P.S. This is a compressed version. The lossless version can be downloaded at Google Drive:
https://drive.google.com/drive/folders/1hKTGrDuFD1mlW-FsGSyZmxLjcZ3KDiHW?usp=sharing

Portfolio

Junxiu Tang

About

Junxiu is an artist and researcher whose work explores the boundary between scientific reasoning and sensory experience. He is a Postdoctoral Fellow at Northwestern University's *Center for Science of Science & Innovation* (CSSI), where he is part of the newly established *CSSI Art Studio*, a research initiative that transforms large-scale scientific data into visual, material, and spatial forms. His practice draws from computer science, digital media, and architectural design, focusing on data-driven storytelling, spatio-temporal visualization, and the translation of complex systems into poetic and culturally resonant art-works.

His work examines how computation, craft, and perception converge to generate new ways of un-derstanding both the physical world and human knowledge. His projects have been exhibited at the SIGGRAPH Art Gallery, IEEE VIS Arts Program, and the Info+ Exhibition, and his research has been published across computer graphics, HCI, visualization, and interactive media venues.

His artistic approach is shaped by close collaboration with the interdisciplinary scientific team at *CSSI*, whose research investigates the structure and evolution of knowledge across scales. Working within this ecosystem, Tang creates artworks that reimagine scientific data not merely as information, but as a site of metaphor, imagination, and aesthetic exploration.

https://ahugh19.github.io/
junxiu.tang@kellogg.northwestern.edu
Center for Science of Science & Innovation (CSSI), Kellogg School of Management, Northwestern University
1800 Sherman Ave., Evanston, Illinois, US. 60201

Preface

In a world where paths diverge, scientists and artists find ways to converge, bringing people together.

In this portfolio, I present four projects that delve into visualization and data-driven art. These projects were developed through collaborations with my team and myself, with me leading three of them.

The projects span a range of topics and employ various techniques and presentations. While their aesthetic appeal is a key focus, they are also grounded in scientific principles, blending artistic expression with data-driven insights to create emotionally resonant experiences for audiences.

We explore the often-overlooked voices on social media, trying to feel their loneliness and isolation. We investigate the essence of information communication and its eternal losses, discussing the philosophy behind it. We celebrate the diversity of individual life paths in contrast to societal expectations and stereotypes. We reflect on the connection between past and present, appreciating the abstract themes and subjective emotions found in Chinese classical poetry.

These projects mark a significant milestone in my artistic journey and stand as a testament to the collective growth of my team and me. As you explore these works, I hope you encounter the limitless possibilities at the intersection of data and art, and feeling the power of the emotions and messages we aim to convey. Join us on this journey of discovery and immerse yourself in the fascinating world where data and art converge.

Catalogue

About.....I
Preface.....II
Catalogue.....III

Posts with no Response: The Island of Loneliness.....1
Loss of *Sonnet 18*.....13
Life Thread.....25
Explore Mindfulness without Deflection.....33

Posts with no Response: The Island of Loneliness

Junxiu Tang, Rui Sheng, Yifang Wang, Xinhuan Shu,
Xiaojiao Chen, Tan Tang, and Yingcai Wu
2023

IEEE VIS Arts Program. Library at the Dock Gallery, Melbourne, Australia. Oct. 04-29, 2023.
Information is Beautiful Awards Long List. Online. 2023.
Information+ Conference. Inspace, Informatics Forum, Edinburgh, UK. Nov. 22-24, 2023.
China Visap. Lee Shau Kee Library of HKUST, Hong Kong, China. Jul. 21-Aug. 01, 2024.



Abstract
Loneliness and isolation are eternal emotions in human beings. Technological advancements create ample avenues, like social medias, for individuals to articulate themselves and record emotions. However, the sense of loneliness has never vanished, as their expressions are easily buried in the digital stream. We analyze tweets that express loneliness during holiday seasons but receive few responses. By superimposing digital charts on physical models, we visualize these lonely posts and generate the island of loneliness. We aim to reveal the complexities of human emotions in the digital age and reflect on the interconnections between technology, solitude, and social communication.

Mediums
Wood; Video projection; Data physicalization.

Keywords
Social Media; Loneliness; Data physicalization; Data Visualization; Data Art.

Website
https://osf.io/3kt8d/

Loneliness is not on the mountain
but on the street, not within one
person but among many people.

——Kiyoshi Miki



Introduction

Social media platforms generate a constant flow of digital expressions, opinions, and emotions, forming a mass of data that constantly exists and continuously evolves. While much attention is often given to opinion leaders and posts that receive extensive responses, few works explore the unacknowledged expressions in this vast digital landscape.

These unresponsive posts, seemingly insignificant in their weight or impact, hold an intrinsic value as reflections of individuals' thoughts and emotions at specific moments. Like grains of sand sinking and settling at the bottom of the data sea, these posts form an unseen layer that merits our attention. We explore these overlooked expressions, focusing on posts that convey feelings of loneliness. The absence of responses amplifies the sense of isolation, compelling us to shed light on these silent voices and highlighting their existence and significance. They record human loneliness and indicate its perpetual presence: even in a world where the Internet connects people and everything, individuals can still experience deep feelings of loneliness.

The loneliness experienced on social media is a complex phenomenon that has been studied [3,7]. This project aims to amplify the expression of loneliness in social media through visualization. We first collect social media posts that received no response, particularly during poignant periods like the Christmas and New Year holiday seasons. Through data-driven approaches, we extract underlying topics and unravel the emotional undercurrents embedded within each post. We try to comprehend the multifaceted nature of loneliness within social media, the digital and perpetual landscape. Finally, we undertake the task of visually representing these unattended expressions. By creating a metaphorical island, we symbolically manifest the accumulation of these lonely "sands" in both the digital realm and the physical world. This visual representation serves as a testament to the enduring presence of these expressions and embodies the continuous nature of the data-driven landscape we inhabit.

In the vast mass of the data-centric world, this work strives to unveil the hidden narratives and untold stories within the expansive sea of data. By bringing attention to these neglected expressions and illuminating them through visualization, we aim to deepen the collective understanding of the intricate interplay among data, human emotions, and the ever-evolving fabric of our world.

Data Collection and Analysis

We choose Twitter, one of the most popular online social media platforms, as our data source. Twitter enables users to post tweets and receive responses such as reposts, comments, and likes.

We collect a vast dataset comprising 162,376,361 tweets during the Christmas and New Year holiday seasons between 2016 and 2019. The collection was achieved using a roughly 1% sample rate of the entire tweet stream.

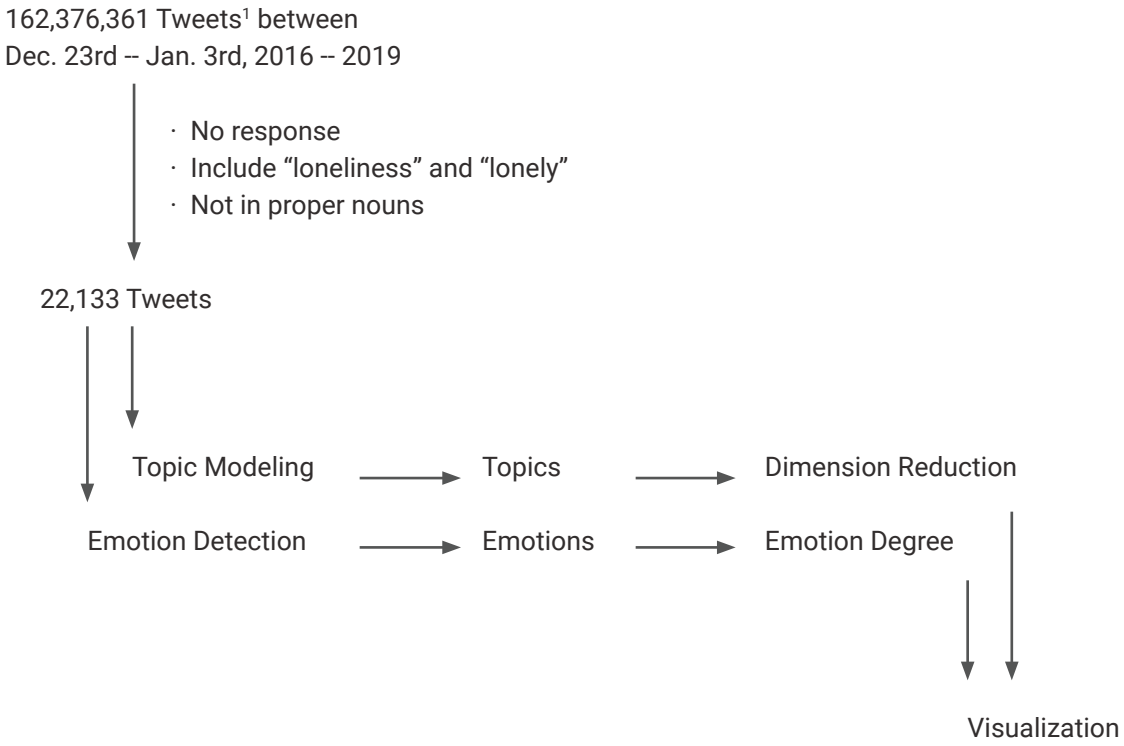
Upon analysis, we find that the majority of tweets receive no responses from others. This leads us to narrow the filter criteria to identify those lonely tweets that were left unnoticed.

Following the data processing methodologies utilized in previous studies [1, 4], we specifically select tweets containing keywords related to “lonely” or “loneliness”. To avoid confounding factors, we exclude tweets where these words appear as proper nouns. As a result, our dataset consists of 22,133 tweets that capture the existence of loneliness.

To extract meaningful insights from the collected tweets, we employ a Latent Dirichlet Allocation (LDA) model [2,6] to extract meaningful insights from the different contents. In order to strike a balance between accuracy and efficiency, we adjust and finally set the number of topics as 20, ensuring a representation of the underlying themes within the dataset.

Additionally, we utilize an Emotion English DistilRoBERTa-base model [5] to detect the emotions conveyed in each tweet. This model provides us with a comprehensive understanding of the emotional nuances encapsulated within the tweet content.

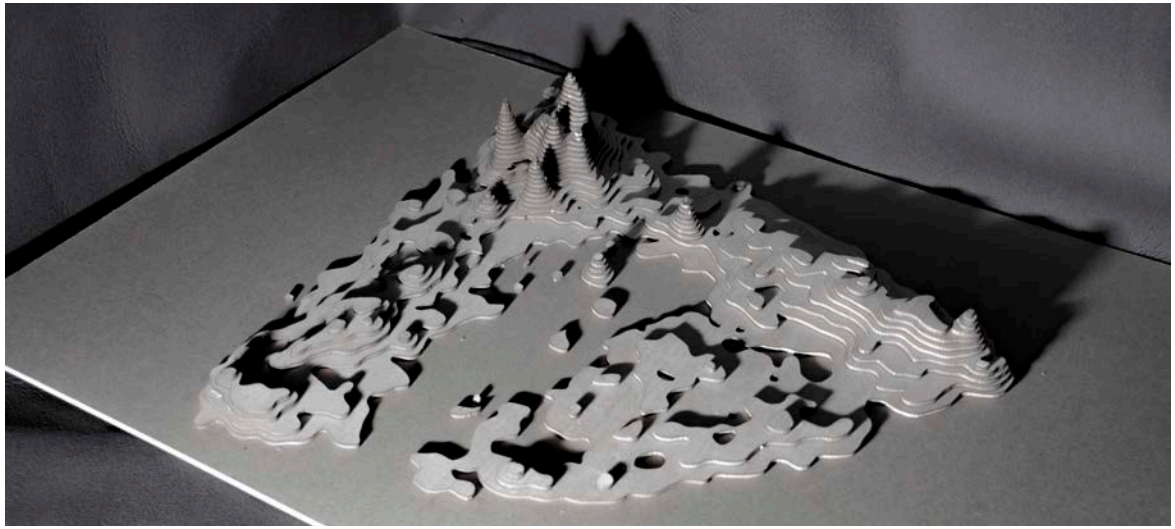
For each tweet, we utilize the LDA model [2, 6] and the emotion detection model [5] to generate a list of topics and a list of associated emotions, respectively. These lists include weights that indicate the relevance or significance of each topic and emotion to the tweet. Based on these multiple weights, we represent each tweet with a multi-dimensional vector. The vector plays a crucial role in our visualization process.

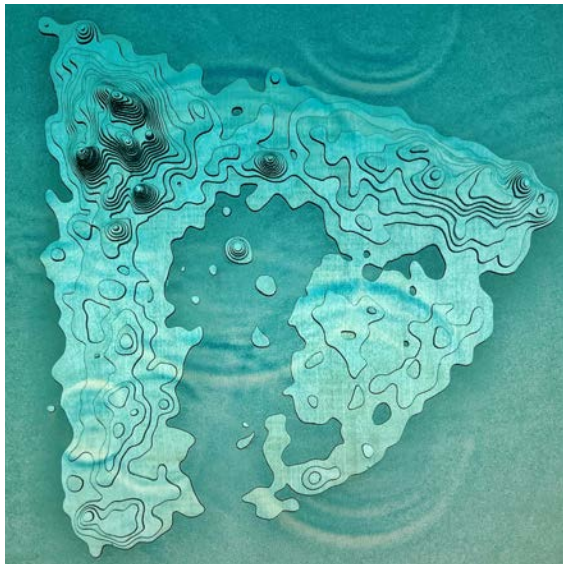


The workflow of data collection and analysis in this work.

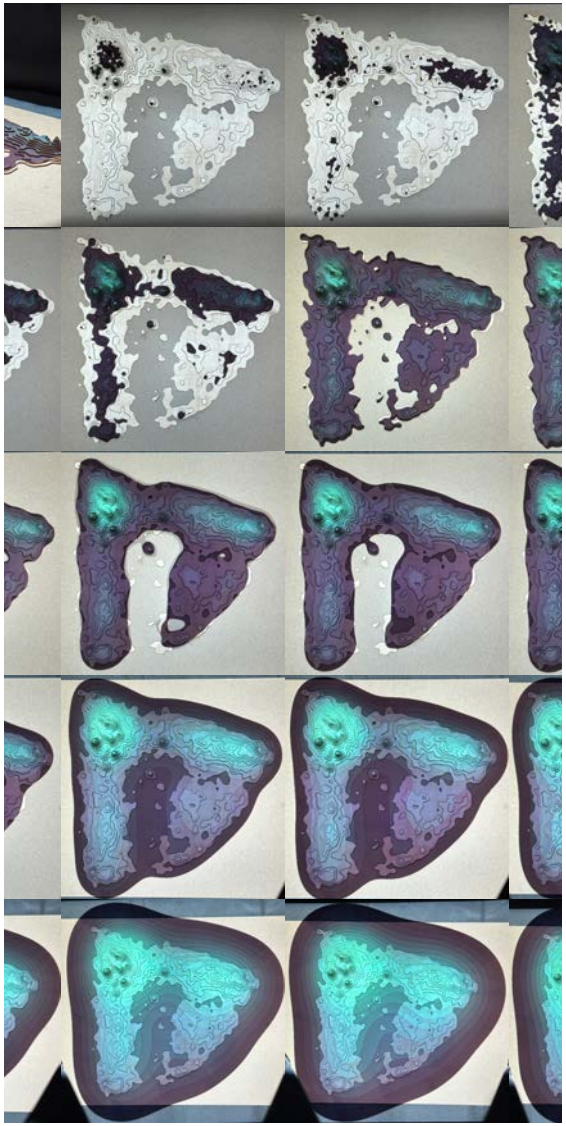
[1] By Twitter Standard APIs v1.1: https://developer.twitter.com/en/docs/twitter-api/v1

Overview









The physical model, projections of ripples and contours.
The supplementary materials can be found in
https://osf.io/3kt8d/.

