-boyle@sydney.edu.au

ABSTRACT

This paper describes the development of *96 Postcards in Real Color* (2022), a virtual reality (VR) work for up to eight singers which features a three-dimensional immersive score generated from image captions scraped from Instagram. The poetic inspiration is briefly presented and various technical elements of the work’s design, development, and implementation are discussed including how the *Selenium* and *Beautiful Soup* Python libraries were used to scrape and parse images and text from Instagram, and how the multiplayer framework of the work is supported with the *Photon Unity Networking* SDK. Various user experience (UX) considerations influencing the work’s design are discussed, together with a discussion of future research directions.

1. POETICS – BACKGROUND

In 1978, French writer Georges Perec composed his playful *Deux cent quarante-trois cartes postales en couleurs véritables* [1], a series of postcard texts generated with a set of simple combinatorial rules [2]. Each text describes a location, either a city, region, or hotel at which various activities and entertainments occur, before signing off with a farewell, see Figure 1.

We’re camping near Wood’s Hole. Sunning ourselves.
Lobster at every meal. I’ve caught a salmon. Many regards.

A big hello from Biarritz. So nice letting yourself go
brown in the sun. I’ve done a bit of sailing. Love.

We’ve finally landed in Nice. Lots of lazing about and
sleep. Really nice (despite the sunburn). Love.

Figure 1. Sample postcard texts from Perec’s *Deux Cent Quarante-Trois Cartes Postales en Couleurs Véritables*, trans. by John Sturrock [3].

The combinatorial writing techniques of Perec and the Oulipo (*Ouvroir de littérature potentielle*) school [4, 5] of whom he is an exemplar member, have produced several widely cited works including Perec’s 1969 novel *La Disparition*, which omits the letter ‘e’, and Raymond Queaneau’s *Cent mille milliards de poèmes* (1961), which is derived from a set of ten sonnets with interchangeable lines. Contemporary readers would perhaps be most familiar with the Oulipo through the work of its current President Hervé Le Tellier whose recent *L’Anomalie* (2020) was one of the bestselling novels of all time in France. Oulipo’s impact outside literature has also been felt through the loose assembly of schools in other disciplines each designated with its own title, e.g. Oumupo (*Ouvroir de musique potentielle*), Ouphopo (*Ouvroir de photographies potentielles*), and Oupeinpo (*Ouvroir de Peinture Potentielle*) [6]. More broadly of course, constraint-based creative practice has a long tradition in a diverse range of practices from film [7, 8] and architecture [9], through generative art and music [10].

The constraint-based approaches of the various Oulipo schools have been a particular inspiration to the author with their application of deceptively simple operative constraints offering unexpected forms of creative expression. In *96 Postcards in Real Color*, henceforth *96 Postcards*, for up to eight singers, I sought to adopt a similar constraint-based approach to musical organization, although contemporizing the medium to explore ways in which Instagram can provide source material for a three-dimensional (3D) immersive performance score presented to singers in VR. In a manner reminiscent of Perec’s combinatorial processes, in *96 Postcards* text captions are scraped from 96 unique locations, as many as possible of which are chosen from locations listed in Perec’s original text. Three locations are scraped for each letter of the alphabet to provide the first 78 data points, e.g. Ajaccio, Antibes, Alhambra, Balearics, Biarritz, Berghof, Cannes, Cyprus, Calabria, etc., while the remaining 18 locations are randomly chosen from this set. The text captions that accompany each of these 96 posts are analyzed with various natural language processing (NLP) techniques and used to define the musical structure of the work.

2. INSTAGRAM SCRAPING

Through their respective Application Programming Interfaces (APIs), data scientists have been able to explore how social media such as Instagram, Twitter, and Facebook can offer unique insights into social, economic, and cultural

trends. Indeed, a veritable cottage industry of data visualization has emerged over the past decade which seeks to make the vast swathes of data and relationships they present, aesthetically pleasing yet informative and functionally useful. In my recent creative work I have sought to explore how the contents of this data and the way it is propagated and transformed through the social metaverse can provide innovative modes of musical organization, especially in extended reality projects such as *The Twittering Machine* (2021), for HoloLens 2 and prepared piano [11], and more recently in the present work *96 Postcards*.

