VOL 3
AUTUMN/WINTER
2021

Design Behaviors

INTERNATIONAL DESIGN RESEARCH JOURNAL

UDA THE GLOBAL DESIGN ORGANIZATION
United Designs Alliance FIRMITAS ▪ UTILITAS ▪ VENUSTAS

Big Data Visualization of COVID-19 using Augmented Reality

Young Ae Kim, MFA

Associate Professor
South Dakota State University, USA

Qiuwen Li, MFA

Assistant Professor
Santa Clara University, USA

ABSTRACT

The Coronavirus disease 2019 (COVID-19) pandemic has led to unprecedented disruption of global daily life. The United States of America is a hotspot for the COVID-19 and leads the world in confirmed cases and deaths. The global impact of coronavirus is an extremely overwhelming experience for everyone, despite the rapid vaccine development and inoculation. A vast quantity of COVID-19 data has been produced and gathered from a variety of data sources.

Careful observations and implementations of the epidemiological big data is valuable for the broad audience and inspires the innovation of identifying, controlling, combating, and preventing the virus. Despite the invaluable function of data, it is very challenging for the public to understand it because it is highly abstract and noninteractive for everyday people. In fact, when it comes to big data, it is even harder to understand and relate the data at the individual level. We argue that the majority of current COVID-19 data dashboards may be useful for researchers, policy makers, and healthcare providers; however, it failed to deliver emotionally meaningful conversation at an individual level to encourage growth and openness to change for the public.

The study takes the visualization process with big data to create visual elements (shapes and symbols) and properties (size, color, position, etc.) that represent the meaning of the big data and encourages the public to engage data with minimal knowledge of statistics. It makes the data comprehensive for the public and easily accessible in order to provide the pattern of data effortlessly. In addition, the AR experience with data visualization introduces a new experience of how to interact with big data in order to increase the user's ability to intuit meaning and relationships from the high dimensional data.

The visualization was curated into a pictorial format that allows vast amounts of epidemiological data immediately using preattentive visual properties and popular visualization techniques. A combination of a qualitative palette and preattentive visual properties encouraged to identify a summary of unseen patterns, revealing insights, discovering areas of needs, finding errors in the current system, and directions in the future. The systematic modular grid provided flexibility to fit the data visualization with various types of data sets.The data visualization of COVID-19 using AR provides users meaningful *interaction and better understanding of confirmed cases, deaths, social distancing, unemployment cases, vaccination rates, and economic impact of COVID-19 epidemiological data. This study is designed for visualization of epidemiological data, comprehensive and interactive for the public and easily accessible in order to provide the pattern of data effortlessly. The technology enhancements will make augmented reality and virtual reality experience feel much more real, in-person, authentic, and effective, which will allow the public to live smarter, safer, and productive lives. This study provides inspiring emotional experience and encouragement to be empathetic for the public while exploring COVID-19 epidemiological data.*

KEY WORDS: Big Data, Data Visualization, Augmented Reality (AR), COVID-19, Epidemics, Time Series Analysis, Preattentive Visual Property

INTRODUCTION

There are more than 219 million confirmed cases and more than 4.5 million people have died globally of COVID-19 since the start day of the pandemic (WHO, 2021). The coronavirus pandemic is an ongoing public health emergency and an economic crisis, unprecedented in the disruption of daily life. The pandemic has created the "New Normal" (Anderson & Vogels, 2021). The pandemic has caused the largest critical global social and economic disruption since the Great Depression of the 1930s (Roberts, 2020). It has elevated concerns of discrimination issues and the balance between human rights and public health imperatives. WHO and CDC recommend continued social distancing, frequent hand washing, wearing a mask, and keeping hands away from the face, due to the rapid rate of spontaneous mutation of SARS-CoV-2, despite the availability of COVID-19 vaccine. The spreading of the disease in different places and types of SARS-CoV-2 variant demands active solutions to control virus progressions. A massive amount of COVID-19 epidemiological data has been pouring to the public daily since the pandemic began. The development of dashboards to present the COVID-19 epidemiological data in real-time has provided the spatial distribution of the virus, which was accessible to people instantaneously. Many data visualizers and dashboards have been developed in the past year to help the public understand COVID-19 numbers. Interactive dashboards with a variety of charts demonstrated concise ways to present the pandemic growth. John Hopkins University (JHU) was the first to display a COVID-19 dashboard of cases and death totals from worldwide as well as the United States data (see Figure 1). It included a map composed of diverse radius red circles representing the concerns of pandemic around the world, list of counts, and histograms. The New York Times (NYT) used choropleth maps for indicating geographical regions (states) and colors for representing measurement (number of cases), which help to observe data trends comprehensively.

Figure 1. COVID-19 dashboards and interactive tool by Johns Hopkins University

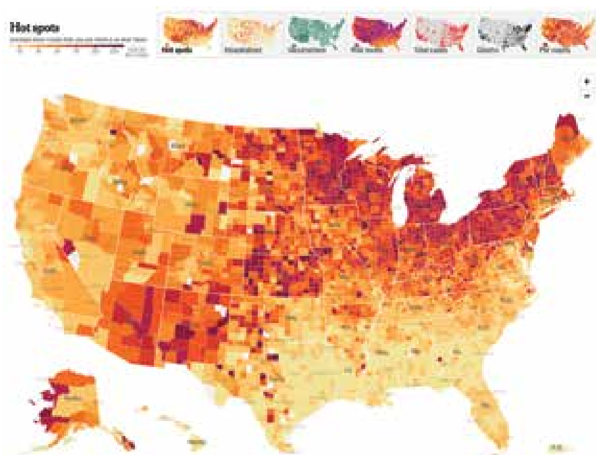

Figure 2. United States COVID-19 Hot Spots dashboards by the New York Times

There are various collections of COVID-19 dashboards and interactive tools available. Some provide only the data visualization of COVID-19 confirmed cases and deaths. Others visualized COVID-19 related impact on different sectors such as economy, education, etc. (Zuo, F., Wang, 2020). However, people were struggling with the enormous amount of CO-

VID-19 epidemiological data that grew by the second and the complexity of data (Geoffrey et al, 2018). In fact, the current methods of data visualization of COVID-19 epidemiological data have been criticized and have developed misinformation and misinterpretation about the scale of the pandemic, treatment, diagnosis, and prevention of the virus through social media (Murphy et al., March 2020).

