Digital Data Visualization using Augmented Reality (AR). Review of the Current State of Art

Ratnesh Chaturvedi[1*], Nitya Darshanil[1*], Yash Kumar Gandhi[1], Neelak Ghosh[1], Shantanu Gupta[1] and Archit Sen Gupta[1]
[1]*Narsee Monjee Institute of Management Studies - [**NMIMS** Deemed to be University]*
V. L, Pherozeshah Mehta Rd, Vile Parle, Mumbai, Maharashtra 400056
**Contact: ratnesh.chaturvedi@nmims.edu, nitya.darshanil24@nmims.edu.in*

Abstract— **This paper presents an overview of the research issues and results for data visualisation techniques and its instruments in a multidisciplinary way. The primary aim of this review is to sum up the all the existing data visualization tools with the challenges in those methods and to provide new solution for the next era. This paper audits the current state of intelligently and immersive innovation for data visualization which is for open and enormous information, to highlight potential projections of information visualization based on variables such as client encounter. Situating the foremost basic information in an expanded reality within the central range of human visual areas permits one to get the data presented within a brief time, without any critical information misfortune due to human perceptual issues. Moreover, the audit will give an knowledge towards the impacts of modern advances, such as Virtual Reality shows and Augmented Reality head wears on the Big Data visualization as well as to the classification of the most challenges of joining the innovation.**

I. Introduction

Humanity's entire former history is considered to be a huge collection of data. For thousands of years, information has been preserved and used in various fields of industries. This data is now a key component in history, politics, science, business, and even in social life. This trend can be seen in our daily life on social networks, like Facebook, Twitter and Instagram. The users of these social media accounts generate a huge amount of information everyday (music, images, text...) which we call as the “RAW DATA”. In today 's information society, resource exploitation and data replication has been developed as a complex scenario. Institutions and agencies, on the one hand, have an increasing tendency to leave the data open. The same reason gives rise to disciplines related to data access and management, some examples for the same can be data science, data mining, big data, or deep learning

Many existing computing frameworks have tried through more immersive user interfaces, to become more adaptable to human perceptual vision. There are also many significant issues when we talk about these Information Visualisation Tools. In this sense, Augmented Reality (AR) tests different modes of interaction and simulation and makes a seamless experience for the users as well as for the data on which it is being processed on.

One of the few questions which emerge from this point concerns the way in which we really need to speak to and visualize the data included, and which component ought to we consider to make it more interactive. In other words, regardless of whether the data may be a blend of absolutely subjective nature, the way in which we need to construct it depends upon the story we need to tell, which in turn is reliable in related factors such as Client Involvement and Interaction Plan.

In this regard, we discover the interface element that we want the user to be Immersed in a futuristic, digital environment. That is, how we can envision the possible situations in which the user is submerged in a world of data where s/he automatically understands and interprets the visualized data. How this user interface is actualized or how the interfaces are amplified in a sense, so distant, to the components related with the utilize of this innovation, gives rise to various issues, such as: what interaction systems does technology right now provide to represent and collaborate with the user? How this user interface is actualized or how the interfacing are opened up in a sense, so far off, to the components related with the utilize of this advancement, gives rise to different issues, such as: what interaction frameworks does technology right presently give to speak to and collaborate with the user?

II. Literature Review

The systematic literature review for this research started in August 2020. First of all, data and information about the topic was kept in mind. We searched for keywords such as data mining, data science, visualization tool, open access, representation of data etc for primary search of information on this topic. The issue of the subject and the field of study were also established. In this case, questions were focused in the role of data in fields such as data visualization using Augmented reality.

The efficient review connected highlights the have to be compelled to actualize unused strategies for utilizing and collecting data. Furthermore, other perspectives of the systematic literature review handle have been combined with other regions, such as client-technology partnerships, data analysis, user experience, interfaces and their future utilize in

data visualization. This reality has permitted us to examine the consistency of the data and to distinguish the more important highlights that show us the genuine background for in-depth review.

The utilize of open data by open institutions and associations is the junction between the openness of legislative bodies and the engagement of individuals (Peled, 2011). In reality, as comparative terms drop into the same context (inside the "data" set), such as big data, data mining, open investigate, or learning analytics, it generally consolidates a wider phenomenon connected to the potential of data revelation, administration, and representation technologies.

But what "data" means? According to writers like (Kitchin, 2014), evidence comes before statement or perception, in rhetorical terms.