In 2020, Instagram deprecated its public API, replacing it with an Instagram Graph API which can only be used to gather data and provide analysis on a user's Instagram Business account. Unlike public APIs developed by platforms such as Twitter, this severely inhibited the ability to perform any social media analysis through the platform. This restriction was exacerbated by an additional update to Instagram's terms-of-service that expressly prohibited the scraping of data or images through any automated service. Despite these constraints and the lack of a public API, workaround techniques have been developed for those conducting social media analysis although such techniques come with the inherent risk of having accounts banned and even exposing a user to potential litigation, although this is widely acknowledged to be a legal grey area. One such technique is with the Python library *Selenium* which provides an automated way of managing web searches.

For *96 Postcards*, *Selenium* automates a search on Instagram for posts at 96 unique geographic locations, e.g. Biarritz or Nice, by invoking and driving Google Chrome. This requires installation of the open-source tool *ChromeDriver*.¹ The critical parts of the Python code for conducting this search are presented in Figure 2. Note that for *96 Postcards* a dummy Instagram account was created from which all Instagram data is obtained.

```
#specify the path to chromedriver
driver = webdriver.Chrome('Path to ChromeDriver')

#open the webpage
driver.get("http://www.instagram.com")

#target username
username = WebDriverWait(driver, 10).until(EC.element_to_be_clickable((By.CSS_SELECTOR, "input[name='username']")))
password = WebDriverWait(driver, 10).until(EC.element_to_be_clickable((By.CSS_SELECTOR, "input[name='password']")))

#enter username and password
username.clear()
username.send_keys("account_name") #Instagram account name
password.clear()
password.send_keys("account_password") #Instagram account password

location = "Cyprus"
driver.get("https://www.instagram.com/" + location + "/")
driver.execute_script("window.scrollTo(0, 2000);")

#select images
images = driver.find_elements_by_tag_name('img')

#images = driver.find_element(By.TAG_NAME, "img")
images = [image.get_attribute('src') for image in images]
```

Figure 2. Python code for automating an Instagram search on “Cyprus” with the *Selenium* library and *ChromeDriver* open-source tool.

HyperText Markup Language (HTML) data returned by *Selenium* is analyzed using the *Beautiful Soup* library,² which provides an efficient, powerful means of extracting data from HTML files. The text caption and URL of the Instagram post is saved to a .csv file, see Figure 3, which is then called to generate the immersive score displayed in VR space within the Unity3D platform. The image associated with the post is also saved.

```
#urls of all scraped images
posts = []
links = driver.find_elements_by_tag_name('a')
desc_texts = []

for link in links:
    post = link.get_attribute('href')

    if '/p/' in post:
        posts.append(post)

for i in range(len(posts)):
    print(posts[i])

#parse html for descriptive texts
for post in posts:
    driver.get(post)
    soup = bs(driver.page_source, "html.parser")
    desc_text = soup.find("title")
    desc_texts.append(desc_text)

for i in range(len(desc_texts)):
    print(desc_texts[i])

df = pd.DataFrame(list(zip(posts, desc_texts)), columns = ['URLS', 'Texts'])
df.replace(to_replace = "<title>", value = "")
df.replace(to_replace = "</title>", value = "")
df.to_csv(96Postcards.csv)
```

Figure 3. Python code for grabbing Uniform Resource Locators (URLs) and text captions.

3. VISUALIZATION AND MAPPING STRATEGIES

After collating the Instagram text captions from 96 locations, a textual analysis is performed to obtain various formal and semantic properties which are then mapped to features of the performance score. In *96 Postcards*, this analysis is performed with a combination of simple Python queries and the use of the Natural Language Processing (NLP) library *spaCy*.³

Two curious insights with respect to typical Instagram text captions shaped decisions on how text might provide useful ways of generating musical structure. Firstly, the overwhelming number of texts polled were positive in sentiment. This effectively mitigated the need for commonly applied analytical techniques designed to provide quantitative data on a text's semantic meaning. Secondly, there is often significant use of emoticons which cannot be readily analyzed with *spaCy* dictionaries nor other popular NLP libraries such as *TextBlobs* or *NLTK*. While emoticon types could certainly be used as a means of providing

¹ <https://chromedriver.chromium.org/home>

² <https://www.crummy.com/software/BeautifulSoup>

³ <https://spacy.io>

structural organization in *96 Postcards*, when they ostensibly replace text, as in Figure 4a, they were not of particular interest, and accordingly such posts were discarded. An example of a caption that might be kept, despite its use of emoticons, is shown in Figure 4b.