Metaphor

In the digital world, social media platforms can be likened to an expansive, boundless sea brimming with many semantic data. Individuals continuously contribute their emotions, feelings, and opinions to the torrential stream of information.

Each post acts as a metaphorical stone cast into the sea, with some carrying significant weight, size, or other captivating features that generate continuous ripples and water flowers, sparking mass responses within the crowd.

However, the majority of posts bear lightweight and unremarkable appearances, failing to generate any notable ripples or water flowers. These posts akin to grains of sand, simply sinking to the seabed. While they may temporarily resurface through topic resonance when certain keywords are searched, they primarily exist as unnoticed and unattended sedimentations of information.





People threw their thoughts and emotions to the data sea.
<1> Some of them are big and weighty,
<2> while some of them are small and unimpressive.



<3> Only a small number of stones make noises and generate continuous ripples.



<4> The majority of stones straightly sink into the bottom of the sea.

<1> Photo by Jerome Ru.
<2> Photo by PAN XIAOZHEN .
<3> Photo by Linus Nylund.
<4> Photo by Yannis Papanastasopoulos.
All above photos are on Unsplash (https://unsplash.com/).

Design Process

We first introduce the major elements of this project, namely **loneliness contour**, **emotion side view**, and **noisy ripples**, and then report the implication in both digital and physical versions.

Loneliness Contour

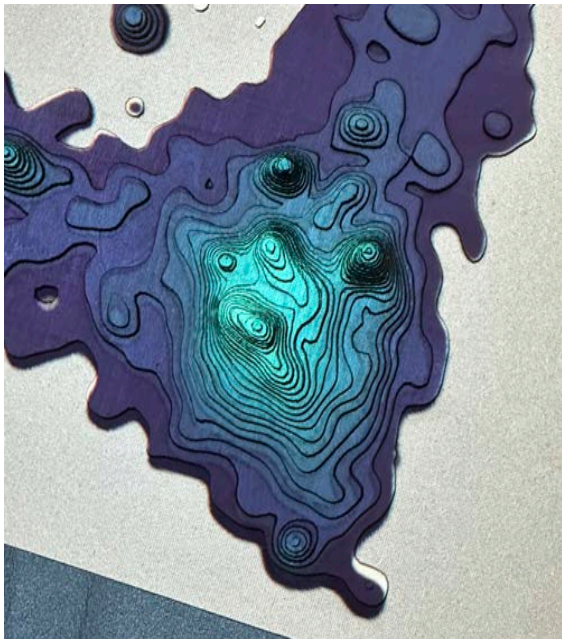
We gather the solitary sands to uncover the hidden island of loneliness based on the collected data. To intuitively convey the results of topic modeling analysis and visually align with the geographical island metaphor, we employ Isomap dimension reduction after testing commonly used non-linear dimensionality reduction methods.

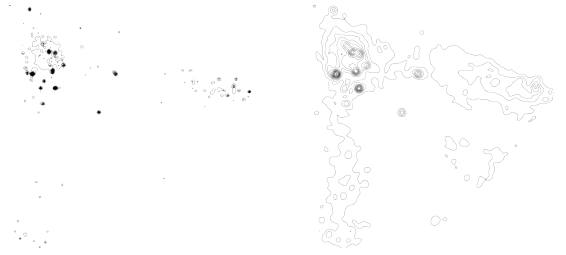
By reducing the vector dimensions to two, we are able to render a scatter plot of all the tweets within an x-y space. The resulting 2D scatter plot is the basis of the island layout. In the scatter plot, each point represents a tweet, and the proximity of points indicates thematic similarity. More adjacent points (tweets) share closer thematic connections. Subsequently, we calculate the density of these tweets and utilize the density contour to delineate the shape of the island. This mapping provides valuable insights into the popular topics frequently mentioned within the realm of lonely tweets.

We simulate the accumulation process of the island through adjusting the bandwidth. In the projection, we utilize a sequential color scale (from dark blue to bright blue) to encode the height of the island.









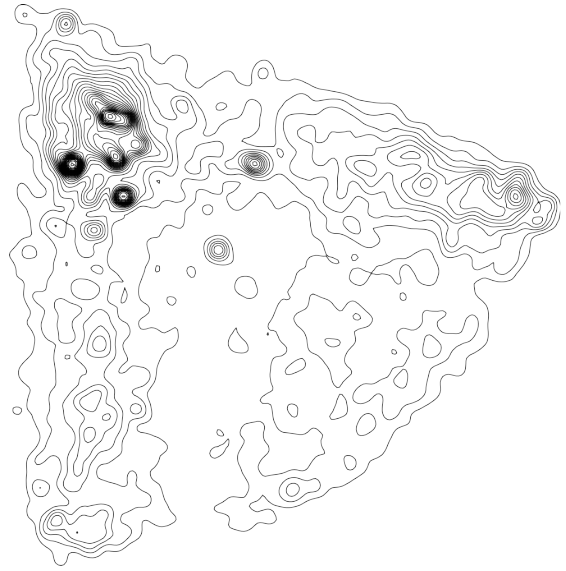
bandwidth = 1

bandwidth = 6

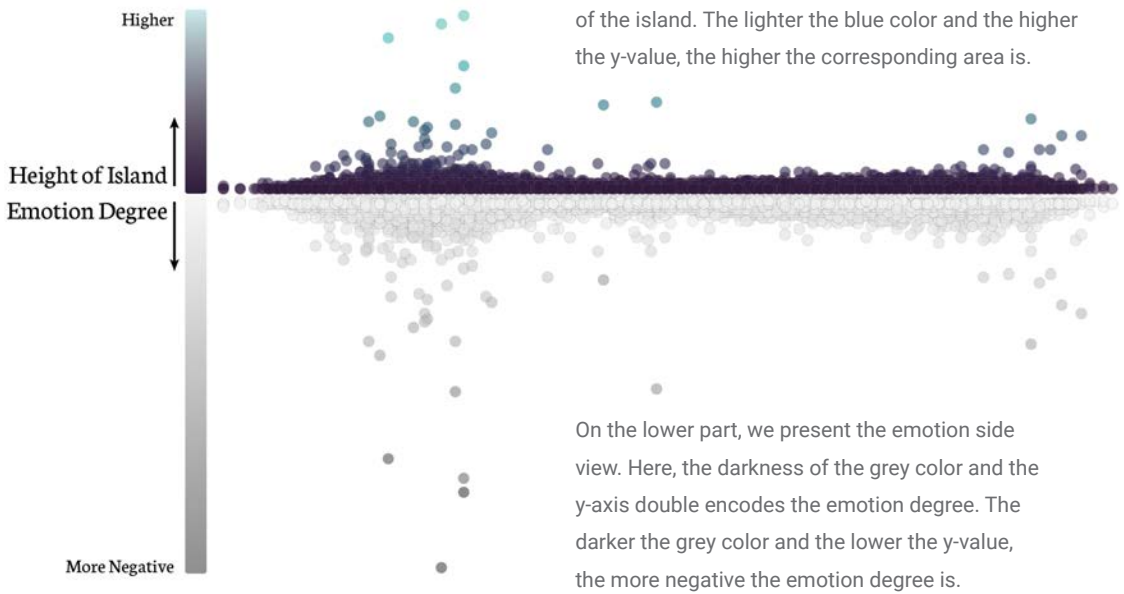


bandwidth = 15

bandwidth = 40



bandwidth = 8



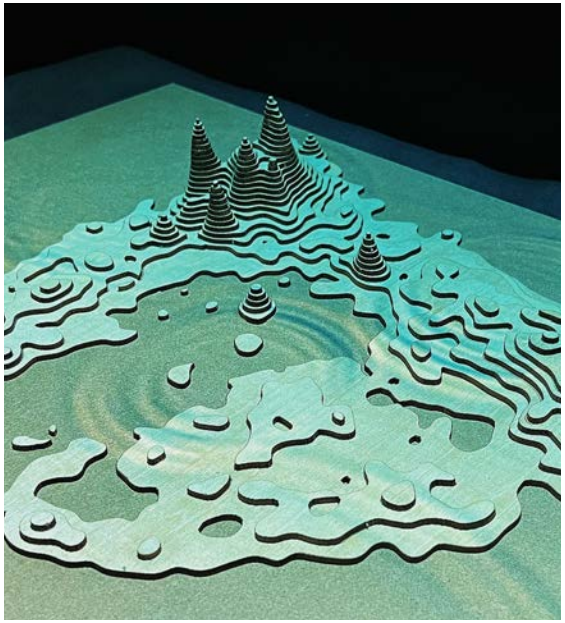
The upper part represents one side view of the island, with a sequential color scale (ranging from dark blue to bright blue) and the y-axis double encoding the height of the island. The lighter the blue color and the higher the y-value, the higher the corresponding area is.

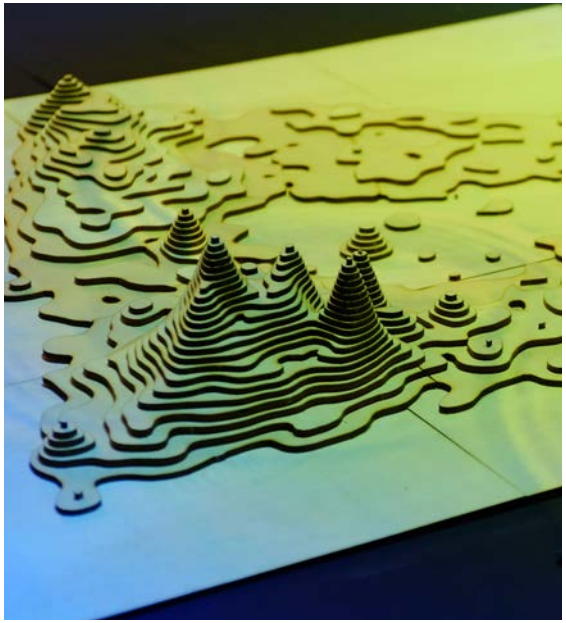
On the lower part, we present the emotion side view. Here, the darkness of the grey color and the y-axis double encodes the emotion degree. The darker the grey color and the lower the y-value, the more negative the emotion degree is.

Emotion Side View

In order to assess the emotional overview embedded within the tweets, we calculate the emotion degree by assigning positive weights to active emotion values and negative weights to negative emotion values. In this context, the “emotion values” are the absolute values of both kinds of emotions, derived from the results of emotion detection [5].

While the tweets themselves are visible, the underlying negative emotions that often accompany lonely tweets remain easily overlooked, much like the submerged portion of an iceberg.

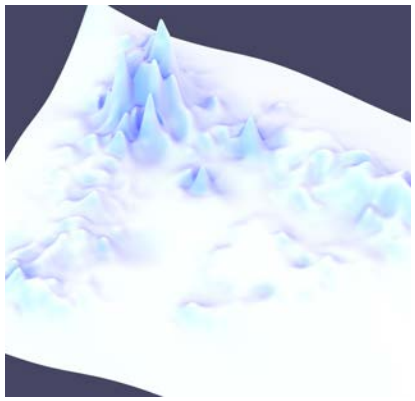




Noisy Ripple

In addition to visualizing the sunken land beneath the sea, we aim to capture the bustling surface of the water through the metaphor of ripples. We encode the intensity of response with the radius of each ripple. The response degree is defined as the cumulative number of reposts, likes, and comments, indicating engagement and communication.

In this context, ripples serve as an indicator of the presence of responded posts when the unresponded posts sediment into the seabed. The noisy ripples are deliberately juxtaposed with the quiet sedimentation to emphasize this contrast.



The model rendered in Rhino.



Laser cutting.



Manual assembly.



Projection setting.

Implementation

The visualization work encompasses a harmonious fusion of digital charts and physical models. The loneliness contour and emotion side view are rendered using D3.js.

To deepen the solid presence of loneliness, we physicalize the data-driven contours. We choose the contour (bandwidth=8) with ample details for physicalization. The initial step involves transforming the contour chart into a threedimensional model using the Rhino software. Subsequently, we finetune the contour lines in AutoCAD, ensuring precision and practicability, and fabricate the individual components of the island using laser cutting techniques with wooden materials. Through the meticulous assembly of these components, a tangible and tactile physical model of the island is brought to life.

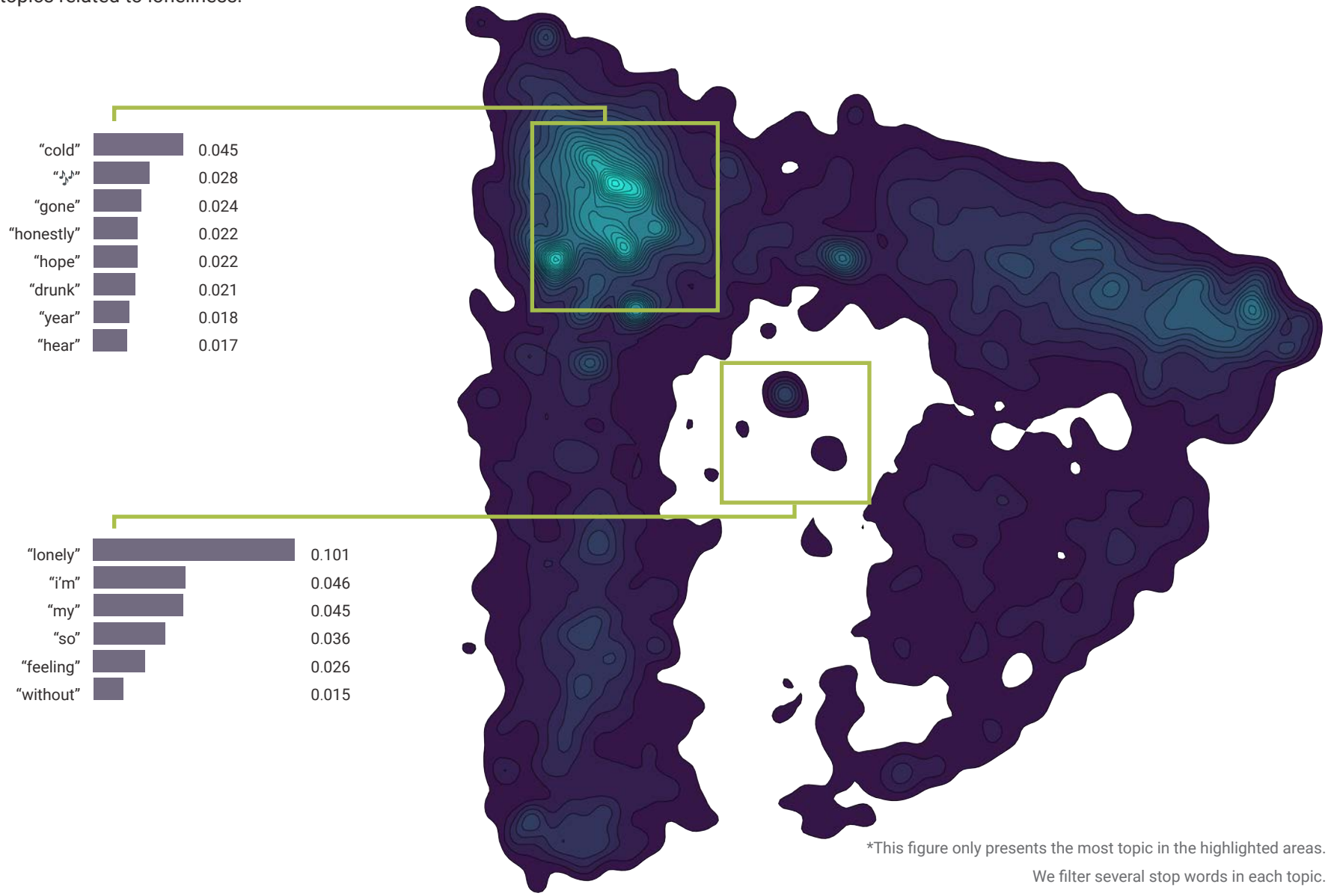
We project the ripples and contours over the island model temporally. By incorporating a physical model, we offer an additional sensory dimension to the experience. Through touch interaction with the physical model, viewers can further engage with the data, extending their understanding beyond the visual channel and enriching their perceptual encounter with the visualization.