The complexity of big data analysis of COVID-19 epidemiological data presents an undeniable challenge and the current visualization techniques, and methods must be improved. The data visualization design outcome of this research provides evidence of the visualization benefits in terms of collaboration, interaction, and sharing insights and information among the public. The interface design supports visualizing complex data in a single chart. Users control the time-series of data by month, which is useful to observe different locations at specific checkpoints. The field of big data visualization focuses on the innovation of tools that help users to construct rapid and effective observation with large amounts of data. We believe that correlation between Extended Reality (XR) such as augmented reality (AR) and big data visualization techniques are improving and becoming a more promising method. Thus, we used AR applications to display big data visualization of COVID-19 epidemiological data to help users identify the pattern in the increase of confirmed cases, death cases, vaccination rates, social distancing, unemployment, and economical impact of COVID-19 in the United States using AR.

AGE OF BIG DATA

Data exists everywhere. Location data, audio and video recordings, social media message flows, and more, in all areas of human activity today are considered as data (Childs et al., 2013). Big data is streaming of data with new data points, which can provide the possibility of a new method to understand, predict, and prevent the decisions related to the dataset. The government has been transferring data onto the web, Data.gov, that allows a variety of government data available to the public since 2019. However, it is very challenging to process provided data with the traditional methods that have been utilized currently.

The importance of big data is not about how much data has been collected, it is about what to do with the data. The concept of big data can be defined with three V's (Laney, 2001): Volume, Velocity, and Variety. Volume refers to collected data from a variety of sources. Velocity represents the unprecedented speed of data growth. Finally, Variety indicates data in all types of forms. Additionally, big data needs to be defined according to the activity of the specific organization such as data size, incompleteness, complex representation, heterogeneity of the data sources (Komodakis & Pesquet, 2015) as it covers various fields and sectors.

This research uses the term Big Data as a high volume, high velocity, and high variety data that is to obtain intended data value to secure the original data and obtained information. To observe the hidden pattern and find answers without overfitting big data, isolation of data is a mandatory process. Advanced forms of data and information analytics for process control and enhanced insight demand isolation of data to ensure originality of obtained information (Demchenko et al., 2014).

BIG DATA PROCESSING METHODS

Big data processing methods comprise a variety of fields including applied mathematics, statistics, computer science and economics, which are the basis for data analysis techniques (e.g., data mining, neural network, machine learning, signal process, and visualization methods). Often, these methods are interconnected and used simultaneously during data processing that increases higher levels of system utilization. Visualization for big data differs from other processing methods and from traditional visualization techniques. Feature extraction and geometric modeling must be employed to envision large-scale datasets, which encourages the reduction of the size of data before the rendering of data (Thompson et al., 2011). Visually represented information is easily acceptable by humans in comparison with formless text-based information (Horton, 1996). The data visualization design outcome of this research provides evidence of the visualization benefits in terms of collaboration, interaction, and sharing insights and information among users.

PREATTENTIVE VISUAL PROPERTY IN INFORMATION VISUALIZATION

People connect and remember information efficiently when visual elements (e.g., video, images, graphics, etc.) accompany words, which improve the information clutter to inspire and motivate viewers. As “a picture is worth a thousand words”(Advice, 1911), a data-driven art and visualization piece not only delivers a message in a highly perceptible way but also reveals patterns in complex information, which stimulates emotional connections to establish trust, loyalty, and positive experience for global audiences. In fact, more than 80 percent of an individual's decisions are based on one's emotional response (Stlouis, 2017). The interaction with data visualization is encouraged by one's unconscious perception (Padilla et al., 2018).

A preattentive visual property is from spatial memory without one's conscious actions, which is extremely rapid (less than 250 to 500 milliseconds) and accurate to process a preattentive property of any image (Bartram, 1997). A preattentive visual property avoids the conscious processing of all the presented information (requires intensive effort) that helps to comprehend information effortlessly. Utilization of preattentive properties of visualizations can encourage users to identify patterns of the dataset because it provides information with a comprehensive structure (Tukey, 1962).

Color, form, movement, and spatial movement are four main preattentive visual properties that play a critical role in human perception (Ware, 2019). Figure 3 exhibits an example of preattentive visual properties in form. Form has six types of attributes that can be manipulated to either call or reduce attention to a member of the data set. For instance, any of these six form attributes can be manipulated to either give or reduce an indication of its significance in a data set. In data visualization, experiments in psychology have used various features to perform preattentive visual tasks. The important two detections that are utilized in the big data visualization design are target detection and boundary detection. Target detection helps to detect the presence or absence of "target" elements with a unique visual feature within a field of distractor elements (Healey, C. G., 2007) (see Figure 4). Boundary detection is the recognition of the texture boundary between two groups of elements. Hues and saturations are useful to separate visual elements from the surrounding elements, which motivates to perform a visual search in the dataset (see Figure 5). Each visual example of six form attributes in figure 3 calls attention to a part of a visualization. Sub-attributes of movement are flicker and motion, which can call attention to a specific visual component; however, it must be carefully utilized as it can become a distraction from the rest of information. Two-dimensional (2D) positioning, stereoscopic depth, and concave and convex positioning must be considered in spatial positioning. 2D positioning is recognized as the best visual method of delivering data for quantitative data representations. 2D positioning in spatial positioning of data visualization was used in this research as a means for reflection in the information that requires comparisons.

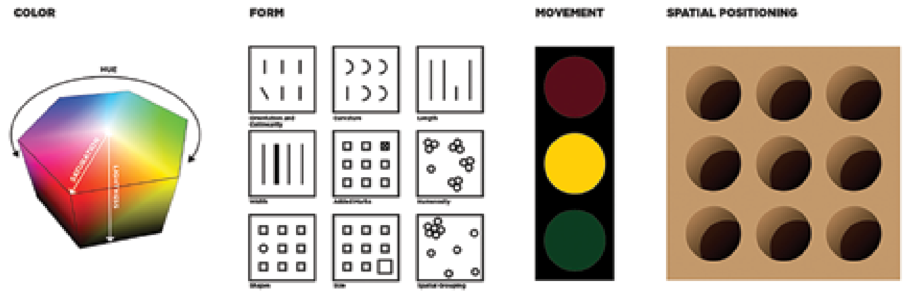

Figure 3. Four most popular preattentive visual properties

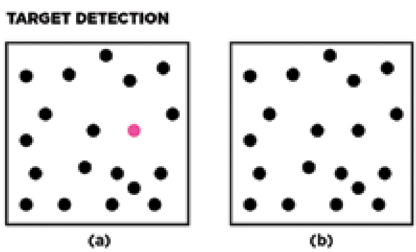

Figure 4. Searching for a target pink circle based on difference in hue. a) on a distractor (pink) in a sea of black colored circles. b) distractor is absent.