The model is considered, based on this think about, as a handle that joins the related components that must be considered as a whole in order to overhaul end-user involvement, to have an impact on spaces or disciplines such as empowerment of citizens or schooling. In a research piece examining data framework, (Kitchin, 2014) assesses how the data environment is seeing a constantly changing stage and gives a differentiation that separates open data from big data. In reality, since data mining strategies in both open data and big data are oppositely distinctive, the integrating strategy and advance processing can depend more on the target gathering of people than on the way the data is collected.

The key concern is, firstly, how the data is collected and stored and, furthermore, how this data is pictured. Both are forbid, which in turn depend on the accessible technical capability in any given environment.

In turn, they are distinctive forms, especially with respect to representation, because it makes different conceivable outcomes that outline the relationship between customer, space, technology, data, and image.

Visualization of the data is in this manner is an vital highlight within the strategy of data interaction. Authors like (Marr, 2016) highlight this noteworthiness of visualisation, noting that it could be a vital portion of the ultimate critical stage in data ventures. Within the light of the endless volume of data accessible nowadays that should be dealt with, this author suggests strategies by which this amount of data can be ideally handled and controlled.

Recent studies (Donalek et al., 2014) deal with the need to discover ways of to represent data in interactive and immersive conditions. In other words, augmented reality, as a platform, could be a tool with specific characteristics, such as its interactive and immersive nature, which empowers us to create relational processes that express new discourses about data use. In either case, there are currently different publications and think abouts related to this integration, including augmented reality, between data and interactive worlds.

Typically the case for (Huang et al., 2001), which consolidates a number of components for spatial data processing, or (Helbig et al., 2014), utilizing 3D computational data visualisations to amplify them to logical areas such as meteorology. Also included in his thesis is (Wexelblat, 2014), which focuses on the future uses of augmented reality.

This literature review helps us to look at how immersive and digital technology could be at present, a phase of change of our connectivity, engagement with our colleagues and our perception of reality. And as (Marr, 2016) says, the quantitative and efficiency leap in the interface scale can dramatically improve the scope of the data we have in our brains when the consumer is absorbed into the digitally-created environment with a 360-grade field of vision and 3D virtual movement.

III. DATA VISUALIZATION TOOLS

Visual representation of the information may be a vital step in Big Data to be able to get all the benefits that this innovation provides.Improved decision-making is one of the key advantages of visualisation applications, but other significant factors to consider include enhanced ad-hoc data processing, improved user collaboration and increased return on investment (ROI). There are a variety of traditional methodology, including charts, histograms, storyboards or columns, for visually representing details. This does not mean that any kind of representation possible should be used at the time of creating a visualisation, but each visualisation approach has a particular purpose and it is also important to decide which sort of detail is implied to be represented and thus select the foremost reasonable graphic representation. For instance, you'll utilize pie charts on the off chance that you need to show numeric parts, whereas bar charts are more accommodating in the event that you need to show comparisons, or in case you need to show hierarchical connections, tree outline charts will be more fitting for this frame of data.

There are modern web-based visualization approaches that can offer assistance to decrease the generation costs and give associations with diverse data sources permitting the representation to be up-to date. These devices are exceptionally well known on business analytics environments but they are not utilized in scientific regions. Examining the current showcase, it is conceivable to classify visualization tools in two primary bunches:

Visualization tools for DEVELOPERS

D3

Characteristic	Interpretation
Data load	There are no constraints in this respect, because a backend for each solution has to be created. This library allows any kind of data to be attached to a DOM (Document Object Model) and, once collected, multiple transformations can be applied to the document depending on the desired visualisation.
Data Visualization	As it is not a graphic library but a JavaScript library that uses web standards, there are no limits on forms of visualisations. It focuses on manipulating the DOM through visualisation and interaction methods in such a way that it has both the browser functionality and the freedom

	to construct the most suitable visualisations.
Advantages	If someone is searching for high-quality graphics, and if the customer knows just what he or she is looking for, the ultimate solution. It enables a portal/application to be perfectly embedded. Although developers are expected to use this method, this enables users to build almost any form of visualisation and thus get the most appropriate for each scenario against other common and easy-to-use software where the user wants to restrict the functionality offered for each visualisation
Disadvantages	Specialized planners are expected to construct visualisations, the production time is high and output limits are presented as soon as the number of items to envision increases.