Case: Lonely Tweets

We present what people tweet when they express loneliness. The higher places on the island indicate popular topics related to loneliness.



*This figure only presents the most topic in the highlighted areas.
We filter several stop words in each topic.

Here are some lonely Tweets:

It's 3 am I must be **lonely**.

You are too lovely to be so **lonely**.

He is the **lonely** one ... he is the sad one ... the greedy one ...

still **lonely** at 🏁😪

I'm beyond **lonely**.

I just realized I'm gonna be **lonely** for Christmas. My family is 2000+ miles away. I have no bae. My friends are with me

It...strikes me as a very long and **lonely** path through the year as opposed to looking forward to anything.

Idk if any of y'all felt so **lonely** for ya parents but I am. Like so fxxking bad! I'm honestly not used to this.

lonely I get and am. I'm not perfect and lord knows I...talk about it alot anymore but I don't know what else to do...

lonely it is

lol I'm so broke and so **lonely** and so sad haha :'+)

Mood: **lonely**.

#joinin when am I going to spend Christmas with those I love. Why is it so **lonely** and sad

What a **lonely** day it's been 😩

This Christmas has been **lonely** af for me. Again, not bixxhing.

Man had this morning sucked. One of my friends wants to deactivate and I'm **lonely**. FML.

Well, one of my best Twitter friends wants to deactivate and I'm feeling **lonely** without my bf.

I'm scared of **lonely**

I'm really trying but this **lonely** feeling is getting the best of me.

Shhhh. I know that. I'm just fxxking **lonely**. I hate it.

Idk. This is one of them nights, I feel so **lonely** for everyone.

Fxxk this. I'm too **lonely**. 😭💔

A boyfriend would be nice. Even just a friend. Being **lonely** sucks. I just want someone I can share my life with

i just want to not be **lonely** this new year's :(

pull up to your function being **lonely** as fxxk!

Being **lonely** during the holidays is NOT a good feeling.

I wish I had someone other than my mom to rely on. I be **lonely**.

It just makes you... kind of really **lonely**

Tomorrow has the potential for greatness but also ennui and self loathing and crippling **loneliness** and depression so lessmakeitagood pls

SINGLE is not a relationship status. Its a word describing your **loneliness** If you ask me, yes I'm SINGLE.

If a star fell each time I thought about you then the moon would truly realise what **loneliness** is really like.

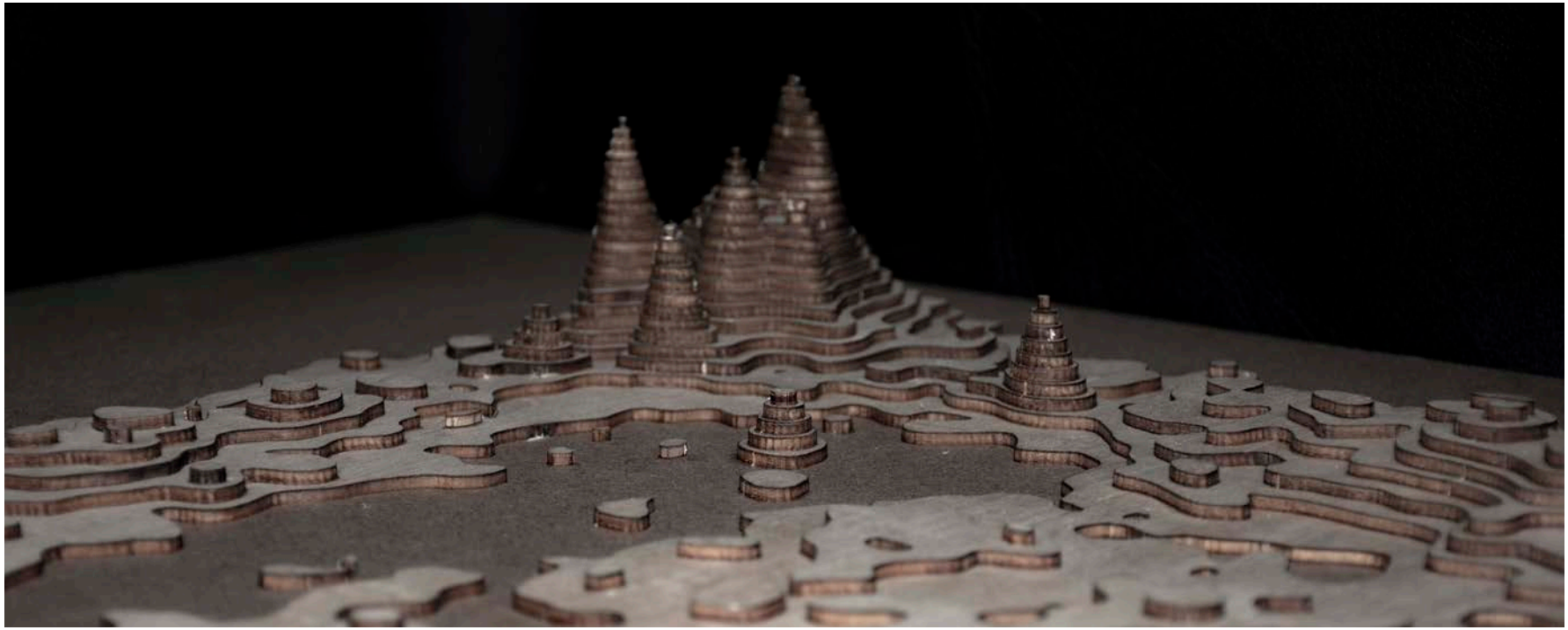
Discussion

By shedding light on the unattended expressions of loneliness, we strive to provide a deeper understanding of the emotional landscape within social media platforms. Through the data-driven approach and visual representation, we hope to illuminate the hidden narratives embedded within the vast sea of social media.

Social media serves as a unique intersection between public information and private life. It provides individuals with a platform to express their opinions on hot topics while also offering a space for personal reflection and documentation. However, not every post is shared with the expectation of receiving a response. Sometimes, individuals simply use social media as an outlet to express their emotions without a specific purpose in mind.

It is crucial to acknowledge that not all posts without responses should be considered isolated cases. The absence of engagement does not necessarily indicate the insignificance of these expressions. Instead, it highlights the complex nature of social media dynamics and the diverse motivations behind sharing personal experiences. Whether individuals who feel lonely seek any form of responses when expressing their loneliness requires further investigation.

In this art project, we collect social media posts using a heuristic approach under a limited temporal period, understanding that it may not capture all the relevant data related to loneliness or other negative emotions. There are likely many hidden posts within the vast expanse of social media that remain unexplored. By acknowledging these limitations, we recognize the ongoing exploration and the potential for future research to uncover additional insights into the intricate interplay between social media and loneliness.



References

1. Anietie Andy, Garrick Sherman, and Sharath Chandra Guntuku. 2022. Understanding the expression of loneliness on twitter across age groups and genders. Plos one, 17(9):e0273636. https://doi.org/10.1371/journal.pone.0273636
2. David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. Journal of Machine Learning Research, 3:993–1022.
3. Verolien Cauberghe, Ini Van Wesenbeeck, Steffi De Jans, Liselot Hudders, and Koen Ponnet. 2021. How adolescents use social media to cope with feelings of loneliness and anxiety during covid-19 lockdown. Cyberpsychology, Behavior, and Social Networking, 24(4):250–257. https://doi.org/10.1089/cyber.2020.0478
4. Sharath Chandra Guntuku, Rachelle Schneider, Arthur Pelullo, Jami Young, Vivien Wong, Lyle Ungar, Daniel Polsky, Kevin G Volpp, and Raina Merchant 2019. Studying expressions of loneliness in individuals using twitter: an observational study. BMJ open, 9(11):e030355. https://doi.org/10.1136/bmjopen-2019-030355
5. Jochen Hartmann. 2022. Emotion english distilrobertabase. Retrieved June 2, 2023 from https://huggingface.co/j-hartmann/emotion-english-distilroberta-base/
6. Hamed Jelodar, Yongli Wang, Chi Yuan, Xia Feng, Xiahui Jiang, Yanchao Li, and Liang Zhao. 2019. Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey. Multimedia Tools and Applications, 78:15169–15211. https://doi.org/10.1007/s11042-018-6894-4
7. Matthew Pittman and Brandon Reich. 2016. Social media and loneliness: Why an instagram picture may be worth more than a thousand twitter words. Computers in Human Behavior, 62:155–167. https://doi.org/10.1016/j.chb.2016.03.084

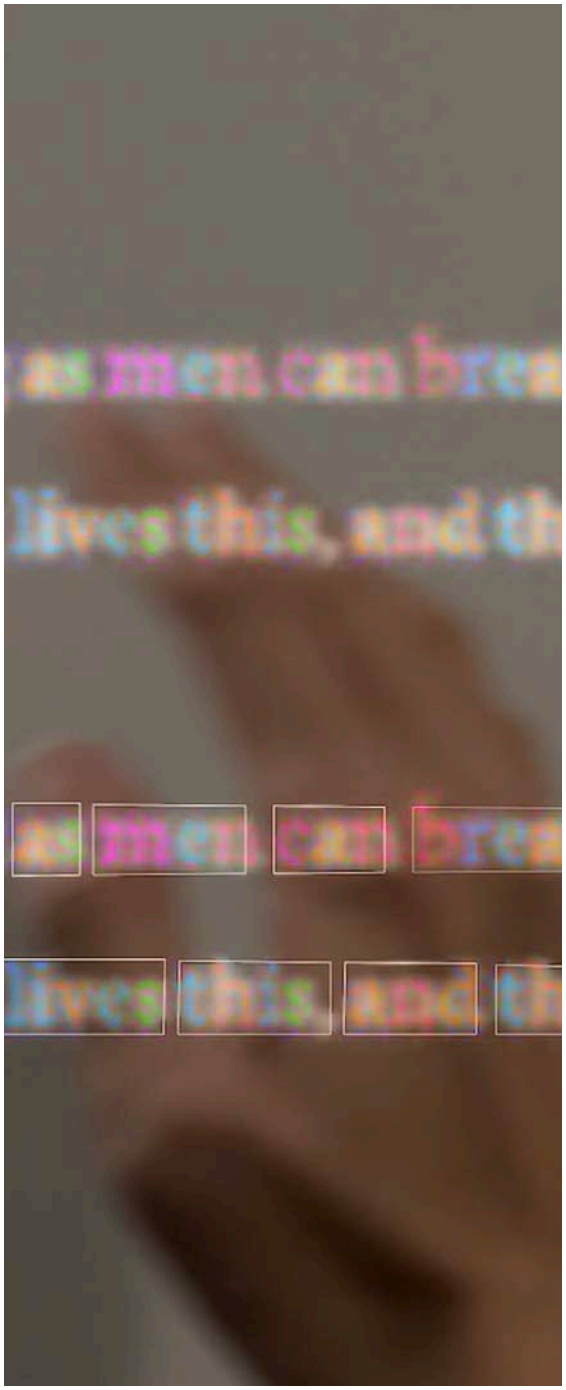




Loss of Sonnet 18

Junxiu Tang, Jiayi Zhou, Yifang Wang, Xinhuan Shu,
Peiquan Xia, Xiaojiao Chen, Tan Tang, and Yingcai Wu
2024

SIGGRAPH Art Gallery. Colorado Convention Center, Denver (CO), US. Jul. 27-Aug. 1, 2024.



Abstract
This project is an explorative digital art installation that delves into the phenomenon of generation loss in digital signal processing - the inevitable degradation of information (e.g., words) quality through copying and propagation. It transcends the technical realm to probe how this concept applies to the transmission and transformation of words and meanings through time and technology. It consists of a series of videos illustrating the loss of words themselves and their meaning, combined with real-time interpretation from both humans and AI. The project allows the audience to find a border of loss, beyond which, information of words might change.

Mediums
Video Installation.

Keywords
Generation Loss; Information Communication; Noise; OCR; Data Art.

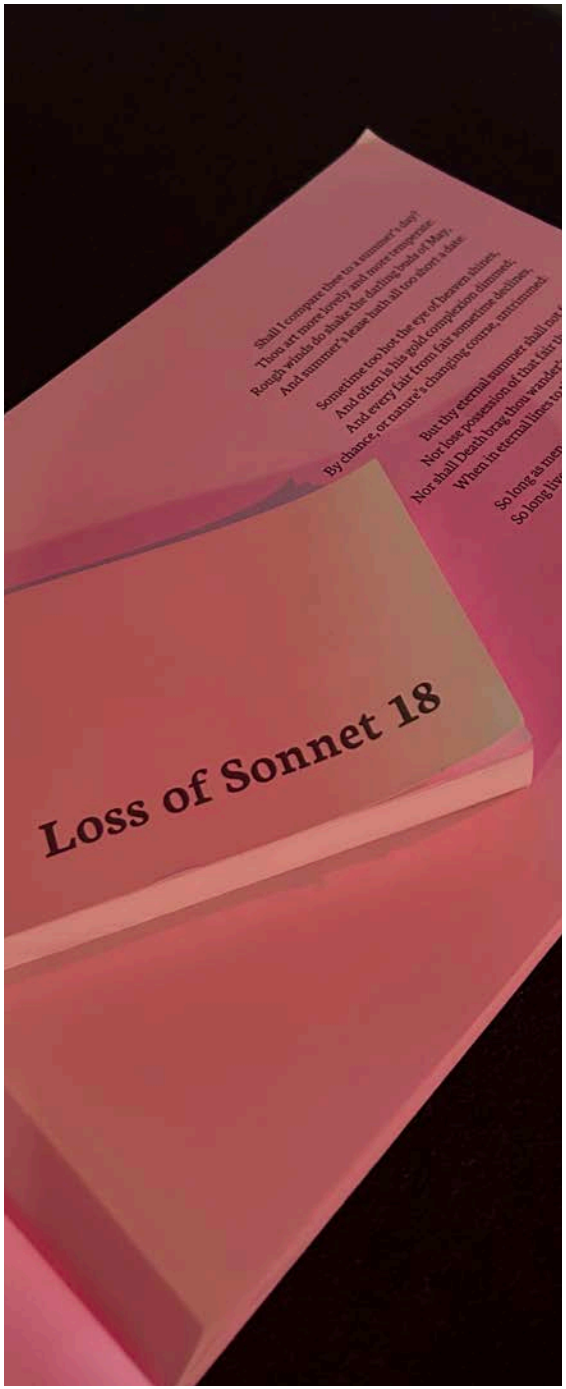
Website
https://osf.io/kjtrb/

Conceptual Foundation

At the heart of "Loss of *Sonnet 18*" lies an acknowledgment of the inherent loss in information storage and spreading, echoing through the corridors of human history and our methods of preserving and transmitting words. Although there is an inherent loss in information transfer, the thought behind information (i.e., the "word" in this project context) storage hardly changes.