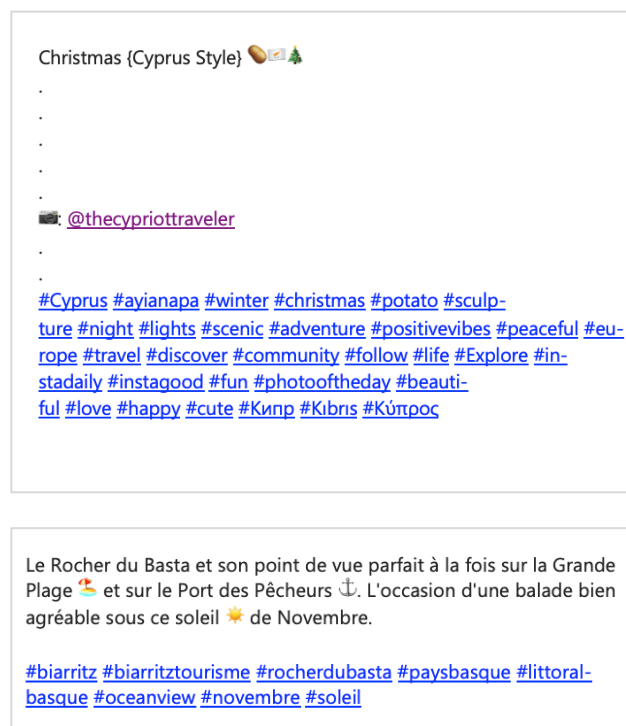


Figure 4. a) Sample Instagram caption on #Cyprus with no useful textual caption other than emoticons. Note in addition the relatively large number of hashtags which are also ignored in text analysis (upper), b) Sample Instagram caption on #Biarritz with an acceptable textual caption for inclusion (lower). In both examples, the Instagram user account name has been removed.

In *96 Postcards* the results of the textual analysis drive parameters of an immersive score which surrounds the performers, see Figure 5a. The notational schema builds on previous creative work by the author with 3D scores [12] and adopts a similar design aesthetic where a 3D construct of colored nodes, each of which denote the onset of a musical event, are connected by thin white lines which denote the duration of those events. In *96 Postcards*, this node/line construct is itself surrounded by twelve curved panels, or canvases in Unity3D nomenclature, partitioned into various subsections. Each of these partitions contain the Instagram image scraped from one of the 96 locations while the text caption accompanying one of these image postings is positioned above each panel, see Figure 5b.

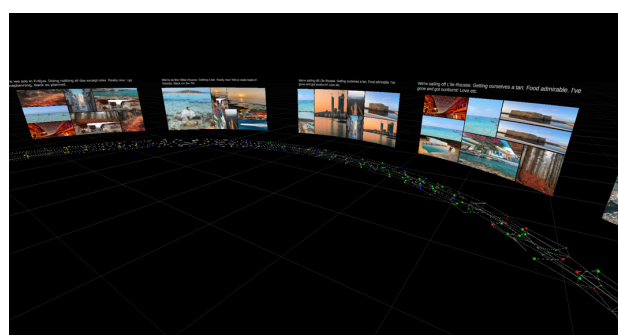
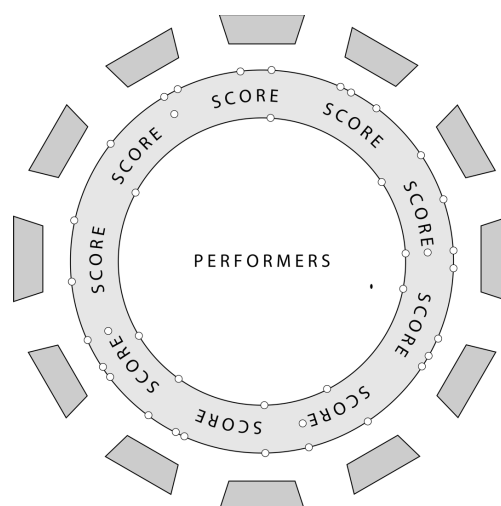


Figure 5. a) A bird's-eye view schematic of the immersive score for *96 Postcards* with performers positioned in the center of the VR scene and surrounded by the 3D score and image canvases (upper), b) a VR scene detail from Unity 3D's scene view (lower). Note the node/line construct of the immersive score below and slightly in front of the image canvases.