High Charts

Characteristic	Interpretation
Data load	This tool loads the data into a JavaScript list to be used in the visualisations later on. A specialised developer must programme this load. You will find the data array either in the local configuration, or in a different file or in a separate path.
Data Visualization	Both desktop and mobile browsers have a wide range of visualisations available.
Advantages	Complex visualisations are not easy and fast to use as long as what you want to create is not complex. It is a library that represents various types of graphs as JSON objects, so changing the display type from a bar map to a circle one, for example, is as easy as changing a property from 'bar' to 'apple' in the JSON file. It allows the visualisations to be dynamically updated and the graphs to zoom in.
Disadvantages	It does not allow new graphics to be made. To build visualisations, professional developers are required.

Leaflet

Characteristic	Interpretation
Data load	You can use data from different sources since it is a JavaScript library, so you need a skilled programmer to create the code for loading it. To load the data, the library uses GeoJSON files.
Data Visualization	The dynamic map view for mobile devices is focused on.
Advantages	Enabled on both the desktop and smartphone, the most relevant browsers. Light and without external addictions. Developers may not have to have experience of GIS.
Disadvantages	Required specialized developers. Only maps are displayed.

R

Characteristic	Interpretation
Data load	This requires data from various file types such as CSV, SAS, TXT, JSON and SPSS to be downloaded, or directly from software such as Microsoft Excel, Microsoft Access, Oracle, MySQL and SQLite.
Data Visualization	High-quality images that can be exported to PDF, JPG, PNG, SVG, and even LATEX and HTML tables. Developers have dedicated graphics libraries for
Advantages	It has a broad variety of mathematical and graphical methods focused on statistical analysis. Since it is a programming language, it encourages users, aside from what it already supports by design, to create their own features, such as classification and grouping algorithms, statistical analyses, time series analysis. As it is an open source (and an open repository), packages that expand the features offered by R can be shared by all users. There are many graphical interfaces, editors, and IDEs that make the job simpler, considering the difficulty. For both 32-bit and 64-bit processors, it operates for all major operating systems. It handles huge data volumes and, due to the automation of many processes by scripts, preparing them (to be loaded) is simpler than for other developer tools.
Disadvantages	Specialist developers are required because it is a complex language and has a high learning curve.

Visualization tools for GENERAL USERS

Tableau

Characteristic	Interpretation
Data load	Supports most of the data sources.
Data Visualization	Quick. Fast. For the consumer, several visualisations are usable, not only appealing but also practical. Enabled on many platforms (desktop,

	mobile, browser) The most common use is for internal dashboards that allow simple data analysis by the user. It is necessary, however, to indicate that it is not an analytical instrument.
Advantages	Easy to use and manipulate to get more appropriate visualizations
Disadvantages	Knowing R is important if more complex visualisations are pursued (allows the connection with R to use a series of expressions). Some types of visualisations are accessed via the Tableau Server, such as dashboards, and only those supported by the tableau are formatting functions. If the user needs to render those visualisations that are not supported by Tableau, this method cannot be used to build them.

Google Fusion Tables

Characteristic	Interpretation
Data load	You can use text files separated by a character to import data from a file, files encoded as UTF8, and the user can choose a URL to locate data, among other alternatives. It is a web service that automatically displays the user a map or a schematic of it until the data has been loaded.
Data Visualization	The user is restricted to the kinds of views available, and before making visualisations, the user has to know a visual programming language, but it is not complicated. There is a possibility to incorporate filters
Advantages	This tool makes it easy, for example, to use fusion tables with other users to allow data comparisons between the same entities. It is possible to transform localization tables into maps very easily. Simple to use, it gives easy visualisations, but for more sophisticated projects, the user has the option to use Google libraries or Java script. It is very effective to detect the same data when handling the data, even though they have distinct formats (Spain and SPAIN).
Disadvantages	The user may need to prepare the data before being imported to follow a format understandable by the tool (directions, percentages, formulae)

Quadrigam

Characteristic	Interpretation
Data load	It uses Google Drive to store data and JSON files to load it.
Data Visualization	There are more than 50 types of interactive visualizations, from bars to trees or animated maps.
Advantages	It is possible to create interactive visualizations without having programming skills. Its handling is very intuitive
Disadvantages	User options are limited to the types of visualisation available in the tool, without the possibility of creating new graphic classes. To construct visualisations Knowing a specific programming language that is easy to learn is necessary.

Datawrapper

Characteristic	Interpretation
Data load	Data can be loaded through CSV files.
Data Visualization	This tool provides a basic visualization which allows the user to get a result in a fast and easy manner.
Advantages	Fast and easy to use, it does not require a lot of user training
Disadvantages	The tool provides a list of visualizations not very configurable and there is no option to create new types.