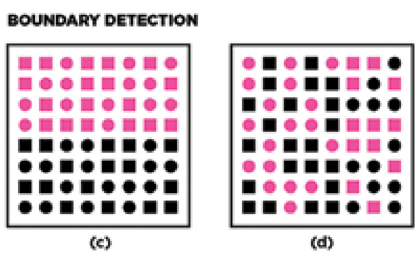

Figure 5. A boundary detection from Treisman's experiments

These four preattentive attributes can be applied to utilize the science of vision to improve information visualization (Ware, 2019). This research uses all four main preattentive visual properties to design the big data visualization of COVID-19 to reduce the cognitive effort to explore and understand the dataset and encourage only sensory memory. The effectiveness of using preattentive properties in this design is evaluated by the process of visual mapping, which helps to create structures that support the representation of the dataset. Further usability testing of this visualization design can be done in the future study to provide additional evidence of the appropriate impact to the user group.

DATA VISUALIZATION TECHNIQUES

The choices of visualization techniques influence the relevance and importance of data sets. Big data visualization methods can be classified into two categories – conventional visualization and interactive visualization. Most conventional visualization techniques are static and are lacking in handling big data (Chandra at el., 2019). Charts, plots, maps, timelines, and diagrams and metrics are commonly used for conventional visualization types. Depending on factors of the dataset, different data visualization techniques and configurations can be utilized. Various charts can be used from bars to lines that demonstrate the correlation between elements over time and periodic comparative data. Plots illustrate the distribution of two or more data sets in a two-dimensional or three-dimensional space. Scatter and bubble plots are commonly used in big data visualizations, which often explains relationships between large volumes of data. Maps are often used for data organized or segmented by physical geography. Maps allow viewers to locate elements on relevant objects and areas which is one of the idealistic techniques for COVID-19 data to predict and prevent from spreading the virus.

Kim and DiSalvo's research in 2019, recognized the interconnection between data visualization and graphic design fields that can create meaningful, insightful, and impactful data (Kim & DiSalvo, 2010) as graphic design strengthens argumentation with objectives. This research utilized three visualization techniques – charts, plots, and maps to create the visualization of COVID-19 data sets. The data sets require data comparison, trend illustration, and geographical locations. Furthermore, the objective data observation process influenced the structure of the systematic modular grid, which helps to create utilization methods of three data visualization techniques as well as four preattentive visual properties.

INTERACTIVE VISUALIZATION

Interactive visualization emphasizes graphical representations of data sets and inspires users to interact with information. Interactive visuals include graphic displays with rapidly changing data, which allows users to explore, extract, and understand information in an actively shifting data. The data visualization must have human input (e.g., clicking, scrolling, rotating, and zooming) and fast responsive time to display data input and visual output (Cook & Thomas, 2005). Most importantly, interactive visualization encourages easier exploration of real-time data. There are many visualization tools (see Table 1) for producing dynamic, interactive data visualization and bringing data to life.

Table 1. Big Data Interactive Visualization Tools

Tools	Type	Based on	Used for
R	Programming Language	Statistical techniques	Statistical computing and graphics
D3	JavaScript Library	Document Object Model (DOM) Manipulation	Data-driven transformations in browsers
Google Charts	Web Service	Google Visualization API	Customizable Charts
Tableau	Software Tool	Data Analytics platform with Business Intelligence	Interactive analysis of large-scale datasets

Dashboard visualization played an important role in enhancing efficiency and helping target patients' needs and monitoring COVID-19 trends, nationally and globally, during the unpredictable demands of a global pandemic. Beautifully rendered visualization made by D3.js for the New York Times in 2008 is one good example of interactive visualization (see Figure 6). It illustrates the box office receipts of all major commercial films released between 1986 and 2007. The visualization offers insight into how release patterns have changed over the past decades; and users can hover and click any "wave" to see more details for each movie (The New York Times, 2008). Tools like D3.js provide opportunities to generate fascinating interactive visualizations; however, they have accessibility issues due to compatibility of browsers and embedded software.

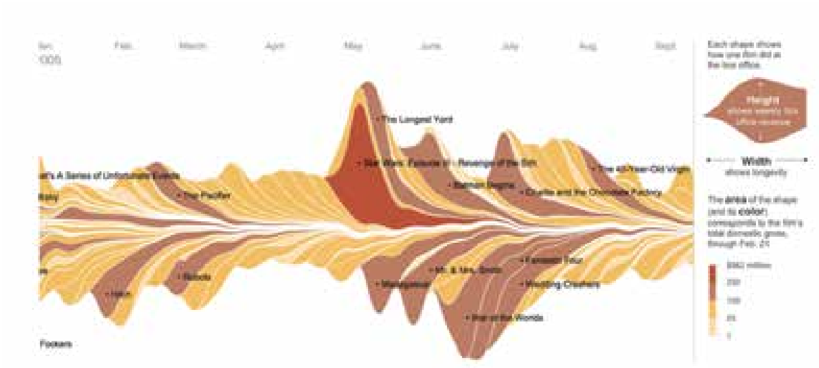

Figure 6. Movie grosses in New York Times using D3.js (February 23, 2008)

Big data processing methods are advancing as a division of the Human-Computer Interaction field today, thus the continuous innovation of new methods is necessary. We believe that big data requires multidisciplinary methods and techniques that allow us to explore complicated structures of interconnected data sets. Therefore, utilization of AR in big data visualization as a function of interactive visualization is rational. Interactive visual experience through AR can assist users to shape mental models of the data sets to understand and interpret data as well as identify patterns of the obtained information.