The mapping of linguistic data to features of the performance score was guided by various perceptual and musical considerations. In initial development, for example, the correlation of vowels within a text to pitch and harmonic content was thought to have organizational potential. But from a statistical perspective, the distribution of vowels within a series of text captions has negligible difference and hence any direct correlation of vowels to pitch or harmonic content was unlikely to result in a highly differentiated musical structure or at least a musical structure that might prove aesthetically satisfying. For this reason, it was decided that harmonic structure in *96 Postcards* would not be driven by NLP data but rather predefined. This had the additional benefit of ensuring a degree of continuity across performances.

In *96 Postcards* each node denotes the articulation of a pitch with the node's color indicating the specific pitch to be sustained along a line. During the performance, performers are instructed to freely read around the score from one node to another, musically exploring the myriad range

of pathways presented as they traverse the full 360-degree node distribution. As in previous works, the performance per se thus becomes an actualization of the latent possibilities presented by the non-linear, open form notational schema.

As noted, the harmonic content in *96 Postcards* is predetermined and not driven by the NLP analysis. In the score, this effectively means that the color distribution of nodes falls within a predefined map with weighted probability distributions applied, see Figure 6.

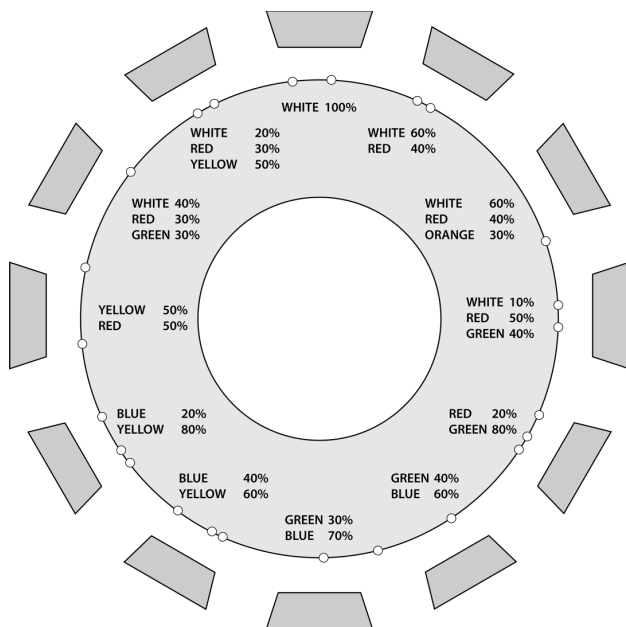


Figure 6. Harmonic mapping with probability weights around the performance score with node color mapped to pitch as follows – White (W) → Unpitched (whispered), Red (R) → C, Orange (O) → C quarter-tone sharp, Green (G) → G, Blue (B) → F, Yellow (Y) → B-flat. Pitches may be sung at any octave.

Unlike color distribution, which is not directly correlated to data returned by the text analysis, the spatial distribution of nodes in the score is driven by the spatial position of vowels within the Instagram text captions and other data returned by *spaCy* on the text's formal properties. Musically, vowel distribution therefore becomes a means of rhythmically striating a harmonic complex [13]. The spatial/temporal mapping is represented in the performance score by the thin white lines that connect nodes. These effectively function as a proportional notation with a line's spatial extension within the VR scene correlated to the temporal duration of a musical event. In *96 Postcards*, the mapping is prescribed at approximately 20 seconds across each canvas with smaller subdivisions resulting in proportionally shorter temporal divisions. Given that line is purely a rhythmic denotation, there was no need to ascribe additional graphic properties such as color or variety of thickness when developing the work, and hence each line is simply colored white with uniform width.

In the VR scene, the performers are surrounded by twelve panels partitioned into seven, eight, and nine sub-sections. Each partition is in turn filled with one of the 96 images collected in the Instagram scrape, see Figure 7. While the images themselves are not “interpreted” by the performers in any strict sense, they do provide an invaluable visual anchor as the performers traverse around the 360 degrees of the score. The text caption that accompanies each image post generates a discrete node/line construct that is positioned in front of its respective panel with the color and spatial distribution of nodes following the procedure previously outlined. As there are up to nine node/line constructs generated per panel, these are interleaved in order to be contained within the finite space at the front of each panel. While the interleaving of node/line constructs means that performers are not able to distinguish the individual constructs themselves, it does permit them to traverse across the visualizations, or across Instagram locations, as they navigate their way around the score. In this respect, as in previous work, the *dérive* model of performance is a particularly apt metaphor [14, 15].