IV. COMPARISON OF DATA VISUALIZATION TOOLS

	Load data files	**Visualization types**	**Development time**	**Easy to use**	**Configurable**	**Option to create**
D3	Any	No limit	High	No	Yes	Yes
High Charts	Any	Large	High	no	Yes	Yes
Leaflet	Any	Only maps	High	No	Yes	no
R	Multiple	Any	High	No	Yes	yes

Tableau	Multiple	Large	Low	Yes	Yes	No
Google Fussion Tables	Multiple	Large	Medium	Yes	Yes	No
Quadrigam	JSON	Medium	Medium	Yes	Yes	No
Data wrapper	CSV	Low	Low	yes	No	No

Different visualisation procedures, in any case, are right now utilized to represent distinctive data types. The number of forms, clearly, is as it were obliged by human creative ability. Clarity and ease of the interpretation of the data represented is the key prerequisite. Procedures for visualisation can be both basic (line charts, charts, bar charts, etc.) and complex (based on the numerical apparatus). In expansion, visualisation can be utilized as a blend of diverse methodologies. Visualized data representation, in any case, is abstract and exceedingly compelled by the capacities and requests for recognition (See Figure 1)

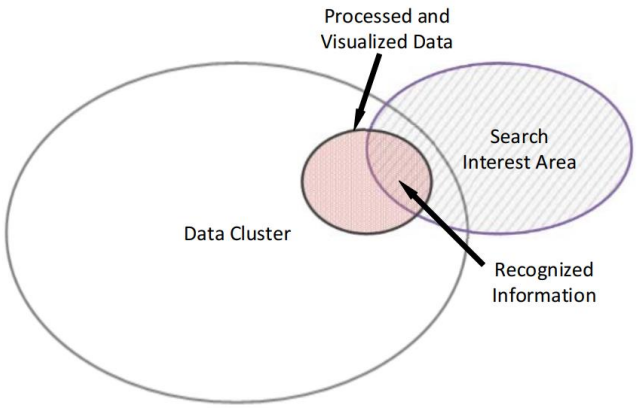

Figure 1: Human Perception Capability

V. NEW APPROACH FOR BIG DATA VISUALIZATION – AR

Big Data and AR have a sensible relationship that produces them an start for a modern strain of challenging application space. By completely interacting with the audience, AR has the capacity to process, improve and visualise the data in an out-of-the-box way without any external diversions. AR applications require more data to process and picture it within the most intelligently way than any other visual applications. Specific data AR applications also require external data on the environment and the environment, in expansion to the application. The source of big data for AR applications would be the zettabytes of data produced each day over the web, as already said. For numerous analysts, the collaboration between AR and Big Data is presently a major research point and both give a more extensive opportunity for innovation within the data visualisation areas. The potential collaboration opportunities give users and associations with intellectual applications and frameworks that give personal suggestions, information analytics, etc. To get the best results out of it, any data must be analysed and it is a genuine task when working with large sets of data. In addition, for analysis purposes, Big Data requires human-machine collaboration to input the information. Prevailing collaboration helps to superior understand the data and extricate some valuable data.

2D/3D modelling AR stands out on its possess, indeed with numerous visualisation procedures such as tables, charts, as the data visualisation in its case is implanted within the genuine world to provide a more prominent knowledge into the data. For illustration, it would be more appealing to show a 3-dimensional building plan on the building location through an AR environment than a 2-dimensional plan on a sheet of paper or a screen that makes a difference in workplace diagnosis within the case of situation visualisation. Data can be upgraded as AR enables clients to be soaked in a entire modern world mixed with interactive data and their viewpoints. Such data will be considerably more easily comprehended on the off chance that it relates original content and there doesn't arise a need for interpretation of the data. It is vital to be up to date indeed while rendering Big Data within the real-world environments with AR. Numerous applications specific and content particular necessities ought to be considered while doing so. Data must be handled and implanted into a real-time environment in a way that it should consistently fit into the environment. AR features a 3-dimensional visual interface which depicts the data without diverting the client far from reality.

VI. APPLICATIONS OF AUGMENTED REALITY IN CURRENT SCENARIO

Augmented reality and Big Data have presented new trade rules. But not all the businesses are able to execute this collaborative visualization procedure due to the insufficiency or restricted information. In spite of the fact that there are only constrained markets where this visualization is pertinent, it is clear that it might go distant past.

