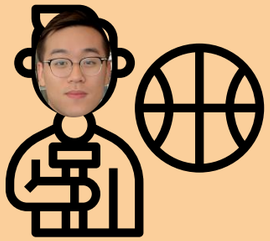
VIS2024

More Than Data Stories: Broadening the Role of Visualization in Contemporary Journalism

Yu Fu
Georgia Tech

John Stasko
Georgia Tech

Motivation



Journalism

采访/撰文：傅予
策划/编辑：王正宇

编者按：在休斯顿的三年，或许是PJ·塔克生涯里最绚烂魔幻的日子。在这里，他携手儿时的好友保罗，以1.95米的身高在禁区的肌肉丛林里厮杀，体验到了前所未有的曝光和舞台。在场上，他是最具竞争力的北卡硬汉。在场下，他又是名满天下的世界鞋王。对此，“不止篮球”内容团队通过视频连线采访到了塔克。面对过去，面对当下，他究竟说了些什么？

01 “这是一段让人永远铭记的日子。”

在一切重新开始之前，塔克需要让自己的内心平静下来。

他整理着刚刚才寄到奥兰多的球鞋，“已经差不多有一百双了，这两双刚刚才到，等到我离开的时候，应该会更多。”他在休斯敦的中城买下了一间公寓专门用来屯鞋，鞋盒围成了几堵墙，但过道里仍然零散地堆着还没来得及收拾的球鞋。塔克恨不得把所有宝贝们都挪到大佛罗里达人酒店的房间里来，这样他就能睡在自己的鞋床上，俯瞰窗外的七海泻湖。

没有篮球的四个多月里，这个世界发生了太多事情，每一件都远比篮球沉重。新冠病毒改变了所有人的生活，塔克也不例外。他人生当中第一次面对无球可打的情况。“没球打对我来说是最难的，”塔克回想起这四个月，“我从来没有这么久没打过球。”

“这段时间本来应该有各种篮球，但是因为新冠病毒，我生命当中第一次面对无球可打的情况……这是一段让人永远铭记的时期，往后的一切都会不一样了。”

从这一切回归到篮球，并非易事——有人会担心健康，有人更在意家人，有人希望自己的声音不被盖过：也有人抱怨，抱怨饮食不够精良，抱怨戴口罩多此一举……

塔克理解这一切，“那仅仅是隔离餐，”提到园区内的伙食，塔克跟我解释到，“隔离餐会被单独放在袋子里，你不能跟任何人接触，然后直接放在房间门口。我们只有到那的前48小时会吃隔离餐，其实也没什么好抱怨的。”

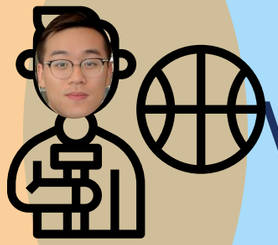
“这些都不是什么大问题，的确有些不太方便。但（没有观众）对我来说也并不是什么问题。我们这些球员从小就适应了在没有观众的情况下打球。我的孩子们还能在电视上看我打球。是的，离开他们很难，但是过去四个月我都陪在他们身边，他们会没事的。”

“联盟已经做得很不错了。在短短的时间之内就把这些布置规划好……所有设施都安排妥当，安全，各种消毒，清洁。让我们每名球员都能放松下来，专注在篮球上。”

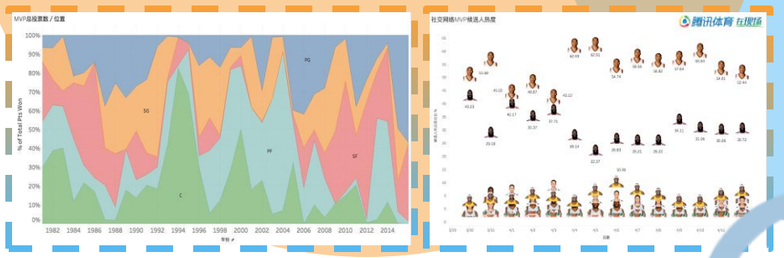
这并没有夸张，NBA的确对奥兰多复赛园区进行了严格的管控，上周的检疫报告，已在园区内322名球员没有一名感染。考虑到美国，尤其是奥兰多所在州佛罗里达的现状，着实难得。