INTEGRATION WITH AUGMENTED REALITY (AR)

The human brain capacity of vision perception has limits (Field et al., 1993). The current equipment for displaying data visualization processes is expensive. The market is flooded with wearable devices and display devices that can be used for a new image interpretation. Augmented reality was invented in 1992 and the application has been found in the military, education, healthcare, and gaming industry (Ma et al., 2014). The devices such as Oculus and VIVE provide opportunities to practice for AR. Apps with smart devices deliver additional opportunities to have experience with AR, which makes it easier to implant virtual content into the physical world through these devices. The use of AR and VR platforms in the visualization can be a solution for a variety of issues such as navigation, scaling, angles, etc. However, the navigation inside of a virtual platform can be an additional influential issue as a user has control of moving around (Fonseca et al., 2013).

Scaling issues in big data visualization cannot be resolved due to multidimensional systems where a need to explore into a division of information to attain individual value or knowledge takes its place. AR displays can be limited by the visual system of the human eye. Considering most activities of the human eye are central vision (reading or driving) (Deering, 1998), we chose AR as a tool to create interactive digital data visualization into a real environment, which encourages users to learn and explore the big data sets in the interactive virtual environment.

ANALYSIS AND DEFINITION OF RESEARCH

The research used augmented reality as a visualization output of COVID-19 data from February 2020 to August 2021. Confirmed cases, deaths, social distancing, economy, unemployment, and vaccination were the primary focus of the big data visualization of COVID-19. The data were collected from the six national data sources. Data points of each data set were standardized to be positioned in the systematic modular grid that made the embed interactive visualization experience possible using AR through users' smartphone device. Artivive app was used as a tool of AR experience for users with their smartphone device. It is an image target-based tool to display animations and videos of how the visualization moves.

MATERIAL AND THE CONTEXT

The COVID-19 data sets utilized in the visualization design are from the six national data sources as below between February 2020 and August 2021 (see Table 2). Internationally and nationally reliable data sources were used only to prevent any possible bias and inputs outside of data sets. The base of the visualization analysis consists of confirmed cases, deaths, social distancing, economy, unemployment, vaccination. Monthly aggregated data trends were significantly important to identify data trends, which allows the prediction of the virus spread as the COVID-19 spread over a year (since February 2020). Daily transformed data sets were simplified as monthly data sets provide easier access to the annual trend of data points.

Table 2. Data sources of the data visualization of COVID-19

Confirmed Cases	**United States COVID-19 monthly confirmed cases in 2020 by state data source: World Health Organization (WHO)**
Deaths	**COVID-19 monthly deaths number in 2020 by state data source: Centers for Disease Control and Prevention (CDC)**
Social Distancing	**COVID-19 state stay-at-home orders in 2020 by state data source: Kaiser Family Foundation (KFF)**
Economic Impact	**COVID-19 monthly economic data in year 2020 data source: Nasdaq, Inc**
Unemployment cases	**United States unemployment correlation to race in 2020 data source: U.S. Bureau of Labor Statistics**
Vaccination rates	**COVID-19 vaccinations data source: CDC**

ANANALYTIC PROCEDURE

Systematic modular grid. Each dataset required different visual mapping for the data visualization. However, it was important to create the grid system that works for a variety of data sets to provide compatibility in smart device platforms. The systematic modular grid consists of horizontal partitions from top to bottom and vertical partitions from left to right (Lupton, 2014). For instance, a 12 x 12 grid affords 144 modules with over 20,000 possible combinations that allows flexibility. In other words, more modules mean the more possibilities. The 50x50 grid with 2500 modules was utilized to be a canvas for accepting various data sets (e.g., locations, hierarchy order, and population or data groups) flexibly and universally as a systematic modular grid (see Figure 7). The concept of modularity allows breaking the whole into smaller chunks (combined into limited groups) (Samara, 2017). It becomes a building block for flexible combinations that can be modified in a way that supports the characteristics of the data set. This process encourages flexibility, efficiency, and a consistent visual language in data visualization.

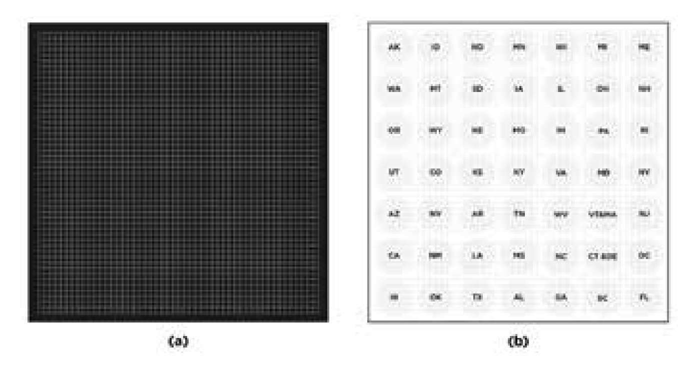

Figure 7. (a) systematic modular grid, a 50 x 50 grid. (b) a 7 x 7 column grid using a 50 x 50 grid with a geographical location system

Colors. Color is one of the important preattentive visual properties that is strategically utilized in the data visualization in this research. It is an extensively important factor in data visualization technique (charts, plots, and maps). Target and boundary detection are frequent use in the color property that can be utilized to highlight the story of data sets by detecting the presence and absence of the elements while poor use of them can create a distraction of data visualization. A qualitative palette is used for the variables with categorical nature. All variables in this research have distinct labels without inherent ordering. Each value of the variables was assigned one color from a qualitative palette (see Figure 8). Typically, the qualitative palette includes 10 or less colors. In particular, the vaccination dataset has groups by each state in the United States, which requires the color coding system to differentiate and recognize each state. Although use of these many colors can be troublesome, we controlled possible issues by using hue in the order of color spectrum followed by alphabetical order. It stimulates one's mental map to follow the group of information gradually to distinguish between groups of data sets.