Medicine:

Within the field of medicine and healthcare, the need for data and the prompt handling of the dame is exceedingly essential. AR has been utilized for around 10 years within the healthcare division. Huge amounts of patient information are promptly accessible, which can be visualised in an individual or geological zone to extricate frequent sicknesses and can be utilized for therapeutic education purposes. People have utilized smartphones and other observing gadgets for wellbeing and activity that create information about blood pressure, heart rate, cholesterol levels, etc. on the go. It is conceivable to handle

all such data and AR can visualise it with real-time notifications in an interactive way. There's a plausibility of whee, as well. There's too a plausibility where AR can foresee the nutritional value of a supper just by checking it through a mobile camera.

Retail:

The retail markets have too been influenced by Augmented Reality, but in a few topographical ranges it is restricted. AR-enabled layers overlay data on the current screen by taking account of the environment. IKEA is one of the distributors in the United States that has executed AR in its mobile application that virtually places a furniture thing within the house or office of the client to suit the environment.

Tourism:

Another segment where AR is broadly related is tourism. From spatial investigation, GPS-based tracking, examining patterns from tweets or check-ins, surveys, etc., expansive data sets are produced that can be utilized to examine and give an interactive tour planner/guide for customers. In arrange to create intuitively 3-D maps for progressed navigation systems, topographical data is additionally utilized.

VII. CONCLUSION

When examining sorts of interactive and immersive innovations and their advancements, we essentially consider the potential scope of application and lines of research into how information is utilized and interpreted, and how interaction frameworks are outlined. Interactive and immersive advanced innovations are a transition technology that requires studies of, one the one hand, its potential for representation and, on the other, its capacity to associate with the client and create scenarios in which these can be represented through pictures, use of space and separations, sound.

The current improvement and advancement in Big Data and Visualization algorithms have assured an ease in visualizing the Big Data and reducing the impact of this challenge. The paper enlightens us about the overview of Big Data, introduces the use of Augmented Reality and its concepts when it meets Big Data and Visualization. The paper also highlights the advantages, techniques of Big Data and illustrates the applications of Augmented Reality in various fields and domains.

VIII. REFERENCES

BOWMAN, D. A., & MCMAHAN, R. P. (2007). VIRTUAL REALITY: HOW MUCH IMMERSION IS ENOUGH? *COMPUTER*, *40*(7), 36–43. HTTPS://DOI.ORG/10.1109/MC.2007.257

CORDEIL, M., BACH, B., LI, Y., WILSON, E., & DWYER, T. (2017). DESIGN SPACE FOR SPATIO-DATA COORDINATION: TANGIBLE INTERACTION DEVICES FOR IMMERSIVE INFORMATION VISUALISATION. *IEEE PACIFIC VISUALIZATION SYMPOSIUM, MAY 2018*, 46–50. HTTPS://DOI.ORG/10.1109/PACIFICVIS.2017.8031578

KITCHIN, R. (2014). THE DATA REVOLUTION: BIG DATA, OPEN DATA, DATA INFRASTRUCTURES AND THEIR CONSEQUENCES. *SAGE*.

DONALEK, C., DJORGOVSKI, S. G., CIOC, A., WANG, A., ZHANG, J., LAWLER, E., YEH, S., MAHABAL, A., GRAHAM, M., DRAKE, A., DAVIDOFF, S., NORRIS, J. S., & LONGO, G. (2015). IMMERSIVE AND COLLABORATIVE DATA VISUALIZATION USING VIRTUAL REALITY PLATFORMS. *PROCEEDINGS - 2014 IEEE INTERNATIONAL CONFERENCE ON BIG DATA, IEEE BIG DATA 2014*, 609–614. HTTPS://DOI.ORG/10.1109/BIGDATA.2014.7004282

HELBIG, C. (2015). CONCEPT AND WORKFLOW FOR 3D VISUALIZATION OF MULTIFACETED METEOROLOGICAL DATA. NOVEMBER 1984.

HUANG, B., JIANG, B., & LI, H. (2001). AN INTEGRATION OF GIS, VIRTUAL REALITY AND THE INTERNET FOR VISUALIZATION, ANALYSIS AND EXPLORATION OF SPATIAL DATA. *INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE*, *15*(5), 439–456. HTTPS://DOI.ORG/10.1080/13658810110046574

WEXELBLAT, A. (ED.). (2014). VIRTUAL REALITY: APPLICATIONS AND EXPLORATIONS. *ACADEMIC PRESS*.

MARR, B. (2016). HOW VR WILL REVOLUTIONIZE BIG DATA VISUALIZATIONS. *FORBES ON LINE*.