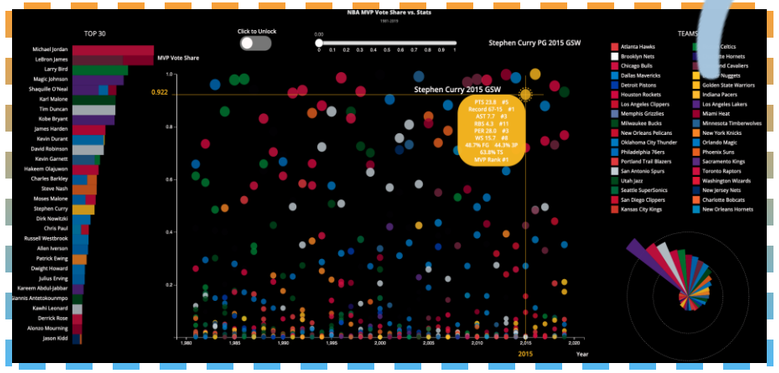
Motivation

Journalism



Visualization



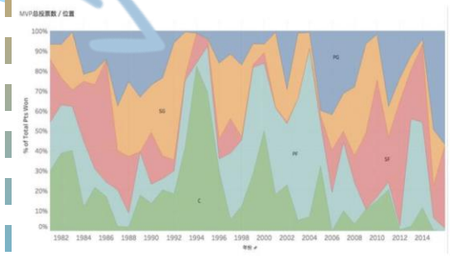


新闻 篮球世界杯 球员 球队 赛程 数据 社区

如果说总冠军是联盟最高的团队荣誉，那么MVP则是联盟最高的个人荣誉。

MVP的得主往往是那个年代里联盟最有统治力的球员，至少也是联盟的招牌球员之一，不像总决赛MVP那么变化多端。

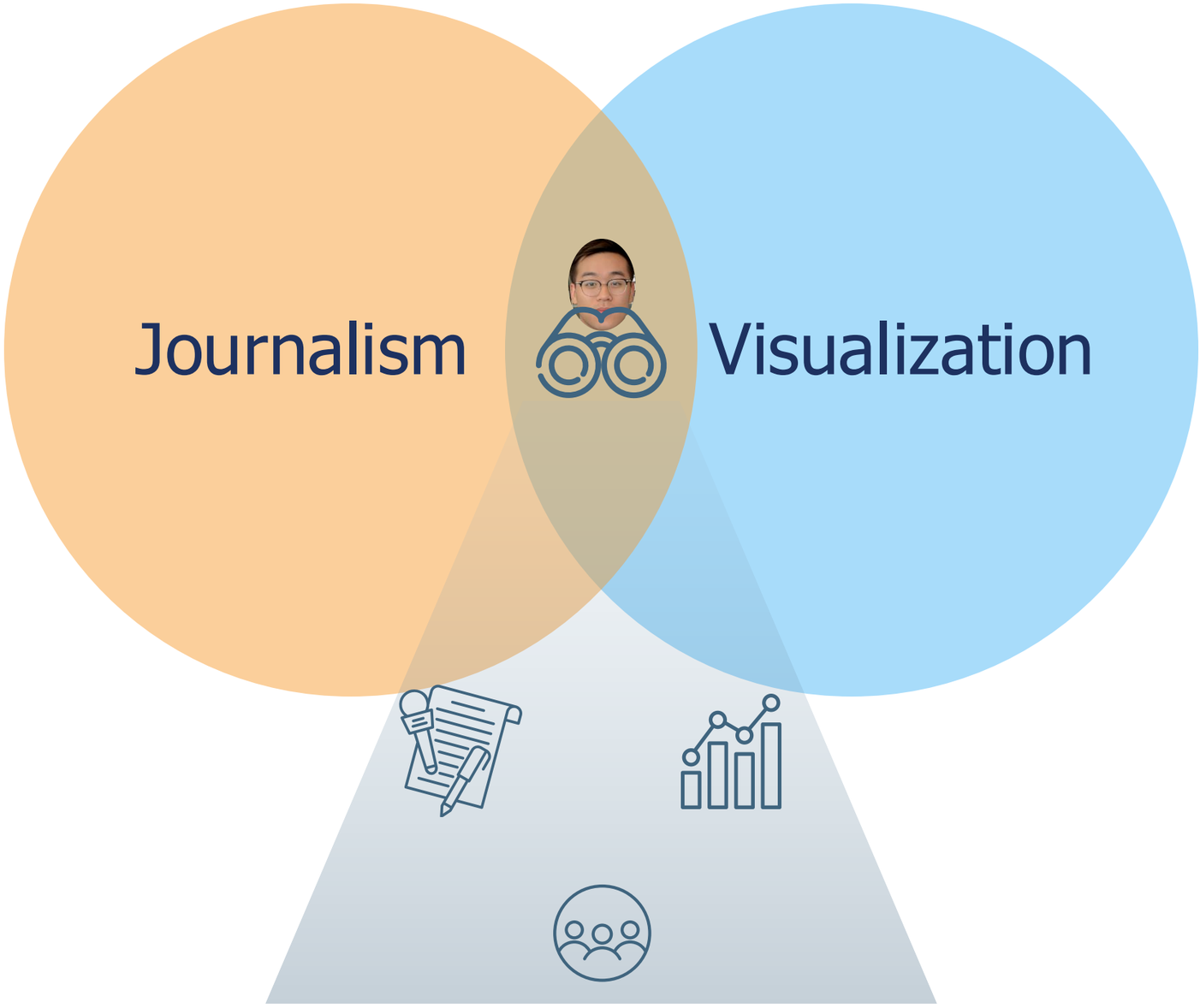
如果我们再进一步，不仅仅关注MVP得主这个奖项，而是细看每年的MVP投票率，我们能获得更多的信息，能帮助我们更了解一个时代，看清NBA整个联盟的变迁。