Depicting the everlasting beauty, regardless of any loss, Shakespeare's *Sonnet 18* serves as our inspiration, a poetic testament to the fleeting nature of beauty and the supposed immortality bestowed through verse. Yet, this project asks a poignant question: **what happens when the medium of this immortality - the words themselves - undergo generational loss?**

We try to describe the loss of information through time and technique in an intuitive way, and invite the audience to feel both the changing and unchanging of thought behind different iterations of storage and spreading.



Inspirations

Drawing from the pioneering works of David Elliott (JPEG compression for 600 times) [1], Alvin Lucier (I'm Sitting in a Room) [2], and Dietmar Offenhuber (Autographic Visualization) [3], "Loss of *Sonnet 18*" represents a synthesis of art, data, and technology. Specifically, we focus on the generation loss of text data through image medium. We choose the last sentence of Shakespeare's *Sonnet 18*, *"So long as men can breathe or eyes can see, So long lives this, and this gives life to thee,"* as the main text content in our piece.

Shall I compare thee to a summer's day?
Thou art more lovely and more temperate:
Rough winds do shake the darling buds of May,
And summer's lease hath all too short a date:

Sometime too hot the eye of heaven shines,
And often is his gold complexion dimmed;
And every fair from fair sometime declines,
By chance, or nature's changing course, untrimmed:

But thy eternal summer shall not fade,
Nor lose possession of that fair thou ow'st;
Nor shall Death brag thou wander'st in his shade
When in eternal lines to time thou grow'st:

So long as men can breathe or eyes can see,
So long lives this, and this gives life to thee.

—William Shakespeare, Sonnet 18

Background

——About "Generation Loss"

In analogue recordings, a progressive loss of quality that occurs every time a tape, film, or vinyl disc is copied. The problem of generation loss has been reduced but not eradicated in digital media.

— *A Dictionary of Media and Communication (1 ed.), 2011*

Generation loss is the loss of quality between subsequent copies or transcodes of data.

— Wikipedia. Generation loss.
https://en.wikipedia.org/wiki/Generation_loss



4 MB



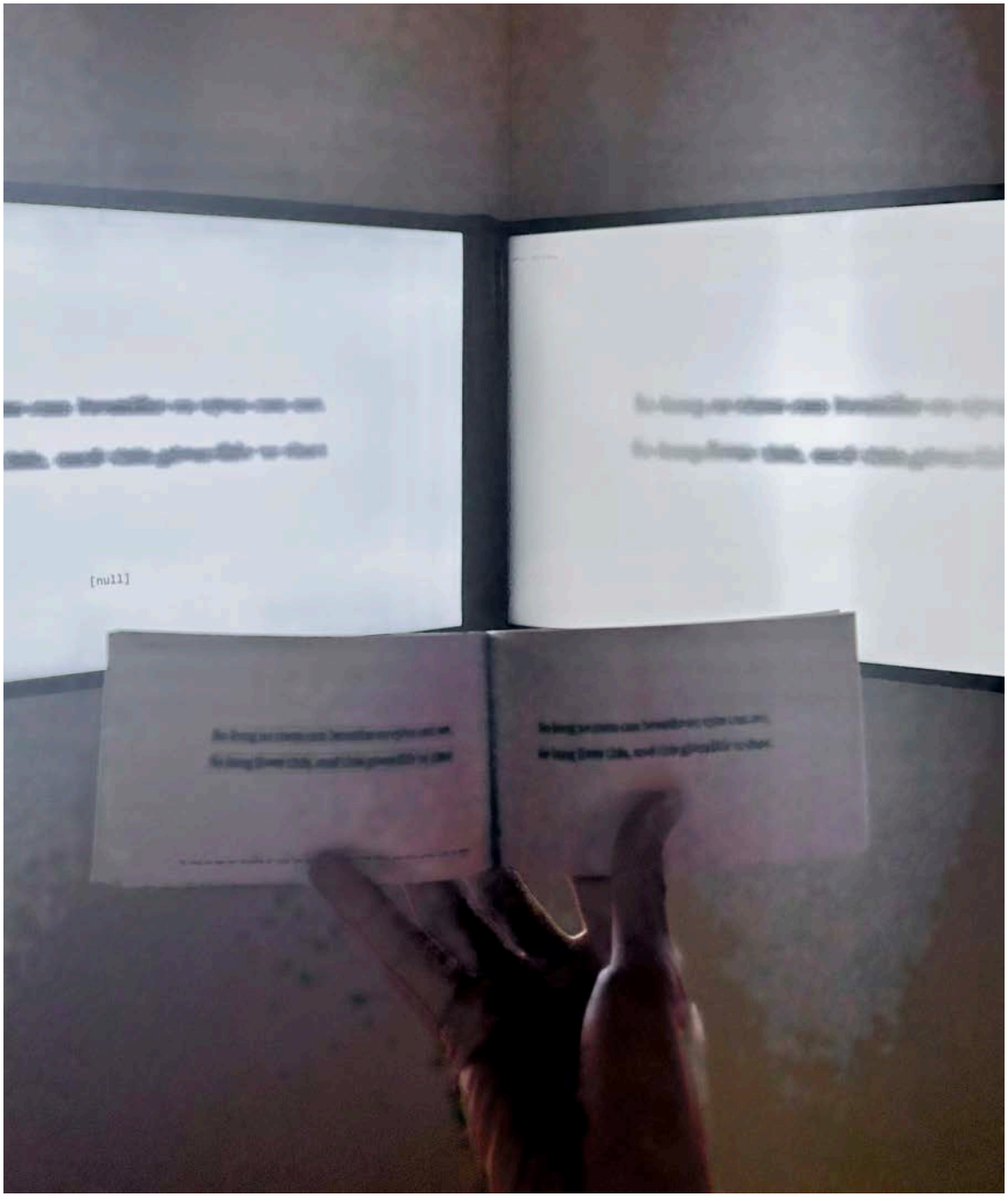
13 KB

The image becomes blurred after being shared many times on social media.
The file size can also become smaller, which means the information within the image decrease.

Artistic Interpretation

Through a blend of traditional poetry and modern technology, we simulate the blurring and compression of text, such as the distortion experienced in social media and digital communication. In a nod to the pre-digital era, we present these texts upon images, like the visual nature of historical mediums written expression – ink on paper, prints, etchings on stone, all of which can be considered as images. This deliberate juxtaposition illuminates the transformative journey of words and meanings, guiding the audience through an experiential narrative where the known gradually morphs into the enigmatic.

Through this artistic process, we invite the audiences to become witnesses to the subtle yet profound metamorphosis of language. They will observe familiar words and phrases slowly erode into realms of ambiguity and reinterpretation, and try to find a border of recognizable loss within this process.

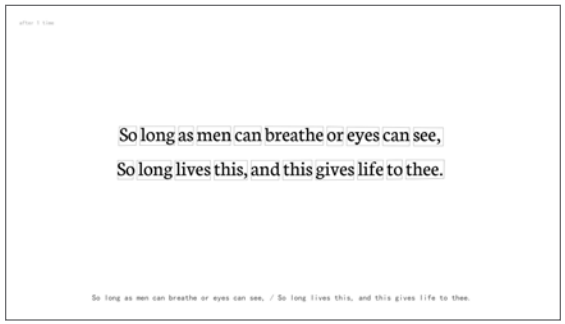


Technical Aspects

The major medium of "Loss of *Sonnet 18*" is a series of videos that record the loss of words being stored and spread. Each video within this artwork is a sequence of 999 images, arranged according to their respective compression iterations. In each iteration, we program to add different types of noises to deliberately degrade the image quality and infuse it with a touch of blur, simulating the gradual erosion of clarity and precision.

This iterative process is cumulative; the output of one iteration becomes the input for the next, creating a layered narrative of loss. Each subsequent image is not just a standalone representation but an aggregate of all the preceding levels of distortion and decay. This continuous progression encapsulates the core theme of our project - the relentless, incremental diminishment of data integrity, mirroring the inexorable passage of time and the transformative nature of memory and perception that might occur on words.

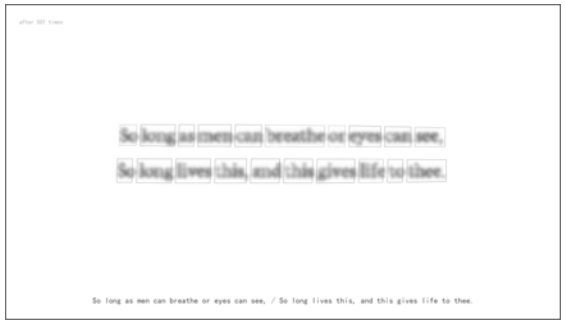
The background sounds in videos consist of two parts. One is the morse code of the poem *"So long as men can breathe or eyes can see, So long lives this, and this gives life to thee."* It is an echo of machine signals, in contrast to the sound in the audience's brain, i.e., the echo of human thought. The other is environmental music we subjectively select to match the main idea of this project. With loss deepening, the background sound also becomes distorted and noisy, generating a kind of sound and picture synesthesia feeling.







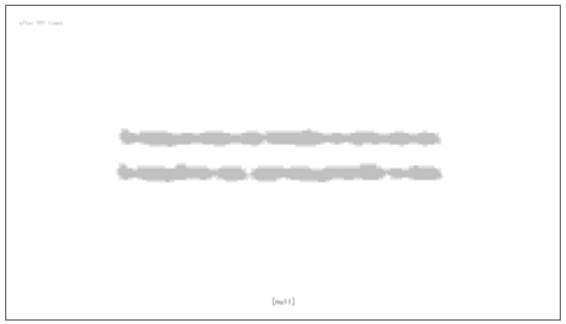






Implementation

We choose the JEPG compression with low quality and blur for the visual effect. We utilize Python with the Pillow library for image processing and Microsoft Azure services [4] for Optical Character Recognition (OCR) and computer vision. We render the video presentation through an HTML5 application.

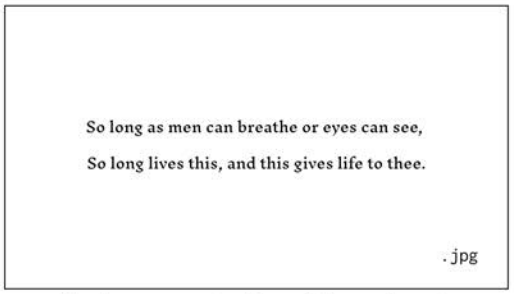


Loss of Sonnet 18

Major Pipeline

“So long as men can breathe or eyes can see,
So long lives this, and this gives life to thee.”

Two original poems
from Shakespeare's *Sonnet 18*.



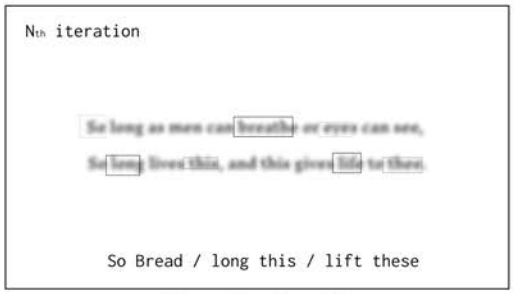
Words are presented through image format.



Image is compressed and blurred to simulate the loss process in communication iteratively. The iteration including 999 times.

“So Bread
long this lift these”

Recognizing words through
Optical Character Recognition Technique



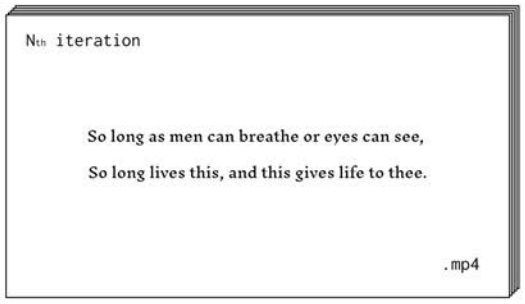
Viewed by AI model.

“[words observed]
and [meaning learned]”

Recognizing words through
reading and understanding.



Viewed by human audience.



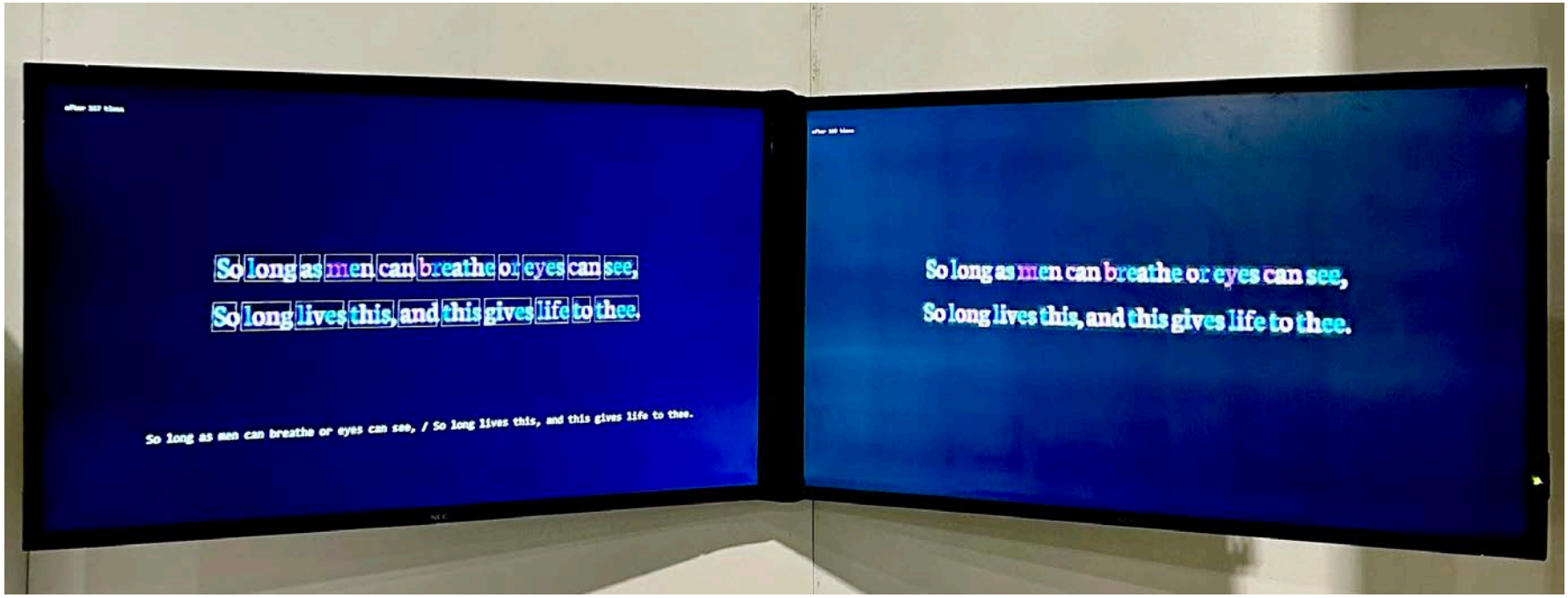
These imges sequentially compose a video, illustrating
the process of visually loss.