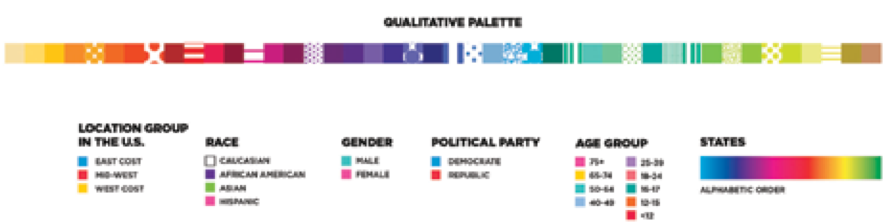

Figure 8. Qualitative palette and assigned colors of each variable

Augmented reality (AR). AR is utilized in this research to create a hybrid environment that connects the real world, data, and actions with elements of the virtual environment (Sonntag, 2019). AR can potentially apply to all senses, including touch, hearing, and even smell (Azuma, 2001). AR visualization has three basic properties: 1) combining reality and virtual information; 2) providing real-time interaction; 3) operating and rendering content in 3D space (Kipper, G., & Rampolla, 2012). Obtained virtual data from a remote device is augmented onto a display in real-time known as indirect vision, which is a process of creating a bridge between real life and digital contents. The first and second properties of AR visualization are primary concerns, and the visual sense is an essential method of interactive visualization in this research. We chose an app, "Artivive", to provide real life experience with big data of COVID-19 with virtual information that is created using Artivive (see Figure 9). Big data visualization designs from this research render the experience between reality and virtual information via AR using Artivive app with a real-time interaction with data sets.

Figure 9. Artivive app design interface (left), interaction output using Artivive app (right)

RESULT

Confirmed Cases. The data has been utilized in the confirmed case data visualization is from the United States COVID-19 monthly confirmed cases in 2020 by World Health Organization (WHO). Data points are carefully placed based on the physical geographical location of each state in the United States into a 7x7 column grid as one of combination modules of a 50 x 50 grid that helps to locate information on relevant areas by stimulating one's mental map (visual map) to observe the data trend effortlessly. Three preattentive visual properties were mainly used – color, form, and movement. Eight circles with size and weight changes (see Figure 10) were the main visual domain to illustrate data scales, significance, and movement of the data set. Lightness properties of color were used only to eliminate the distraction and target detection by use of colors to emphasize the spatial grouping migration with movement of data. The first raw graphics of figure 8 demonstrate how the information scale can be displayed in the visualization system. It was not necessary to display the information scale in every circle element in the visualization as it can be a distraction to call one's attention to the overall trend of visual mapping based on the movement property.

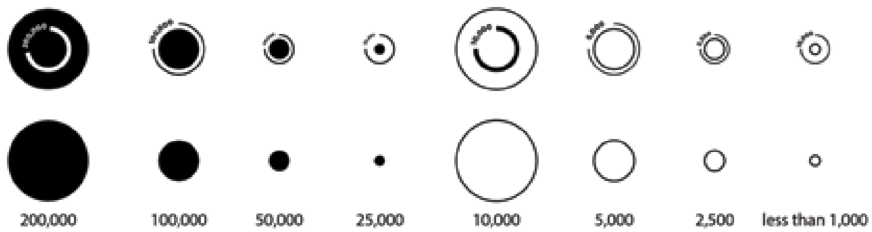

Figure 10. Confirmed case visual mapping units

Figure 11 illustrated the confirmed cases of COVID-19 data trend changes from February to December 2020 based on the geographic locations of the United States. The daily data point transformation was not the concern of this design method. The monthly aggregated data transformation provided a significant pattern of data migration and possible prediction of the virus.

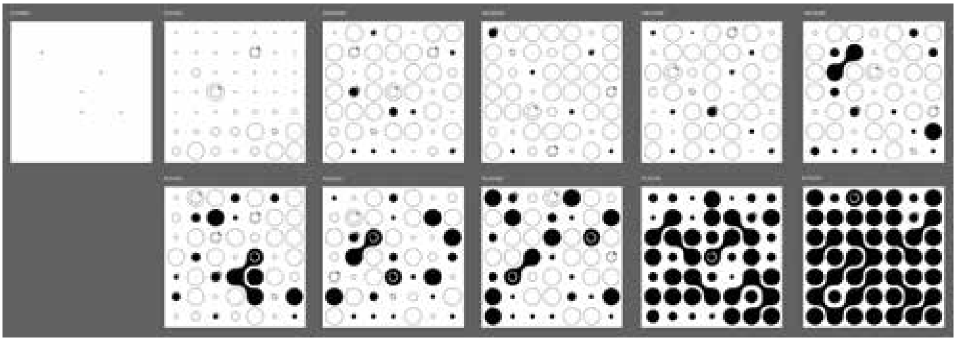

Figure 11. Confirmed cases visual mapping by geographical location in the U.S. in 2020

Deaths. The data source for the monthly covid-19 death number visualization comes from the CDC. The measurement for the monthly data set is based on per one million population of each state. A 7 x 7 column grid was utilized due to the geographical location base data set, which assists users' mental map to position information on relevant areas. Two main preattentive visual properties utilized the science of vision to help viewers compare, contrast, and visualize the information are color and movement. Red and blue were utilized from the qualitative palette, which is a direct representation of political parties – Republican (red) or Democrat (blue) in the United States. As these two colors become public domain to represent each political party, it is wise to use the same color system to reduce the cognitive effort to understand the meaning of colors. Red and blue (primary colors) are located in the near opposite sides of the color wheel. The chosen red and blue color has the same saturation that can provide a bold statement visually with finely tuned visual tensions from their surroundings (see Figure 12). The boundaries of each state were intentionally blended to encourage viewers to focus on the migration of the data set. The glitch effect was employed to symbolize the ruthless and unrelenting emergency of covid-19 as an art form to represent the concept of digital or analog errors. Moreover, the red and blue are set in different values in Z-dimension to effectively call one's attention when they are experiencing the AR with the Artvive app.