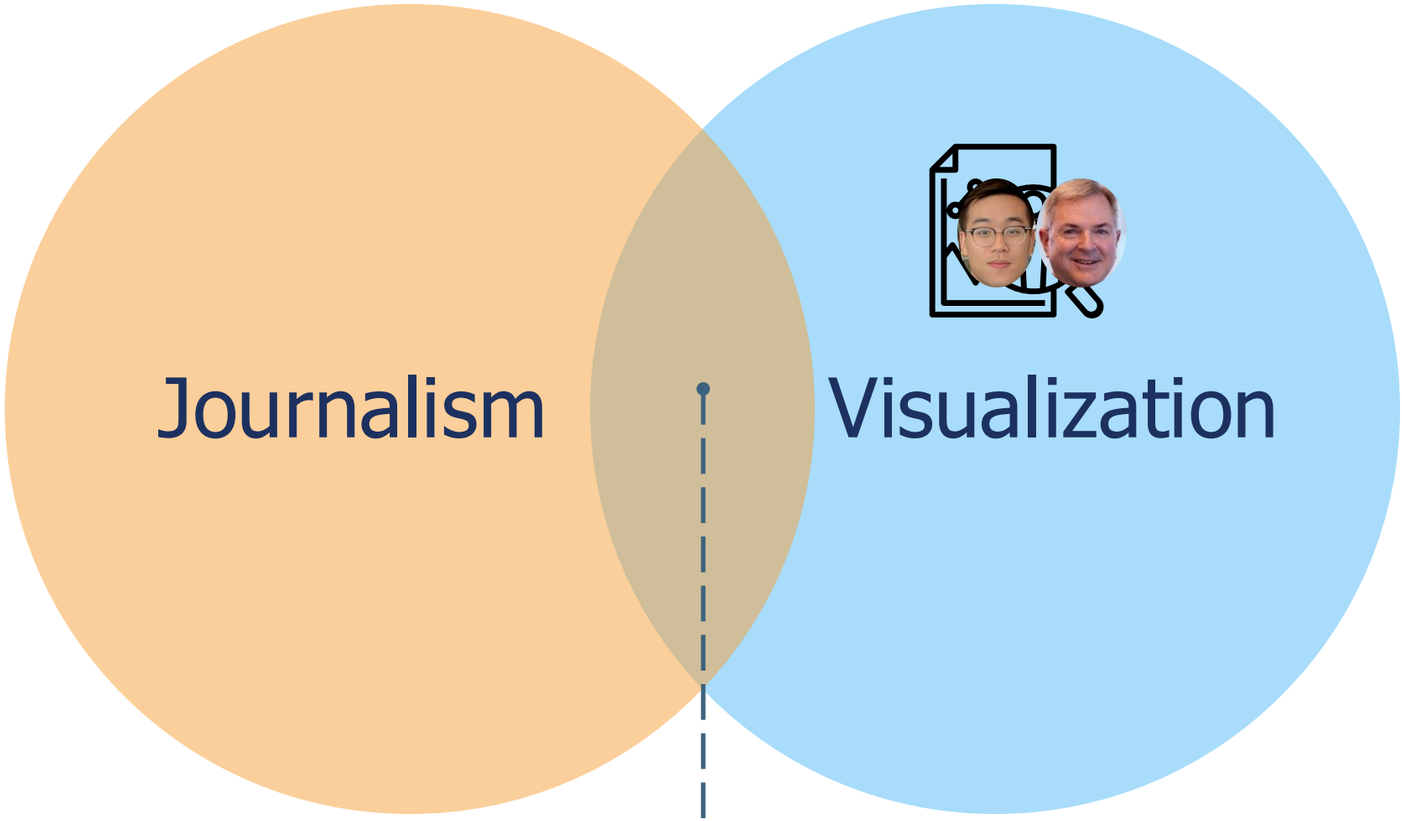
各位置球员MVP得票所占比例（控卫：蓝色、分卫：成色、小前锋：红色、大前锋：浅蓝色、中锋：绿色）