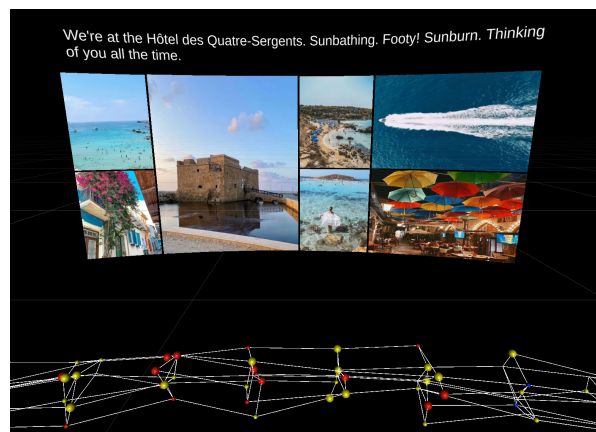


Figure 7. An illustration of a partitioned canvas from the performer's perspective. Note the text caption at top (from Perec's original text for illustrative purposes), the Instagram images in the canvas partitions, and a section of the immersive score's node and line construct slightly beneath.

The *dérive* model is further reinforced through the way in which the Instagram text captions function in the performance score. Above each of the twelve panels which surround the performers, one text caption is placed, see Figure 7. This caption provides a template for how node/line constructs directly beneath are to be enunciated. Considering compositional questions of mapping once again, the use of text was considered a far more natural means of conveying types of vocal articulation than any graphic lexicon especially given the ready source material from which node/line constructs were generated. In performance, the singers enunciate the node/line constructs through the successive vowels displayed in the Instagram text caption. For example, a caption such as "Le Rocher du

Basta et son point de vue parfait à la fois sur le Grande...”, see Figure 4, would require performers to articulate and sustain pitches as represented in the node/line constructs along the vowels “e o e u a a e o o i e u e a a i à a o i u e a e...” etc. Extending the *dérive* model of performance, each sounded text caption, sifted of plosives, fricatives, affricates and other consonants, becomes a filter through which all other visualized captions are sounded. These latent locations, given actuality only through a node/line construct are thus heard as ghost-like echoes through the exploratory *dérive* of the performer’s journey around the score.

4. VR DEVELOPMENT

4.1 Photon Engine SDK

The VR application of *96 Postcards* has been developed on the Unity3D platform for the Oculus Quest 2, a VR head-mounted display (HMD) released in 2020 which at of the time of writing is the world’s highest selling VR headset.⁴ Integral to the work is the use of the Photon Engine SDK, or PUN (Photon Unity Networking), which provides a framework for facilitating shared multiplayer experiences.⁵

Designed for the development of multiplayer games, PUN manages the networking back-end that allows players to be immersed in a common VR scene. The process through which this is achieved is straightforward and adopts a relatively standard protocol. Upon launching the *96 Postcards* application on the Quest 2, each performer is positioned in the center of the same VR scene through a simple series of automated synchronization calls managed by the PUN SDK. In *96 Postcards*, each of the performers passively observe the scene in which they are immersed. They do not interact with any of its constituent elements with hand controllers, nor do they interact with the avatars representing any of the other performers. Furthermore, each performer’s individual movements are constrained to a limited set of movements which enable them to observe the score that surrounds them. The limited set of interactions afforded in this novel system architecture, minimizes the data synchronization calls made to the PUN network and ensures that the scene information between performers is not only consistent, but presented with minimal latency.

4.2 UX considerations

Unsurprisingly, there is a paucity of experiential data that can be drawn from to help guide VR design for musical performance. Amongst the more generally cited UX concerns, is the need to ensure that the space in which players are immersed is comfortable [16]. Thus, it is important to minimize visual effects that may induce motion sickness or other vestibular disorders, through careful attention to

factors such as object movement, image resolution, field-of-view, and exposure time [17]. For *96 Postcards*, vestibular effects were not a pressing concern as the imagery surrounding the performers is entirely static, thus acting as an anchor to reduce any incumbent disorientation. Nevertheless, smooth headtracking where there is a one-to-one correspondence between the head movement speed and camera rotation within the VR scene, and the reduced number of PUN calls from the overall system architecture, further mitigated any negative vestibular effects.