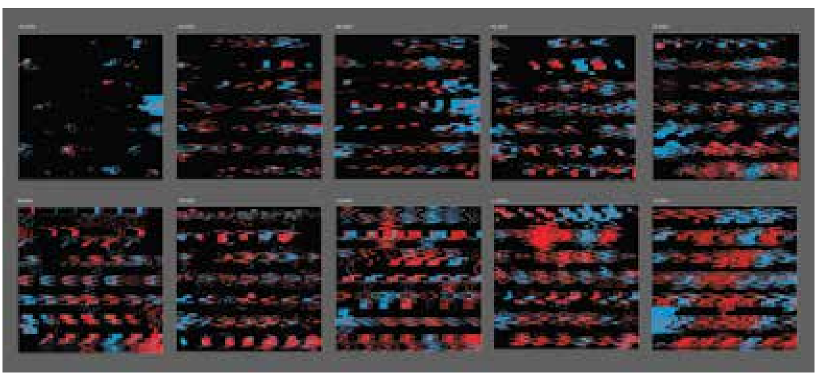

Figure 12. Number of deaths mapping by geographical location in the U.S. in 2020

Social distancing. COVID-19 stay-at-home orders in 2020 by state in the U.S. originate from the Kaiser Family Foundation (KFF). A 7x7 column grid was used to provide the geographical location mapping system for the data set. Each building block was divided into 10x10 small squares to indicate measuring distance in space. As the Figure 13 displays, there are four statuses of mandatory stay-at-home orders across states: no stay-at-home order; implementing stay-at-home order start; stay-at-home order end; and start and end stay-at-home order at the same month. 2D spatial positioning was the key preattentive visual properties utilized in the social distance category to deliver data that can be easily recognized and processed visually. By varying the distance between the lines inside each unit, it stimulates people's mind maps to find trends and discover changes over time. When the lines get closer together (e.g., the space between individuals) the more solid and brighter the area will appear. The more widely spaced the lines, which means decreasing the frequency of contact to reduce the risk of spreading the covid-19 virus, the darker the area will appear. White colored lines with backgrounds creates the motion illusion that increases an attention to the visual element and data migrations within the data set. Moreover, the experience of motion illusion in AR or VR can enhance the spatial positioning of information (Riecke at el., 2015), which is a significant behavioral benefit.

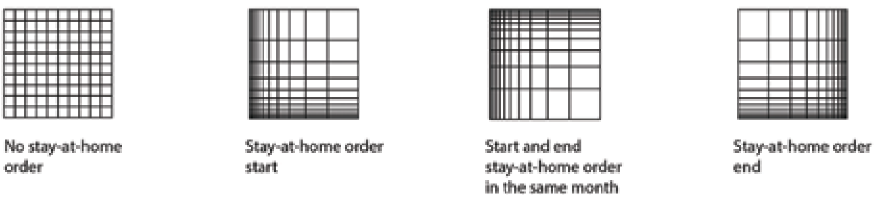

Figure 13. stay-at-home data visual mapping units

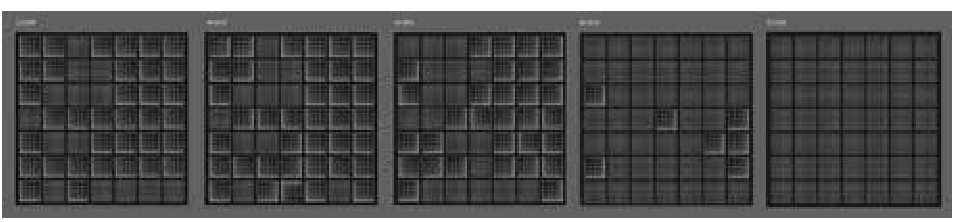

Figure 14. Social distancing visualization by geographical location of the U.S. in 2020

Economic impact. The 144 modules based on the standard systematic modular grid (50 x 50) were produced to formulate a 21x 6 table. The monthly economic data from Nasdaq, Inc (see Figure 15) was used to represent the economic impact. Ten U.S. indices and 11 global industry classification standard (GICS) sectors in a fiscal year (e.g., Q1, Q2, Q3, Q4, week-high, and week-low) was included in this visualization. From top to bottom the 10 US indices includes Nasdaq 100; Russell MicroCap; Nasdaq Composite; Russell Growth; Russell 2000; Russell 3000; S&P 500; Dow Jones Industrials; S&P Midcap 400; and Russell Value. The 11 GICS sectors include Technology; Consumer Disc; Basic Materials; Industrials RE-ITs; Financials; Staples; Communications; Energy; Healthcare; and Utilities. The order does not change to encourage users' mental map within the data set to reduce the learning curve of data locations. Colors, especially the HSL scale, are carefully chosen to make the economic data become easily recognized and processed visually. While red is associated with trading lower than the previous day's close in the stock market and green color connects with trading higher than the previous day's close. Yellow and white indicate no change in price. The characteristics of a stock data set is inherently ordered due to the trading activities. It is practical to use a sequential palette and colors are assigned to data values in a continuum based on both hue and lightness. The light colors represent lower values and darker colors represent higher values in light backgrounds (see Figure 16).

Figure 15. Economic impact data visualization of the U.S. within a fiscal year 2020

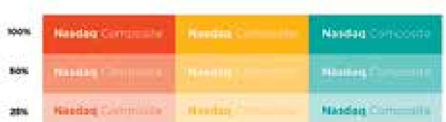

Figure 16. A sequential palette of visual mapping units

Unemployment. The data from the United States unemployment correlation to race in 2020 from U.S. Bureau of Labor Statistics was utilized in data visualization for this category. This data set displays the unemployment case statistics by race – Caucasian, African American, Asian, and Hispanic. Data was standardized by percentage that illustrated the proportion of each race in the total unemployment case number of each month. The standardized systematic modular grid of a 50 x 50 grid provides 81 diagonal lines (see Figure 17) as a unit of data point. Each line in 81 units represents 1.236 percent of the proportion of total data points by each month. Monthly unemployment data points were standardized based on 81 units of proportion placement of each race (see Table 3), which demonstrates the visual mapping of the proportional unemployment case by each race monthly. For example, 6,738 (total of 81 units) unemployment occurred in January 2020 that included 3,863 (46 units) unemployment case in Caucasian, 1,275 (15 units) unemployment case in African American, 315 (4 units) unemployment case in Asian, and 1,282 (15 units) unemployment case in Hispanic.

Table 3. Standardized unemployment cases example

Month/ Year	Total unem-ploy-ment case number	Caucasian		African American		Asian		Hispanic	
			81-unit ratio		81-unit ratio		81-unit ratio		81-unit ratio
01/2020	6,738	3,863	46	1,275	15	315	4	1,282	15
02/2020	6,703	3,869	47	1,251	15	257	3	1,326	16

The Color property in preattentive visual properties was significantly important in the unemployment case visualization. Each race has given colors from the qualitative palette – Caucasian (white), African American (purple), Asian (green), and Hispanic (pink). These four colors were carefully chosen to avoid the color association with each race and possible legibility issues as this data moves in an AR environment. The dark background was introduced in this data visualization to enhance legibility of four chosen colors in both two-dimensional space and three-dimensional space with AR. The proportion of lines is adjusted by changes of data proportion by race. The organic movement of lines represents the unemployment case increasing direction and it represents the sequential movement in AR (see Figure 18).