The vitality and beauty of youth,
the poem, and
other eternal things.

Loss of *Sonnet 18*

A gradual accumulation of
change or damage
compared to the
original data, form, or information.

The loss in our piece appears in different ways. Images become fuzzy when shared across media. The original meaning of orally transmitted poetry fades over time. The clarity of youthful beauty in timeless lines diminishes. All of them show that loss is ongoing, not just a single event. It leaves traces of the past and opens up new understandings.



Audience Engagement*

The installation features two screens and one web camera. The primary screen shows the iteration of text loss. The secondary screen presents Optical Character Recognition (OCR) results from AI, including word bounding boxes, to simulate the new understanding of text loss.

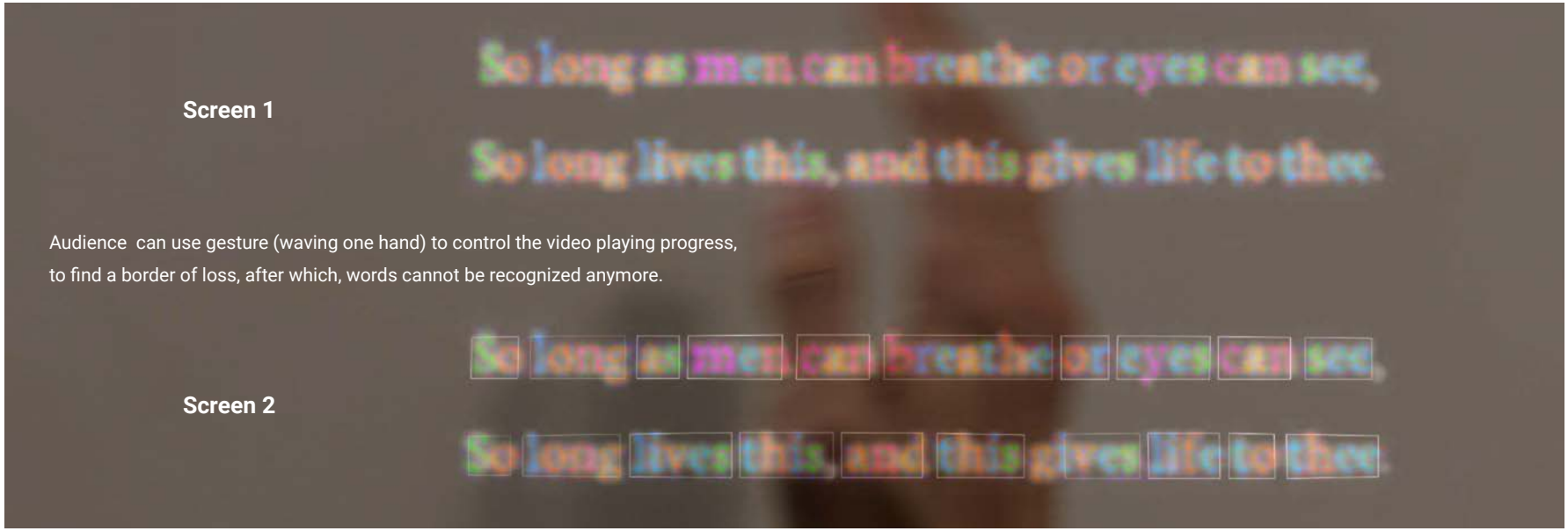
Except for OCR AI, the audience also becomes recognition agents, controlling the video's progression with gestures. The web camera is used to capture the number of audience and their gesture.

There are two playing modes in this installation. When there is no audience gesture detected by the web camera, the loss process will automatically play. When the audience's gesture is recognized, then the automated playing is interpreted and will be controlled by the audience's interactions. Only one audience's hand that is the earliest to wave and detected by the camera will be taken into account for interactions. By waving hands, the audience can control the loss process to a specific moment on the primary screen, and observe when words cannot be recognized anymore. The secondary screen will also be synchronized with such control, providing a comparison for the audience on the moment in human's and AI's perspectives.

They could find the frame after which they cannot recognize the text anymore, comparing it with what AI recognized at each frame. This interaction symbolizes the subjective nature of interpretation, as each person's understanding shapes their perception of loss. The number of audience dynamically influences video playback speed, reflecting the collective impact on interpretation and possible meaning deviation. The loss between individual thought and the original idea behind words introduces an accumulative impact on information storage and spreading. The more individual interpretations, the more possible information will be changed during transfer. Hence, the more loss. This setting aims to highlight the loss introduced by the diversity of interpretations among the audience.

In the secondary screen, the bounding box opacity in the results symbolizes AI's confidence in OCR results, a visual metaphor for the clarity of understanding AI.

* The version exhibited in SIGGRAPH Art Gallery 2024 had no interaction modules.



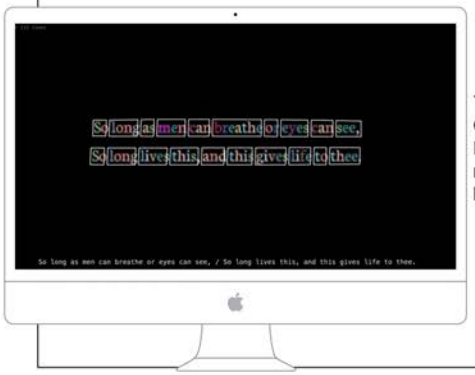
SCREEN 1



The primary display shows the iteration of word loss. Audience act as a terminal, observing and recognizing the lossy information of words during the process.

synchronous

SCREEN 2



The secondary screen presents OCR results, including word bounding boxes. It simulates the new understanding of the word loss, where AI acts as a terminal.

There are two playing modes in this installation.

Mode 1. When there is no audience's gesture detected by the webcamera, the loss process will automatically plays by our pregenerated results.

Interaction 1

When the audience's gesture is recognized, the automated playing is interpreted, and the loss process be controlled by audience's interactions.
Only one audience's hand that is the earliest to wave and detected by the camera will be taken into account for interactions.



Mode 2. By waving hands, audience can control the loss process to a specific moment on the primary screen, and observe when words cannot be recognized anymore. The secondary screen will aslo be syncronized with such control, providing a comparison for audience on the moment in human's and AI's perspectives.



Interaction 2

The number of audience dynamically influence video playback speed, reflecting the collective impact on interpretation and possible meaning deviation.



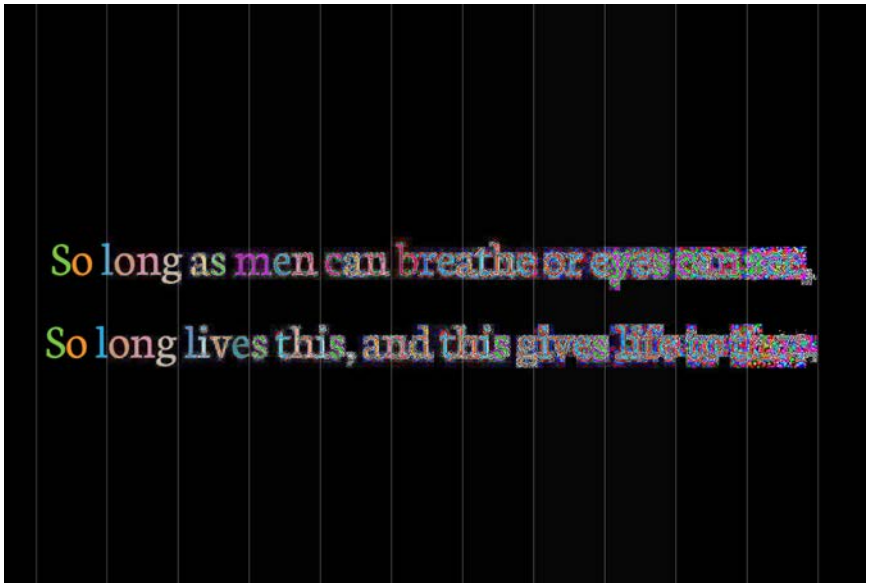
Slow Down

Speed Up





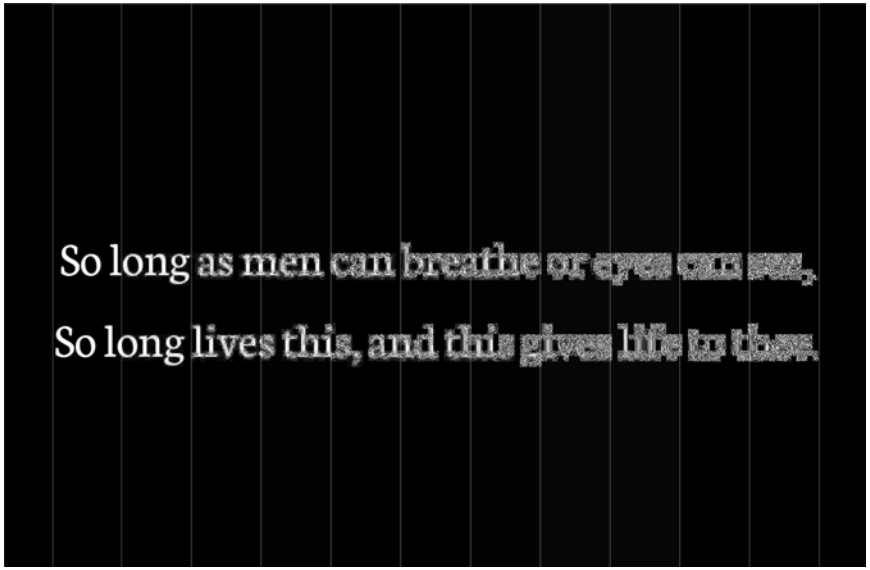
Black background + Colorful texts + Blur



Black background + Colorful texts + Noises



Black background + White texts + Blur



Black background + White texts + Noises

Philosophical Aspect

The Ship of Theseus, known as *Theseus's Paradox*, is about whether an object is the same object after having had all of its original components replaced over time, typically one after the other. The Athenians preserved the ship that Theseus used to escape, by replacing the parts one by one as they decayed. After several hundreds of years of maintenance, if each individual piece of *the Ship of Theseus* was replaced, one after the other, a question is proposed: Was the ship still the same as the original one?



Photo by Milind Kaduskar on Unsplash.

In this project, the information conveyed in the original poem is the ship, each time of interpretation is like replacing one part of the ship. After many times of communication, propagation, and interpretation, the same question appear:

Was the information still the original one?

If not, when did it change?

Everyone may have their answers while machine also have its own.

Similar things also happen in the field of artwork restoration. After restoration, is the painting still the original artist's work?

When we go back to the term behind the project title, *generation loss*, we think that the term can be took apart into two words. *Generation AND loss*, which means during information communication, there are not only information lost but also information generated. Both of the status compose the current information.

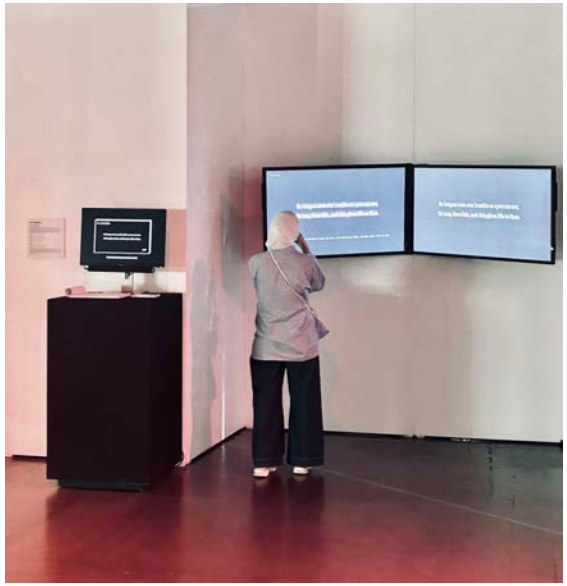




Leonardo da Vinci, *Salvator Mundi*, c.1500, oil on walnut Getty Images
Left: Original painting after cleaning. Right: Reproduction of the painting after restoration.
Public Domain. https://commons.wikimedia.org/w/index.php?curid=64103353







Disucssion

"Loss" is a general concept and phenomena, which exists everywhere. For a broader implications of our art piece, when paper act as the medium to hold the poem, there can be loss of paper when they have gotten mildew, been broken or folded. At the same time, some new visual information appears on the paper, along with the poem.

In our current artwork, we only explore the loss of text through image format but there can be many media for information saving, propagating, and receiving. The kind of loss can include physical, chemical or digital ones, and etc. We are working on a space of loss based on the dimension mentioned above, and expand it to more modality, like audio, image and video. This can be a very general problem with a wide range of exploration space.

Conclusion

This project invites reflection on the essence of communication in our digital age. It highlights the ongoing yet ever-changing nature of loss, leaving a lasting impression of fleeting beauty and the transient nature of understanding. As an art piece, it's an invitation to experience, ponder, and engage with the ever-shifting landscape of communication, technology, and human perception.

Reference

1. Scott Beale. 2009. Generation Loss, What JPEG Compression Looks Like After 600 Saves. https://laughingsquid.com/generation-loss-what-jpg-compression-looks-like-after-600-saves/
2. Martha Joseph. 2015. Collecting Alvin Lucier's I Am Sitting in a Room. https://www.moma.org/explore/inside_out/2015/01/20/collecting-alvin-luciers-i-am-sitting-in-a-room/
3. Dietmar Offenhuber. 2020. Data by Proxy — Material Traces as Autographic Visualizations. IEEE Transactions on Visualization and Computer Graphics, 26(1), 98-108. https://doi.org/10.1109/TVCG.2019.2934788
4. Azure products. Microsoft. https://azure.microsoft.com/en-us/products

Life Thread

Junxiu Tang, Nathalie Bressa, Elise Bonnail,
Xinhuan Shu, Yifang Wang, Yingcai Wu, and Petra Isenberg
2024

An ongoing project.