A less often cited UX principle relates to text presentation. To ensure legibility, VR designers generally try to ensure that text is always static not just with respect to spatial location, but also appearance, for example not blinking and maintaining a stable color and style. In addition, the resolution of most HMDs requires font sizes to be larger than what UX designers may be typically accustomed. For example, the Quest 2 has a display resolution of 1832 x 1920 per eye whereas other popular VR HMDs have resolutions of 1440 x 1600 (HTC Vive Pro), 1280 x 1440 (Samsung Gear VR), and 960 x 1080 (Playstation VR). Even given the relatively high display resolution of the Quest 2, fonts need to be no smaller than 40 point to ensure minimal legibility from five feet (1.5m) away [18, 19]. As such, to include Instagram text captions within each canvas partition in the VR scene of *96 Postcards*, often means that some text could not be fully contained. Consequently, a decision was made to present text from one caption only on each canvas positioned directly above and offset from the images below to ensure maximum legibility, see previous Figure 6. While the score for *96 Postcards* does not employ common practice, stave-based notation, for composers who wish to use traditional schemas in VR scenes, issues of legibility and resolution are of even more pressing concern [20].

Unlike Augmented Reality (AR) technologies, the wearer of a VR HMD is fully immersed in a virtual world and not expected to interact with any physical objects in the real-world. To ensure wearer safety, Quest 2 users usually start their virtual experience by tracing a safe work area on the floor around them with a hand controller. Should the wearer approach the boundary of the safe space, a warning grid appears in their visual display. Full immersion in a VR scene naturally limits the types of instruments musicians can comfortably perform to those not requiring rapid eye-hand coordination or occupying large physical spaces. While the designers of the Quest 2 have recently permitted developers to access the pass-through camera of the HMD and project this image into the VR scene, the resolution is exceptionally poor at 1080x1200 and is also black-and-white. Both these factors make it difficult for the wearer to interact meaningfully with any objects in the real world while continuing to be present in VR space. These various constraints were instrumental in the pre-compositional decision to develop *96 Postcards* for an

⁴ Meta/Facebook, the developers of the Quest 2, do not publish official sales figures although current estimates are in the order of more than 1.2 million units per quarter.

⁵ <https://www.photonengine.com/pun>

ensemble of vocalists rather than other musical instruments.

Finally, while perhaps not a UX consideration from the perspective of the performer, the use of VR HMDs in musical performance raises a challenging UX question from the perspective of the audience. Unlike a traditional performance dynamic, the performers of *96 Postcards*, and indeed any VR work, are visually detached from the audience and other performers. Indeed, in VR space the traditional proscenium-based performance model breaks down. With the audience always situated outside the VR space in which the performers are immersed, the performance itself becomes a form of spectacle upon which the audience can only partially observe. How UX might be optimized for the audience is a challenging question. Whether they should be invited into this world, perhaps through a shared screen projection of the VR scene, or remain passively outside is ultimately, perhaps, an aesthetic question to be uniquely resolved for each VR experience.

5. FUTURE WORK

While my previous work in mixed reality has explored the use of AR technologies such as Microsoft's HoloLens as a means of displaying performance scores [21], *96 Postcards* is my first work with a graphic score to be specifically designed for a VR environment. While the technical affordances of AR and VR systems have proven especially attractive, it is their ability to offer new forms of creative expression which has provided the richer field of possibilities.

As a tool for immersive visualizations of large data sets, whether they function as graphically notated performance scores, visual works of art, or more scientifically focused efforts to enrich understanding, VR offers particularly exciting possibilities. In his recently published text *Cultural Analytics* [22], Lev Manovich explores many of these and particularly how computational analysis of images, whether gathered from Instagram, print media, or other sources, can help facilitate cultural analysis. The creation of new forms of aesthetic expression from analyses of other forms of aesthetic expression would itself seem to offer exciting possibilities for innovative, creative enquiry. For the author's creative practice, this is perhaps the most fertile area for future exploration, aestheticizing multi-player agency and refracting it back into the social media platforms upon which it reflects and is drawn. Further work might also seek to obtain measures for optimizing the UX experience for the audience, and perhaps finding ways in which their agency might itself provide a means of transforming visualized sets of data.