Figure 17. 81-line units based on a 50 x 50 grid

Figure 18. the migration of unemployment data by race

Vaccination rates. The United States COVID-19 vaccinations from the CDC was applied in the vaccination rate data visualization. A 50 x 50 grid employed to create an alphabetic order of states in the U.S (see Figure 19). Color property took the dominant role in vaccination rates visualization. Full range colors were used to indicate each state in the U.S. Hue in the order of color spectrum was used to reference the order of alphabets that encouraged the one's mental map to explore information in hierarchical order. Patterns are used to create visual interests. Figure 20 demonstrates movement in AR illustrating the migration of vaccination rates data.

Figure 19. vaccination rates by states

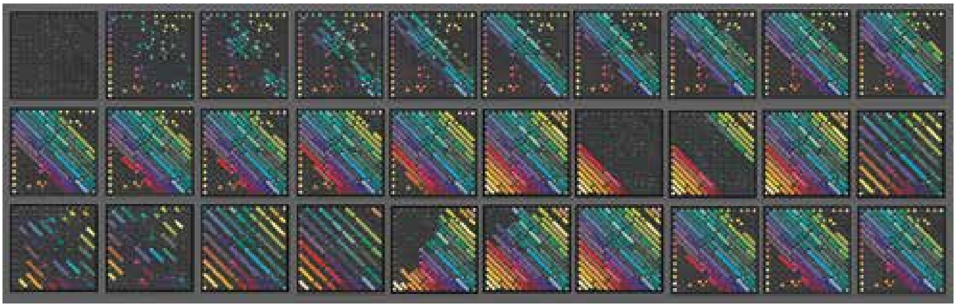

Figure 20. the vaccination rate by states with real-time sequence in AR

CONCLUSION

In this research, six national and international data sources were employed to create big data visualization of COVID-19 using augmented reality. Six categories of COVID-19 data – confirmed cases, deaths, social distancing, unemployment cases, vaccination rates, and economical impact, were special interests in this research. We used three preattentive visual properties (e.g., colors, forms, and movement with AR) with a unique visualization form based on the popular visualization techniques that helped to build a systematic modular grid system to deliver complex information efficiently and encourages comprehensive data exploration while avoiding the conscious processing of the presented information.

Utilization of AR in big data visualization presents advantages and disadvantages; however, it is possible to optimize strategies to reduce disadvantages. Data visualization with AR experience amplifies the function of data visualization and interaction with the big data that encourages the public to understand meaning and relationship at individual and society level from the high dimensional data. It can be concluded that data visualization methodology and the natural interaction with data sets can be improved by implementation of fundamental cognitive design principles. Furthermore, a smart device AR interaction with data visualization provides spontaneous and effortless interaction for users. The AR experience with big data visualization with COVID-19 will be available on the Emotion Lab website (www.emotionlab.org) by October, 2021. We hope this data visualization method is valuable for a broad range audience to encourage the innovation of identifying, controlling, combating, and preventing COVID-19 as well as any future epidemic. Further, we hope that the in-person and authentic data visualization experience with AR encourages the public to reflect spatial and personal relationships to spark empathy from the public.

REFERENCES

Advice, S. G. S. (1911). Syracuse Post Standard. March, 28, 18.

Anderson, J., Rainie, L., & Vogels, E. A. (2021). Experts say the ‘new normal’ in 2025 will be far more tech-driven, presenting more big challenges. Retrieved from PEW: https://www. pewresearch. org/internet/2021/02/18/experts-say-the-new-normal-in-2025-will-be-far-more-tech-driven-presenting-more-big-challenges.

Ajilore, O., & Thames, A. D. (2020). The fire this time: the stress of racism, inflammation and COVID-19. Brain, behavior, and immunity, 88, 66.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. IEEE computer graphics and applications, 21(6), 34-47.

Bartram, L. (1997, November). Perceptual and interpretative properties of motion for information visualization. In Proceedings of the 1997 workshop on New paradigms in information visualization and manipulation (pp. 3-7).

Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. The lancet, 395(10227), 912-920.

Buheji, M., & Ahmed, D. (2020). Foresight of Coronavirus (COVID-19) opportunities for a better world. American Journal of Economics, 10(2), 97-108.

Chandra, A. N. R., El Jamiy, F., & Reza, H. (2019, December). Augmented reality for big data visualization: a review. In 2019 International Conference on Computational Science and Computational Intelligence (CSCI) (pp. 1269-1274). IEEE.

Childs, H., Geveci, B., Schroeder, W., Meredith, J., Moreland, K., Sewell, C., ... & Bethel, E. W. (2013). Research challenges for visualization software. Computer, 46(5), 34-42.

Cook, K. A., & Thomas, J. J. (2005). Illuminating the path: The research and development agenda for visual analytics (No. PNNL-SA-45230). Pacific Northwest National Lab. (PNNL), Richland, WA (United States).

Deering, M. F. (1998, May). The limits of human vision. In 2nd International Immersive Projection Technology Workshop (Vol. 2, p. 1).

Demchenko, Y., De Laat, C., & Membrey, P. (2014, May). Defining architecture components of the Big Data Ecosystem. In 2014 International conference on collaboration technologies and systems (CTS) (pp. 104-112). IEEE.

Dur, B. I. U. (2014). Data visualization and infographics in visual communication design education at the age of information. Journal of Arts and Humanities, 3(5), 39-50.

Fairchild, G., Tasseff, B., Khalsa, H., Generous, N., Daughton, A. R., Velappan, N., ... & Deshpande, A. (2018). Epidemiological data challenges: planning for a more robust future through data standards. Frontiers in public health, 6, 336.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local “association field”. Vision research, 33(2), 173-193.

Fonseca, D., Villagrasa, S., Martí, N., Redondo, E., & Sánchez, A. (2013). Visualization methods in architecture education using 3D virtual models and augmented reality in mobile and social networks. Procedia-Social and Behavioral Sciences, 93, 1337-1343.

Healey, C. G. (2007). Perception in visualization. Retrieved February, 10, 2008.