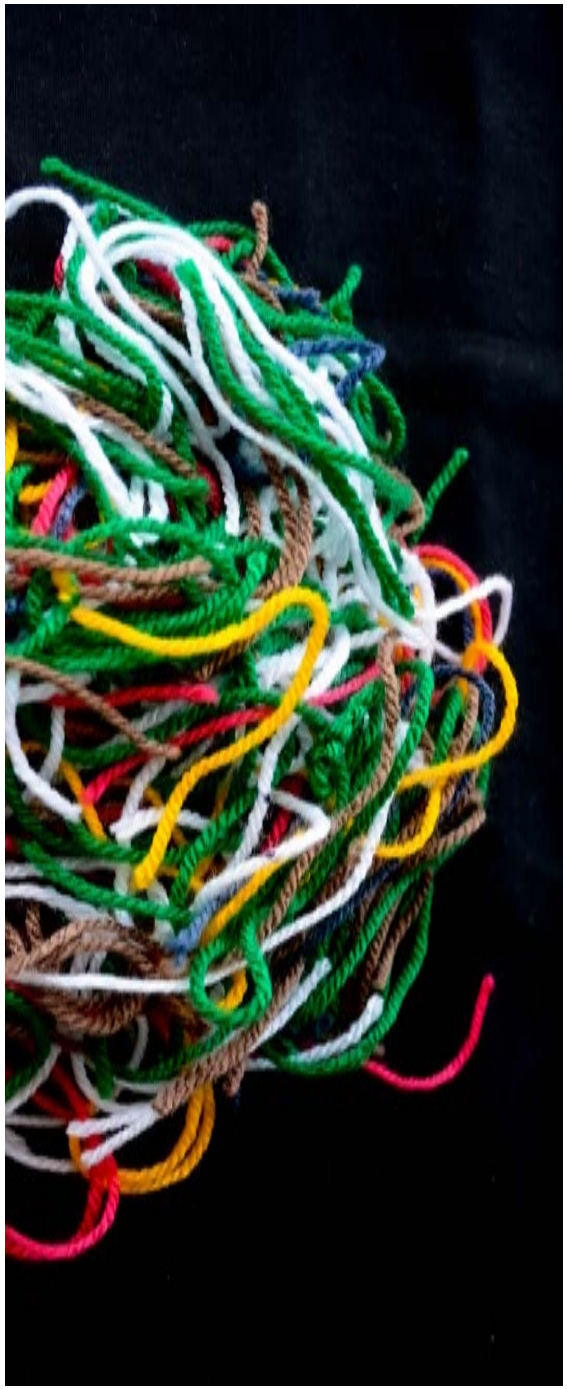


Abstract
The concept of a “Social Clock” represents a timeline of societal expectations for achieving life milestones at specific ages. This timeline, shaped by regional values, can put enormous pressure on individuals as they approach these ages. Despite the pressures of a social clock, individuals lead unique lives that deviate from societal norms. Our visualization piece contrasts personal life timelines with a social clock, highlighting these deviations. It invites audience participation, allowing them to share their own timelines and social clocks from their background, reflecting diverse life paths. This interaction emphasizes the diversity of social clocks and encourages reflection on the reasons behind different life choices. While many adhere to the general trend of a social clock, influenced by biological factors and societal pressures, it is essential to respect the diversity of life paths. Our work focuses on the diversity of life timelines, providing reference and support for those struggling with societal pressures, and serving as an archive to encourage others: weave your own life thread and be the master of your journey.

Mediums
Wool; Cloth; Input Visualization.

Keywords
Social Clock; Input Visualization; Data Physicalization.

Website
https://osf.io/p2fnu/



Introduction

Society often imposes an unspoken rule: one should achieve certain life milestones at a specific age. This concept, introduced by Neugarten et al. in the 1970s [1], is known as the **social clock**. It defines social timelines that mark life's milestones, such as education, employment, marriage, and childbearing. These timelines are shaped by regional social values and vary across genders, cultures, and identities. While these suggested timelines propose suitable ages for certain achievements, they often enforce rigid expectations, pressuring individuals who approach these milestone ages. When personal choices diverge from the scheduled timeline, it commonly leads to external observation and dialogue.

Despite societal pressures, individuals lead unique lives. Our visualization piece juxtaposes diverse personal life timelines against the traditional social clock. This contrast reveals how individual life paths deviate from societal norms in the duration and sequence of life stages, and that some stages may be omitted entirely. Judging one's life against the social clock is challenging because each person's journey is shaped by subjective choices or unavoidable circumstances.

Our piece emphasizes the diversity of an individual's life experiences. It serves as a participatory design, defined as "input visualization" recently [2], inviting the audience to contribute and display their own life timelines with the social clock. Participants can also share the specific social clocks from their cultural, national, religious, gender, and professional backgrounds. This interaction highlights the diversity of different social clocks and encourages the audience to reflect on the reasons behind their and others' life choices.

We might observe that while individual life paths diverge, the majority still follow a general trend, adhering to the social clock – a form of "fixed" diversity. This social clock can be influenced by biological factors, such as the healthiest age to have a child, and societal factors, including elder expectations and peer pressure. As societies become more inclusive of diverse personalities, races, and sexual orientations, it is equally important to respect the diversity of individual life paths.

Detailed observations of life paths within specific groups, such as by gender or country, can reveal unique patterns. However, conducting a comprehensive visual analysis of these patterns is beyond the scope of this piece. Instead, this piece serves to prompt the audience to reflect on the concept of the social clock. It encourages consideration of both personal experiences and those of others, examining conventional norms and the reasons behind individual life choices.

Our work anonymizes personal information, focusing solely on the diversity of life paths. It functions as a reference for younger people to choose life paths, a support for individuals struggling with the societal pressures of the social clock, and an archive to showcase and encourage others.



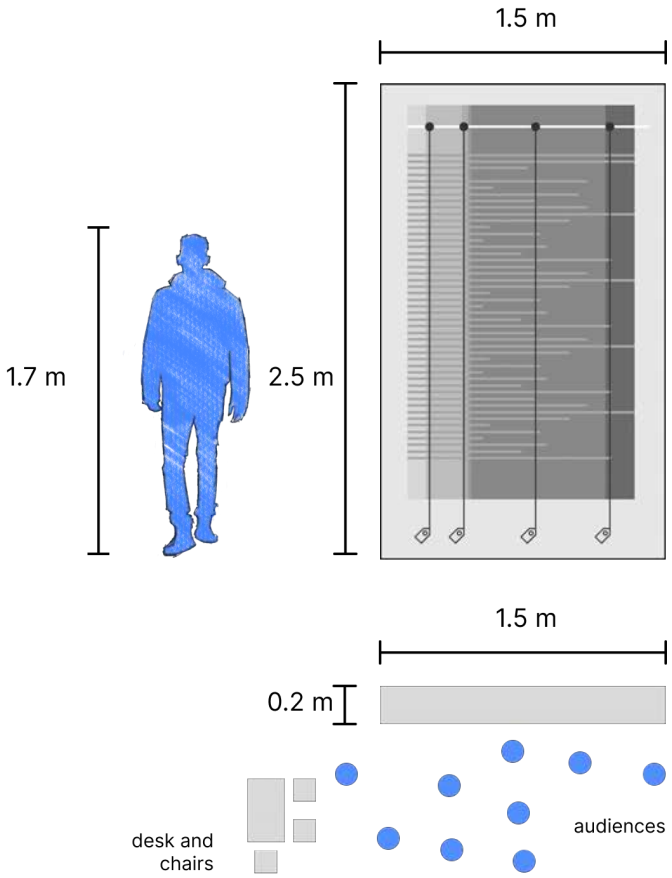


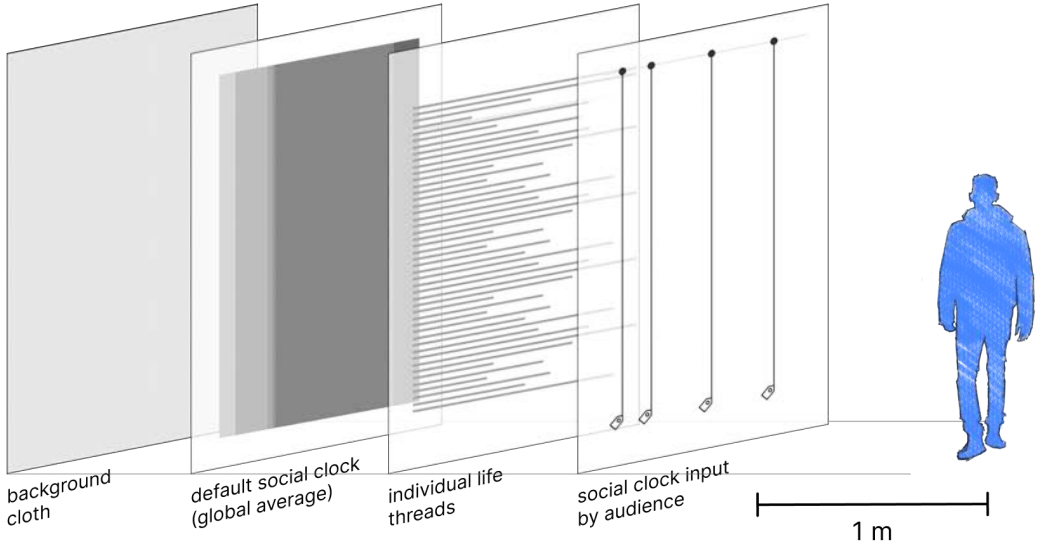


Reference

1. Bernice L. Neugarten and Nancy Datan. 1973. Sociological perspectives on the life cycle. Life-span Developmental Psychology, 53-69. Academic Press. https://doi.org/10.1016/B978-0-12-077150-9.50009-5
2. Nathalie Bressa, Jordan Louis, Wesley Willett, and Samuel Huron. 2024. Input Visualization: Collecting and Modifying Data with Visual Representations. CHI '24: Proceedings of the CHI Conference on Human Factors in Computing Systems. Article 499, 1–18. https://doi.org/10.1145/3613904.3642808

Setup Plan





* The background cloth and default social clock can be rolled up to ensure the scalability.

Junxiu Tang's Portfolio | Updated in Nov. 2025 | Copyright © Junxiu Tang



Background - The Basic Social Clock

We collected data from global sources such as the UNESCO Institute for Statistics (UIS), World Bank Data, and the World Health Organization (WHO). This background visualization represents the average age at which people globally start education, begin employment, retire, marry, and have their first child.

The average age cannot accurately reflect the social clock in each small society, as it is more of a statistical measure rather than a societal norm. To incorporate the diverse and unknown social clocks, we use the current data as an initial prompt and integrate the input from various social clocks (see **Labels on Thread** for more details) during exhibition.





Life Threads - Diverse Life Timelines

We visualized personal life timelines using colored wool. The metaphor is to "weave the life." The white wool represents the basic timeline of an individual's life, while the colored wool indicates specific life stages. For example, **green** could represent education, **brown** employment, **blue** retirement, **red** marriage, and **yellow** childbearing. The length of the wool corresponds to the duration of each life stage, and knots between two colors indicate milestones marking transitions between stages. Since some life stages, like marriage and raising children, may overlap with education and employment, we intertwine the corresponding wool strands to represent these concurrent experiences.

By displaying these life threads against the backdrop of the social clock, we illustrate the diversity and individuality of life paths. We invite audiences to participate in this piece, weaving their own life timelines (with our help) during the exhibition. Their narratives will blend into the input visualization, adding diversity and richness to the piece. The existing threads are made through crochet hooks and wool. The data are from the authors' families and friends, with different gender, countries, and ages.



Labels on Thread - Diverse Social Clocks

Recognizing that average ages might not accurately reflect specific social clocks, we provide an interactive element–a blank white thread–for the audience to share the social clocks they are faced with. This part of the installation allows the audience to express the societal timelines they navigate, offering an opportunity to convey personal and cultural narratives.

The exhibited piece could be largely influenced by where we plan to do the exhibition and who the audience is, as they will participate and are expected to resonate with it. As the audience of the exhibition is mostly researchers, artists, professors, and students, we will be happy to see the diverse life paths and social clocks shared and communicated among the people from this community.



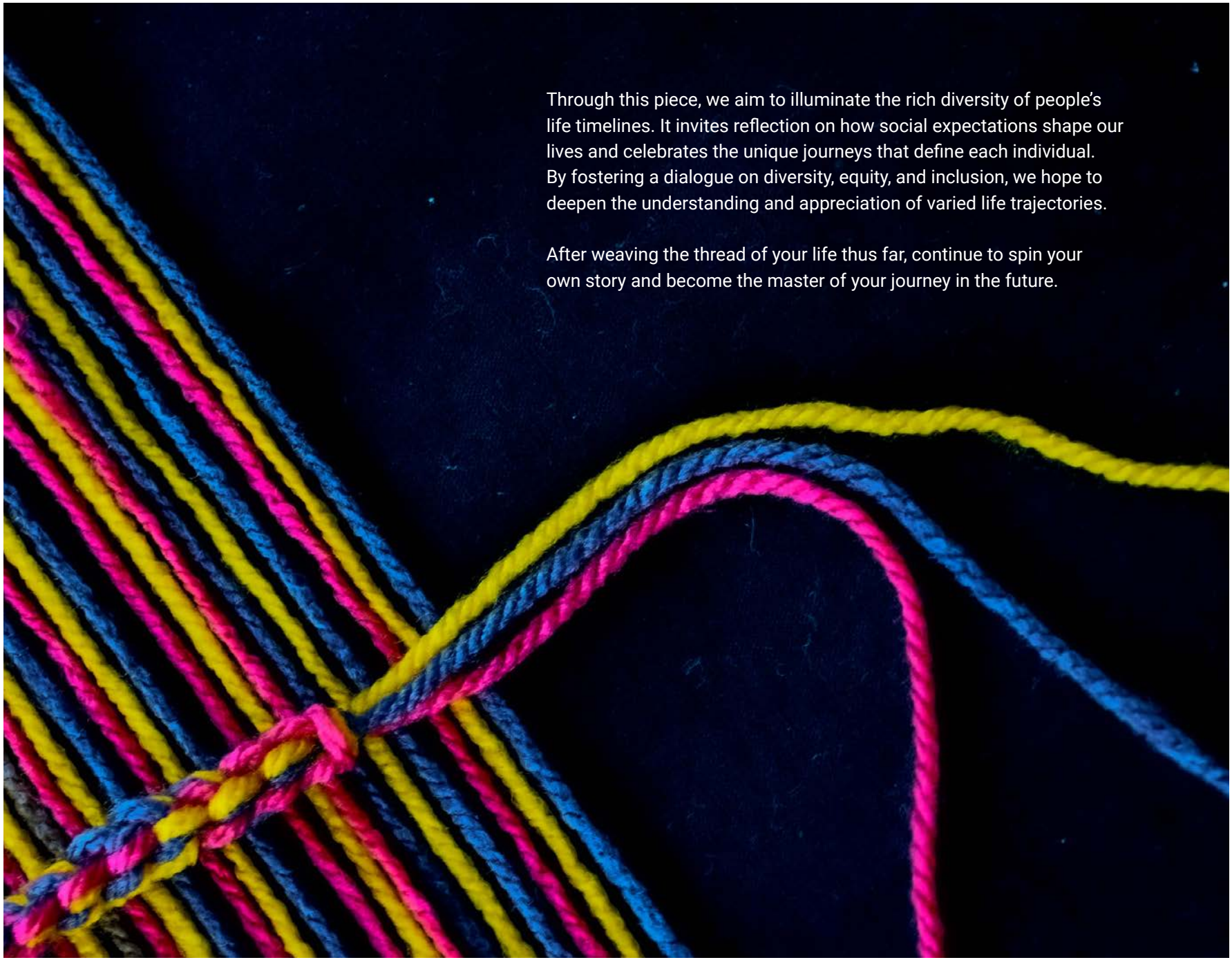
For some, significant life milestones—**graduation**, **employment**, **marriage**, and **having a child**—can all occur within a span of five years.



Some individuals start **working** before completing their **education**.



Some individuals choose not to **be married** or **have children**.