Acknowledgments

Much of my recent research into musical applications of virtual and augmented reality has been made possible thanks to a Churchill Fellowship awarded in late 2019 and

completed in 2021. I would like to thank the Churchill Trust for this opportunity and for their generous support.

Thanks also to Dr. John Taylor from The MARCS Institute for Brain, Behaviour & Development at Western Sydney University for his many helpful suggestions on an earlier draft of this paper.

6. REFERENCES

- [1] G. Perec, "Deux Cent Quarante-Trois Cartes Postales en Couleurs Veritables," *Le Fou Parle*, 1978.
- [2] T. Bonch-Osmolovskaya, "Combinatorial Greetings from Georges Perec," in *Bridges 2018 Conference Proceedings*, Stockholm, 2018. Available: <https://archive.bridgesmathart.org/2018/bridges2018-253.pdf>
- [3] G. Perec, *Species of Spaces and Other Pieces*. John Sturrock (Ed. And Trans.), Penguin, 1999.
- [4] W. Motte, *Oulipo: A Primer of Potential Literature*. Dalkey Archive Press, 1998.
- [5] A. James, *Constraining Chance: Georges Perec and the Oulipo*. Northwestern University Press, 2009.
- [6] P. Terry (Ed.), *The Penguin Book of Oulipo*. Penguin, 2021.
- [7] H. Philipsen, "Constraints in Film Making Processes Offer an Exercise to the Imagination," *International Journal of Media, Technology and Lifelong Learning*, vol. 5, no. 1, 2009.
- [8] K. Waschneck, *Screening Oulipo – From Potential Literature to Potential Film*, PhD Thesis, University of Essex, 2017. Available: <http://repository.essex.ac.uk/22413/1/Screening%20OULIPO.pdf>
- [9] K. Terzidis, *Algorithmic Architecture*. Routledge, 2016.
- [10] M. A. Boden and E. A. Edmonds, "What is Generative Art?" *Digital Creativity*, vol. 20, no. 1-2, 2009, pp. 21-46.
- [11] D. Kim-Boyle, "The Twittering Machine," in *Proceedings of the TENOR2022 International Conference on Technologies for Music Notation and Representation*, Marseille, 2022.
- [12] D. Kim-Boyle, "The 3-D Score," in *Proceedings of the TENOR2017 International Conference on Technologies for Music Notation and Representation*, A Coruña, 2017.
- [13] P. Boulez, *Boulez on Music Today*. S. Bradshaw and R. R. Bennett (Trans.), Faber and Faber, 1975.
- [14] G. Debord, "Theory of the Dérive," *Les Lèvres Nues* No. 9 (Paris, November 1956). Reprinted in *Internationale Situationniste* No. 2 (Paris, December 1958). K. Knabb (Trans.).
- [15] K. O'Rourke, *Walking and Mapping: Artists as Cartographers*. MIT Press, 2013.

- [16] C. Hillmann, *UX for XR: User Experience Design and Strategies for Immersive Technologies*. Apress, 2021.
- [17] D. Saredakis, A. Szpak, B. Birkhead, H. A. D. Keage, A. Rizzo, & T. Loetscher, "Factors Associated With Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis," *Frontiers in Human Neuroscience*, vol. 14, 2020. Available: <https://www.frontiersin.org/article/10.3389/fnhum.2020.00096>
- [18] V. Kurbatov, "10 Rules of Using Fonts in Virtual Reality." [Online] Available: <http://vovakurbatov.com/articles/10-rules-of-using-fonts-in-virtual-reality>
- [19] E. Van der Burg, C. N. L. Olivers, J. Cass, "Evolving the Keys to Visual Crowding," *Journal of Experimental Psychology: Human Perception and Performance*, 43:4, 2017, pp. 690-699. Available: <https://doi.org/10.1037/xhp0000337>
- [20] L. Vickery, "The Limitations of Representing Sound and Notation on Screen," *Organised Sound*, vol. 19, no. 3, 2014, pp. 215-227.
- [21] D. Kim-Boyle, "3D Notations and the Immersive Score," *Leonardo Music Journal*, Vol. 29, 2019.
- [22] L. Manovich, *Cultural Analytics*. MIT Press, 2021.