Ho, J. (2021, January). Anti-Asian racism, Black Lives Matter, and COVID-19. In Japan Forum (Vol. 33, No. 1, pp. 148-159). Routledge.

Horton, W. (1996, April). Designing icons and visual symbols. In Conference companion on Human factors in computing systems (pp. 371-372).

Jiang, F., Leung, C. K. S., & Tanbeer, S. K. (2012, November). Finding popular friends in social networks. In 2012 Second International Conference on Cloud and Green Computing (pp. 501-508). IEEE.

Jonker, D., Brath, R., & Langevin, S. (2019, July). Industry-driven visual analytics for understanding financial timeseries models. In 2019 23rd International Conference Information Visualisation (IV) (pp. 210-215). IEEE.

Kim, T., & DiSalvo, C. (2010). Speculative visualization: A new rhetoric for communicating public concerns.

Kipper, G., & Rampolla, J. (2012). Augmented Reality: an emerging technologies guide to AR. Elsevier.

Komodakis, N., & Pesquet, J. C. (2015). Playing with duality: An overview of recent primal? dual approaches for solving large-scale optimization problems. IEEE Signal Processing Magazine, 32(6), 31-54.

Krum, R. (2013). Cool infographics: Effective communication with data visualization and design. John Wiley & Sons.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. META group research note, 6(70), 1.

Laster Pirtle, W. N. (2020). <? covid19?> Racial Capitalism: A Fundamental Cause of Novel Coronavirus (COVID-19) Pandemic Inequities in the United States. Health Education & Behavior, 47(4), 504-508.

Lupton, E. (2014). Thinking with type: A critical guide for designers, writers, editors, & students. Chronicle Books.

Le Bras, P., Gharavi, A., Robb, D. A., Vidal, A. F., Padilla, S., & Chantler, M. J. (2020). Visualising covid-19 research. arXiv preprint arXiv:2005.06380, 1.

Ma, M., Jain, L. C., & Anderson, P. (2014). Future trends of virtual, augmented reality, and games for health. In Virtual, augmented reality and serious games for healthcare 1 (pp. 1-6). Springer, Berlin, Heidelberg.

Manav, B. (2007). Color□emotion associations and color preferences: A case study for residences. Color Research & Application: Endorsed by Inter‐Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur, 32(2), 144-150.

Murphy H, Di Stefano M, Manson K (2020, March 20). Huge text message campaigns spread coronavirus fake news. Financial Times. Retrieved from: https://www.ft.com/content/34b6df5a-ea4a-471f-8ac9-606580480049

Padilla, L. M., Creem-Regehr, S. H., Hegarty, M., & Stefanucci, J. K. (2018). Decision making with visualizations: a cognitive framework across disciplines. Cognitive research: principles and implications, 3(1), 1-25.

Redelmeier, D. A., & Shafir, E. (2020). Pitfalls of judgment during the COVID-19 pandemic. The Lancet Public Health, 5(6), e306-e308.

Riecke, B. E., Feuereissen, D., Rieser, J. J., & McNamara, T. P. (2015). More than a cool illusion? Functional significance of self-motion illusion (circular vection) for perspective switches. Frontiers in Psychology, 6, 1174.

Roberts, K. (2020). Locked down leisure in Britain. Leisure Studies, 39(5), 617-628.

Sadiku, M., Shadare, A. E., Musa, S. M., Akujuobi, C. M., & Perry, R. (2016). Data visualization. International Journal of Engineering Research And Advanced Technology (IJERAT), 2(12), 11-16.

Sonntag, D., Albuquerque, G., Magnor, M., & Bodensiek, O. (2019). Hybrid learning environments by data-driven augmented reality. Procedia Manufacturing, 31, 32-37.

Souza, J., Leung, C. K., & Cuzzocrea, A. (2020). An innovative big data predictive analytics framework over hybrid big data sources with an application for disease analytics. In Advanced Information Networking and Applications (Vol. 1151, p. 669). Nature Publishing Group.

Samara, T. (2017). Making and breaking the grid, updated and expanded: A graphic design layout workshop. Quarry Books Editions.

Stlouis, M. (2017, August 1). How to Spot Visual, Auditory, and Kinesthetic-Learning Executives. Inc. Retrieved from: https://www.inc.com/molly-reynolds/how-to-spot-visual-auditory-and-kinesthetic-learni.html

The New York Times. (2008, February 23). The ebb and flow of movies: Box office Receipts 1986 - 2008. Retrieved from: http://archive.nytimes.com/www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html.

Thompson, D., Levine, J. A., Bennett, J. C., Bremer, P. T., Gyulassy, A., Pascucci, V., & Pébay, P. P. (2011, October). Analysis of large-scale scalar data using hixels. In 2011 IEEE Symposium on Large Data Analysis and Visualization (pp. 23-30). IEEE.

Tu, J., Verhagen, M., Cochran, B., & Pustejovsky, J. (2020). Exploration and discovery of the covid-19 literature through semantic visualization. arXiv preprint arXiv:2007.01800.

Tukey, J. W. (1962). The future of data analysis. The annals of mathematical statistics, 33(1), 1-67.

Ward, M. O., Grinstein, G., & Keim, D. (2010). Interactive data visualization: foundations, techniques, and applications. CRC press.

Ware, C. (2019). Information visualization: perception for design. Morgan Kaufmann.

WHO (2021, September 3). WHO Coronavirus (COVID-19) Dashboard from 2020 to 2001. WHO. Retrieved from: https://covid19.who.int/

Williams, L. G. (1966). The effect of target specification on objects fixated during visual search. Perception & Psychophysics, 1(5), 315-318.

Wolinski, F. (2020). Visualization of diseases at risk in the COVID-19 Literature. arXiv preprint arXiv:2005.00848.

Yabuki, N., Furubayashi, S., Hamada, Y., & Fukuda, T. (2012, September). Collaborative visualization of environmental simulation result and sensing data using augmented reality. In International Conference on Cooperative Design, Visualization and Engineering (pp. 227-230). Springer, Berlin, Heidelberg.

Zuo, F., Wang, J., Gao, J., Ozbay, K., Ban, X. J., Shen, Y., ... & Iyer, S. (2020). An interactive data visualization and analytics tool to evaluate mobility and sociability trends during covid-19. arXiv preprint arXiv:2006.14882.