Through this piece, we aim to illuminate the rich diversity of people's life timelines. It invites reflection on how social expectations shape our lives and celebrates the unique journeys that define each individual. By fostering a dialogue on diversity, equity, and inclusion, we hope to deepen the understanding and appreciation of varied life trajectories.

After weaving the thread of your life thus far, continue to spin your own story and become the master of your journey in the future.

Explore Mindfulness Without Deflection: A Design Project Based on the Book of Songs

Yifang Wang, Yifan Cao, **Junxiu Tang**, Yang Wang, Huamin Qu, and Yingcai Wu
2021

IEEE VIS Arts Program. Virtual. Oct. 24-29, 2021.

Information is Beautiful Awards Long List. Online. 2022.

Electric Dreams - the 1st CMA International Symposium and Exhibition. Online. Aug. 25, 2022-Jun. 25, 2023.



Abstract

The Book of Songs is regarded as the origin of Chinese literature and has a prolonged impact on Chinese culture, aesthetics, and morality. In this work, we have analyzed the 305 poems in *The Book of Songs* from different dimensions. We aim to learn how various poetic imageries connect abstract themes and subjective emotions at the micro level, and how the poems connect people today and ancestors to understand the universal, everlasting, and poetical human lives at the macro level.

Mediums

Interactive webpage; Infographics.

Keywords

Digital Humanities; Cultural Heritage; Data Visualization.

Website

https://wangyifang.github.io/the-book-of-songs/



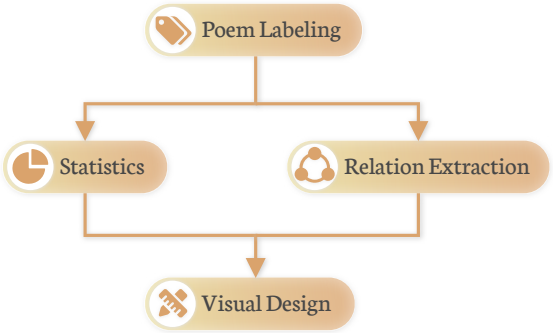
Introduction

The Book of Songs is widely acknowledged as the origin of traditional Chinese culture and highly impacted the whole East Asian culture. It is officially assembled by the famous founder of Confucianist, Confucius, who lived in the Late Zhou Dynasty around 3000 years ago. This epoch is considered paralleled with the Axial Period of world history, and Confucianist is regarded as one of the fundamental East Asian philosophies. *The Book of Songs* represents secular and religious values Confucius approved in accord with his own philosophical system, which brings a profound impact on traditional Chinese culture, aesthetics, and morality. Confucius highly admired *The Book of Songs* [2] , *"The Odes are 300; one expression sums up all: 'mindfulness without deflection.' "*.

Recorded from 1100 BC to 500 BC, *The Book of Songs* consists of 305 poems with three genres, namely, *Songs*, *Odes and Epics*, and *Hymns* [14]. Across different genres, the poems cover various themes to portray the whole society in the Late Zhou Dynasty, including not only ballads, folk-customs of ordinary people, but also sacrifice, banquet, and the musical dance of nobility [1].

Although many phenomena mentioned in the book (e.g., feudalism) are antiquated to Chinese people nowadays, many language styles and underlying thoughts are substantially the same as those we are experiencing today [1], especially the unique euphemistic style lying in the Chinese culture. *Such euphemistic style usually* takes advantage of natural creatures or phenomena to make comparisons or uses them as metaphors to express the feelings of the authors. To express the euphemism, people use tons of poetic *imageries* in *The Book of Songs*. *Herbage*, *Trees*, *Birds*, *Livestock*, *Insects*, and *Fish* have constituted a huge *imagery* database [15] shared by not only Chinese but the whole East Asian cultural circle. They act as essential bridges connecting the abstract themes and the subjective emotions and feelings expressed by the authors.

Savored today, the charm of *The Book of Songs* is still vivid. It represents the circumstances, the thoughts, the habits, the joys and sorrows of individuals from all classes of society in Ancient China. *The Book of Songs* has established a time tunnel to connect people today and their ancestors, which helps us understand the distant and precious cultural heritage and the universal, everlasting, and poetical human lives.



The analytical and design pipeline of the project.

Workflow

To systematically understand *The Book of Songs*, we used the original 305 poems to conduct the analysis and the visual designs. First, we worked with experts in Chinese Language and Literature and labeled these poems in four dimensions: *Functions*, *Themes*, *Emotions*, and *Rhetorical Devices*. In each dimension, the poems were classified into different categories (please refer to the supplementary material for more details). We cross-validated the labels with several authoritative studies related to *The Book of Songs*, including the *Anthology of Poems* [8], *Animism* [15], *Hand-drawn Book of Songs* [9], *Illustrated Handbook of Plants The Book of Songs* [10], and *Descriptive Text of Creatures from The Book of Songs* [11]. In addition to the poems, we also adopted aesthetic paintings drawn by Hosoi Jun in Qing Dynasty [12]. His works also inspired the color selection and designs of our pictorial.

Next, we programmed to obtain statistics of the above dimensions and the relationships among different *imageries*. We further designed and revised the visual representations iteratively and fixed the final version of the work.

Design Process

The pictorial contains two parts, showing the idea of connections in *The Book of Songs* from the micro and macro levels.

At the micro level, we selected the most frequently used poetic *imageries* from six categories (i.e., *Herbage*, *Trees*, *Birds*, *Livestock*, *Insects*, and *Fish*) as representatives and applied an *Imagery Glyph* design to show their importance in terms of connecting the abstract themes and subjective emotions.

At the macro level, a *Poem Flow* exhibits the relationship between *Genres*, *Functions*, *Themes*, *Imageries*, *Emotions*, and *Rhetorical Devices* of all the poems, which takes the form of a traditional Chinese arch bridge (with its inverted reflection in the water) in shape. We hope to use these traditional elements to illustrate how *The Book of Songs* connects modern people with their ancestors through universal and everlasting *imageries*, *Themes*, and *Emotions*.



The *Poem Flow* of all 305 poems in *The Book of Songs*. Dimensions include *Genres*, *Functions*, *Themes*, *Imageries*, *Emotions*, and *Rhetorical Devices* (from left to right).



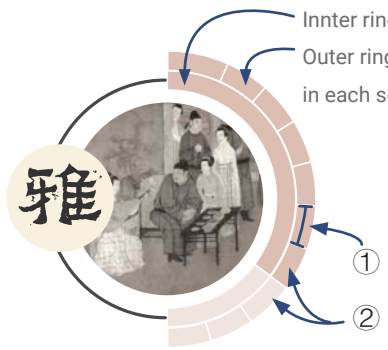
A design project based on *The Book of Songs*.

The *Imagery Glyph* for the top frequently appeared *imageries* in *The Book of Songs*. From left to right, these glyphs represent: Proso Millet, Mulberry, Swallow, Horse, Pest, and Black Bream.

Poem Flow -- Visual Encoding

The *Poem Flow* provides an overview with an enhanced Sankey diagram that presents six dimensions (i.e., *Genres*, *Functions*, *Themes*, *Imageries*, *Emotions*, and *Rhetorical Devices*) of all the 305 poems. We use different node groups in the *Poem Flow* to represent the above dimensions of the poems. The height of each node (except the *Genres* dimension) encodes the number of poems in one category under a specific dimension. The width of a flow between two nodes in two different node groups represents the number of poems belonging to these two node categories. The colors of the nodes and flows follow the style of the paintings drawn by Hosoi Jun in the Qing dynasty, which makes our pictorial similar to the ancient Chinese paintings.

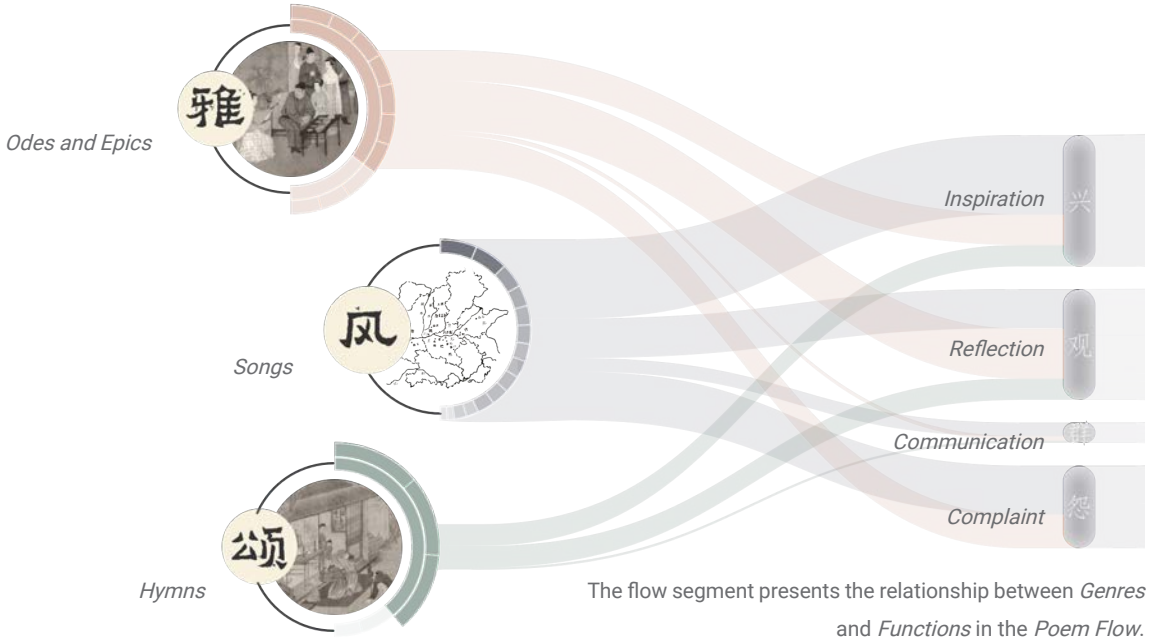
Three sunburst diagrams[1] on the left of the *Poem Flow* present the detailed classification of poems in the three genres (i.e., *Songs*, *Odes and Epics*, and *Hymns*), respectively. In each sunburst diagram, the length and opacity of each arc double encode the number of poems in different categories for a specific dimension. In the *Songs* node, there are fifteen categories representing poems from fifteen ducal states in the Late Zhou Dynasty. In the *Odes and Epics* node and the *Hymns* node, the inner arcs represent the section categories, and the outer arcs represent the related subsections within each section. The above *Poem Flow* segment presents the relationship between the *Genres* (i.e., *Songs*, *Odes and Epics*, and *Hymns*) and the *Functions* (i.e., *Inspiration*, *Reflection*, *Communication*, and *Complaint*) [2] of poems. Each flow represents a group of poems that belongs to a specific genre and a specific function.



Sunburst Structure

Use both the arc length (①) and opacity of each arc (②) to double encodes the number of poems in different categories

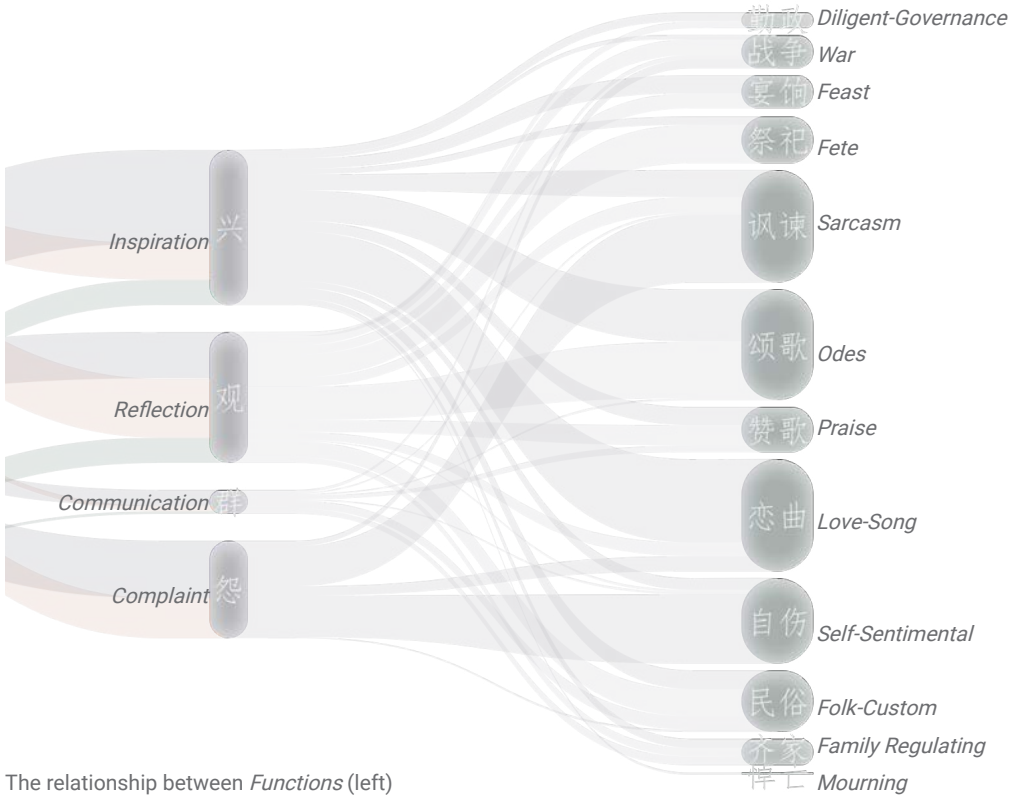
[1] The images used in this flow segment is from [5,6,7].



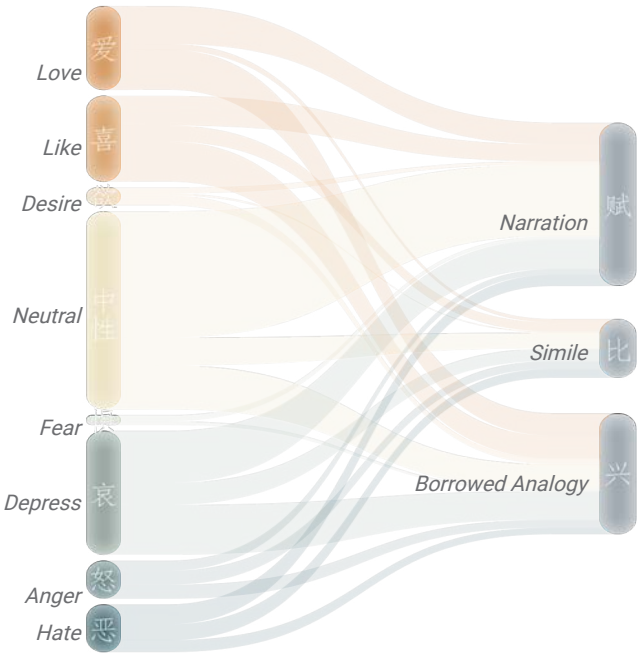
The flow segment presents the relationship between *Genres* and *Functions* in the *Poem Flow*.

Poem Flow -- Findings

As illustrated in the diagram, the *Songs* is the genre with most poems, which records stories and lyricism among plebeians. This may be because the colorful life of the plebeians gave birth to a large number of ballads and poems. These songs and poems are usually started with natural environments or domesticated animals to express emotions since the ancient Chinese people are more euphemistic in expressing their feelings. Such a unique euphemistic style is the reason why a lot of poems in the *Songs* flow to the *Inspiration* function. Another flow from the *Songs* to the *Complaint* also accounts for a nonnegligible proportion because the plebeians may grumble to the ruler. However, most poems in the *Reflection* are from *Odes and Epics*, and *Hymns*. Poems in these two genres record the dynasty history or ritual activities, which are mostly objective.



The relationship between *Functions* (left) and *Themes* (right) in the *Poem Flow*.



The relationship between *Emotions* (left) and *Rhetorical Devices* (right) in the *Poem Flow*.

We take the top three frequent themes, namely, *Love-Song*, *Sarcasm*, and *Odes*, as examples to introduce the relationship between the *Functions* and the *Themes*. First, *Love-Song* appears most for two reasons. On the one hand, love is an eternal theme in literature. Many poems in the *Songs* depict the stories of young people falling in love. Sometimes they are impressed by the scenery, so they expressed their feelings through these objects (*Inspiration*). On the other hand, though several poems are ostensibly about love, the authors use love as a metaphor to convey other ideas, such as the relationship between the monarch and his subjects (i.e., *Inspiration*). Second, besides literature, *The Book of Songs* also acts as moral education materials that Confucius collected for expostulating rulers. Among the poems related to rulers, some criticize the rulers (i.e., *Sarcasm*), while others celebrate them (i.e., *Odes*).

The right flow segment shows that the emotion distribution is balanced regarding the positive and negative sides. The most are neutral emotions, which take a considerable proportion of the *Narration*. However, in the *Simile* and the *Borrowed Analogy*, the emotions are more diverse since people tend to express love or satire euphemistically to avoid offending others.

We use a word cloud to present frequently used *imageries* of plants and animals. The shape is generated from the traditional Chinese arch bridge. The size of each Chinese character encodes the number of poems the *imagery* appears in *The Book of Songs*.

One of the frequently applied plant *imagery* is the Ziziphus Jujuba. It was everywhere in the agricultural age in China: the leaves were picked for teas and the timber was used for small pieces. Select grade Jujube timber were good materials to make traditional Asian instrument components, such as fingerboard and soundposts [4].

The most frequently used animal *imagery* is the Horse. Keeping the Horse in captivity is a conventional practice of ancient Chinese since the Horse is an essential type of animals in scenarios like hunting, sacrifice, and transportation.



The traditional Chinese bridge.
Photo by Charriot ZHAI on Unsplash[1].

One of the frequently applied plant imagery is the Ziziphus Jujuba.



The most frequently used animal *imagery* is the Horse.

Imagery Glyph -- Visual Encoding

The *Imagery Glyph* summarizes the poem information of the most frequently appeared *imageries* among six categories (i.e., *Herbage*, *Trees*, *Birds*, *Livestock*, *Insects*, and *Fish*). It consists of three parts: the circular icicle plot, the picture of the *imagery*, and concurrent *imagery* bubbles.

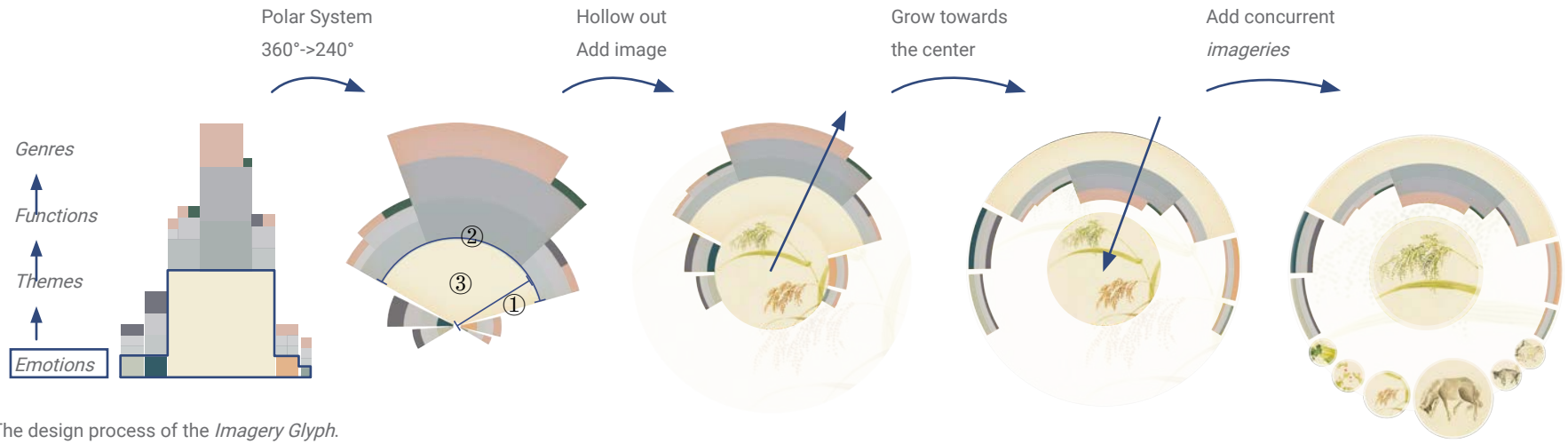
We aimed to summarize all the poems that apply a specific *imagery* to have an overall understanding of that *imagery* from different dimensions (i.e., *Emotions*, *Themes*, *Functions*, and *Genres*). We got inspiration from the icicle plot [3] to present the poem distribution of four dimensions in a hierarchical way. First, we summarized the emotion distribution. In the bottom layer, we used (1) the height, (2) the width, and (3) the fill opacity of the rectangle to triply encode the number of poems in each emotion type to magnify the differences. Following the same encoding strategy, we then classified poems of the same emotion type into different themes, as seen in the second bottom layer. We further summarized the function and genre distributions above the themes, which finally formed into an icicle plot with four hierarchies.

Second, to improve aesthetics and incorporate more information of the *imagery*, we distributed the original icicle plot along the polar coordinate. This circular design also expresses the idea of complete and perfection in the traditional Chinese culture [13]. We only made use of 2/3 of the circle to leave spaces for other information (e.g., concurrent *imageries*).

Third, we further refined the design by hollowing out a circle in the center to hold the picture of the *imagery*. We also provided a larger translucent picture to cover the whole graph.

However, as we can imagine, if we set the same radius for the inner rounded picture in each glyph, the sizes of these glyphs could vary largely due to the different distribution of poems (since the height of each outer arc also encodes the number of pomes). Thus, instead of laying away from the centre of the circle, we laid out the circular icicle plot towards the centre. This could make the sizes of glyphs uniform, easy to compare, and more aesthetic.

Finally, we filled the left 1/3 circle with pictures of the top three concurrent plant and animal *imageries*. The size of the picture encodes the number of poems where this *imagery* has appeared together with the target *imagery*.



The design process of the *Imagery Glyph*.

Proso Millet / *Herbage*

Mulberry / *Tree*

Swallow / *Bird*

Horse / *Livestock*

Pest / *Insects*

Black Bream / *Fish*





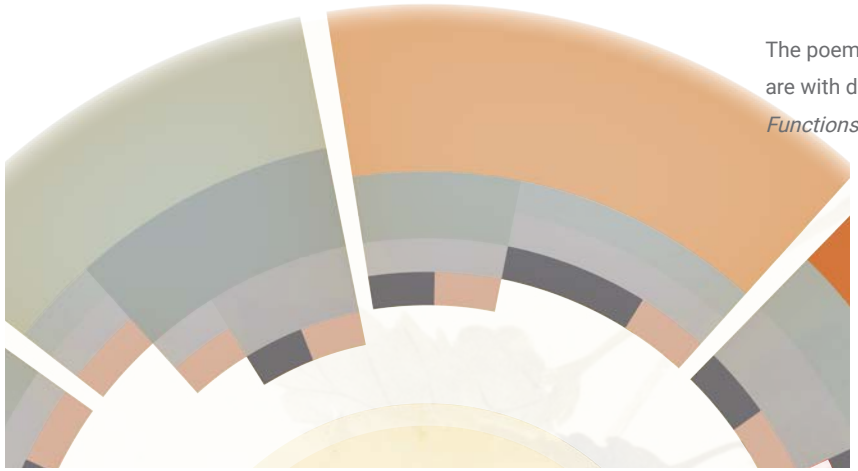
燕燕于飛，
O the swallows onward flying,
差池其羽！
Wings aslant, irregular!

——《國風・邶風・燕燕》
——*Friends in Distress, the Songs of P'EI, Songs*

Imagery Glyph -- Findings

The Swallow is the bird that appears the most frequently in *The Book of Songs*. Besides neutral, it tends to represent more positive emotions (in orange). For the genres, the Swallow appears in the *Odes and Epics* (in pink) most. The Swallow represents auspiciousness and reunion in traditional Chinese culture, which is always applied in the scenes of feast and sacrifice, particularly for nobiliary class. The Broomcorn and Proso Millet always appear together with the Swallow, with the potential to express the joy of harvest and the happiness of being together:

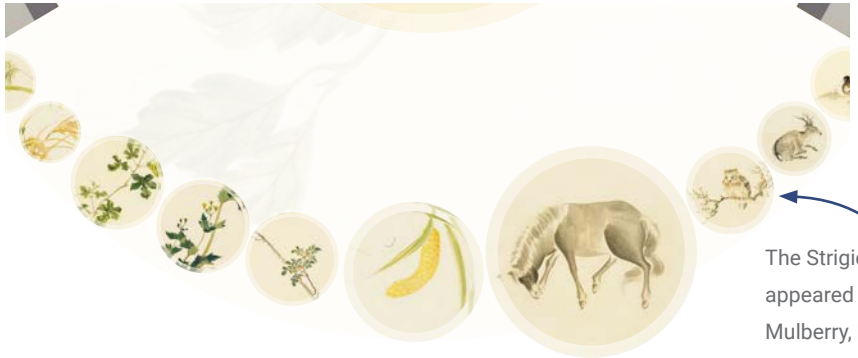
O the swallows onward flying,
Wings aslant, irregular [1]*!*



The poems that involve the Mulberry are with diverse *Emotions*, *Topics*, *Functions*, and *Genres*.



The Mulberry is one of the most important types of trees in Central and North China. The trunks, branches, bark, fruits, leaves, and velamen can all be used for different purposes.



The Strigidae is the frequently appeared animal together with the Mulberry, besides the Horse.

尸鳩在桑，
There in the mulberry-tree the dove,
其子七兮。
Sits on, seven young ones at her side.

——《國風 · 曹風 · 鳲鳩》
——*Praise of An Excellent Ruler, the Songs of TS'ÂU, Songs*

The Mulberry is the most-mentioned tree. The emotions of poems involving the Mulberry are diverse compared with other *imageries*. It may be related to the cultivation culture in the Late Zhou Dynasty when the Mulberry seems to be everywhere: the gain of the Mulberry is a type of tasty fruit, the leaf of the Mulberry is a type of delicious food for the silkworm, the Mulberry trees are great raw materials for houses. Thus, people would always use this common type of tree as a metaphor to express their feelings. Among the animals mentioned together with the Mulberry, the Strigidae appears frequently. This may be due to that the Strigidae nesting on the Mulberry:

There in the mulberry-tree the dove,
Sits on, seven young ones at her side [1].

Discussion

As the pioneer of the poetry anthology, *The Book of Songs* doubtless sets the foundation of Chinese culture. It has a prolonged influence on Chinese literature, serving like a cultural tunnel connecting current Chinese people with their ancestors. The heavy use of poetic *imageries* of *Herbage*, *Trees*, *Birds*, and *Livestock, Insects*, *Fish* successfully connects the intangible themes with subjective emotions.

In this project, we use data visualization to portray the inherent connection nature of *The Book of Songs*. A Sankey diagram deforming into a traditional Chinese arch bridge provides an overview of 305 poems. We also distill the most frequently appeared *imageries* in each category of *Herbage*, *Trees*, *Birds*, and *Livestock, Insects*, *Fish* to show how they assist in conveying the emotions and themes.

The impact of *The Book of Songs* is noticeable. *"I may assure the reader,"* Professor von der Gabelentz of Leipsic in a discussion on these poems remarked, *"in the whole collection of Hymns, and indeed in the whole canonic and classical literature of the Chinese, there is not a line to be found which might not be read aloud without any hesitation in the most prudish society. I know no other literature, of the East or West, on which similar praise could be bestowed"* [1].

Moreover, the cultural heritage of *The Book of Songs* is inherited by poetry, prose, and lyrics from Tang Dynasty to Qing Dynasty. Confucius also highly admired the spirits and rhetoric within *The Book of Songs*. In the Analects [2], he once commented that *"The Odes are 300; one expression sums up all: 'mindfulness without deflection.' "*. The voluminous collections of historical materials such as ancient books are precious resources to connect people today to the past to learn from history. Confronted with data of enormous amount, we can utilize both big data analysis and the aesthetics of arts to facilitate education. We hope that our initial experiment on the visualization of *The Book of Songs* can encourage more trials in digital humanities.

References

1. Confucius. 1891. The Shi King, the Old "Poetry Classic" of the Chinese. (A Close Metrical Translation, with Annotations by William Jennings). London: George Routledge and Sons.
2. Yang Xu, Yuanchong Xu, Bojun Yang. 2012. Thus spoke the master (Di 1 ban.. ed., Classical Chinese poetry and prose). Beijing Shi: China Intercontinental Press : Zhonghua Book Company.
3. Kruskal, Joseph B., and James M. Landwehr. 1983. Icicle plots: Better displays for hierarchical clustering. The American Statistician 37.2 : 162-168.
4. Wikipedia. Jujube. Retrieved August 8, 2021 from https://en.wikipedia.org/wiki/Jujube#Cultivation_and_uses.
5. 周啸天. 1988.十五国风地图:《诗经楚辞鉴赏辞典》附录. Retrieved August 8, 2021 from http://cnki55.sris.com.tw/refbook/R200608074.html.
6. 顾闳中. 五代. 韩熙载夜宴图卷. Retrieved August 8, 2021 from https://www.dpm.org.cn/collection/paint/228200.html.
7. 楼璹. 南宋. 耕织图. Retrieved August 8, 2021 from https://www.wdl.org/zh/item/289/view/1/30/.
8. 朱熹. 1961. 诗集传 (香港第1 版. ed.). 香港. 中华书局.
9. 徐峙立. 2019. 诗经手绘图谱. 济南. 山东画报出版社.
10. 潘富俊&吕胜由. 2001. 诗经植物图鑑 (第1版. ed.). 台北. 猫头鷹出版社.
11. 高明乾, 王凤产, & 毛雪飞. 2020. 诗经动植物图说 (第1版. ed.). 北京. 中华书局.
12. 细井徇&细井东阳. 1848. 诗经名物图解.
13. 段玉裁&许愼. 1981. 说文解字注 (第1版. ed.). 上海. 上海古籍出版社 : 新华书店上海发行所发行.
14. 许渊冲. 2005. Selections from the book of poetry (Illustrated edition) - 精选诗经与诗意画 (英汉对照) (第1版. ed.). 北京. 五洲传播出版社.
15. 细井徇&白云出岫. 2018. 万物有灵:《诗经》裡的草木鸟兽鱼虫 (第1版. ed.). 新竹县竹北市. 品林数位文化出版 : 如是文化发行.