比较有代表性的是1993-94赛季乔丹第一次退役，得分后卫在MVP总票数所在百分比，从92年的接近百分之六十，跌倒94年几乎为零，虽然那从80年代后期到93年得分后卫（图中橙色）这个位置上的球员所占比例很大，但是几乎是靠乔丹一己之力撑起来的，当然德雷克斯勒也做出了一定贡献，但那个年代，不是得分后卫的年代，是中锋的年代。

Motivation



Motivation



- What new visualization research opportunities could emerge at this intersection?
- How could visualization research contextualize its work to address broader challenges in journalism?

Methodology
(scoping review)

Journalism

Publication Venue	Count
Digital Journalism	11
Journalism Practice	4
New Media & Society	3
Journalism	2
Journalism Studies	2
Book Chapters/others	72
Total	**94**

- What recent shifts and challenges has journalism faced?
- How is journalism integrating computational technology?

Visualization

Publication Venue	Count
IEEE TVCG	37
ACM CHI	16
ACM UIST	4
IEEE CGA	4
ACM BELIV	4
Others	28
Total	**93**

- What value does visualization provide?
- How can we contextualize visualization research to support journalism?

Journalistic Transformations and Challenges

Transformations	Challenges
New Media & Interactive Journalism	• Requires skills not traditionally acquired by journalists • Audience may not be aware or utilize interactivity

Journalistic Transformations and Challenges

Transformations	Challenges
New Media & Interactive Journalism	• Requires skills not traditionally acquired by journalists • Audience may not utilize interactivity
Participatory Environment	• Weakens journalism's gatekeeping role • Lowers barriers, enabling misinformation • Challenges in fostering coherent online discussion

Journalistic Transformations and Challenges

Transformations	Challenges
New Media & Interactive Journalism	• Requires skills not traditionally acquired by journalists • Audience may not utilize interactivity
Participatory Environment	• Weakens journalism's gatekeeping role • Lowers barriers, enabling misinformation • Challenges in fostering coherent online discussion
Loaded & Polluted Infosphere	• Public sphere flooded with skepticism • Audiences struggle to discern trustworthy info

Journalistic Transformations and Challenges

Transformations	Challenges
New Media & Interactive Journalism	• Requires skills not traditionally acquired by journalists • Audience may not utilize interactivity
Participatory Environment	• Weakens journalism's gatekeeping role • Lowers barriers, enabling misinformation • Challenges in fostering coherent online discussion
Loaded & Polluted Infosphere	• Public sphere flooded with skepticism • Audiences struggle to discern trustworthy info
News Personalization	• Filtering and biases can worsen the "filter bubble" effect

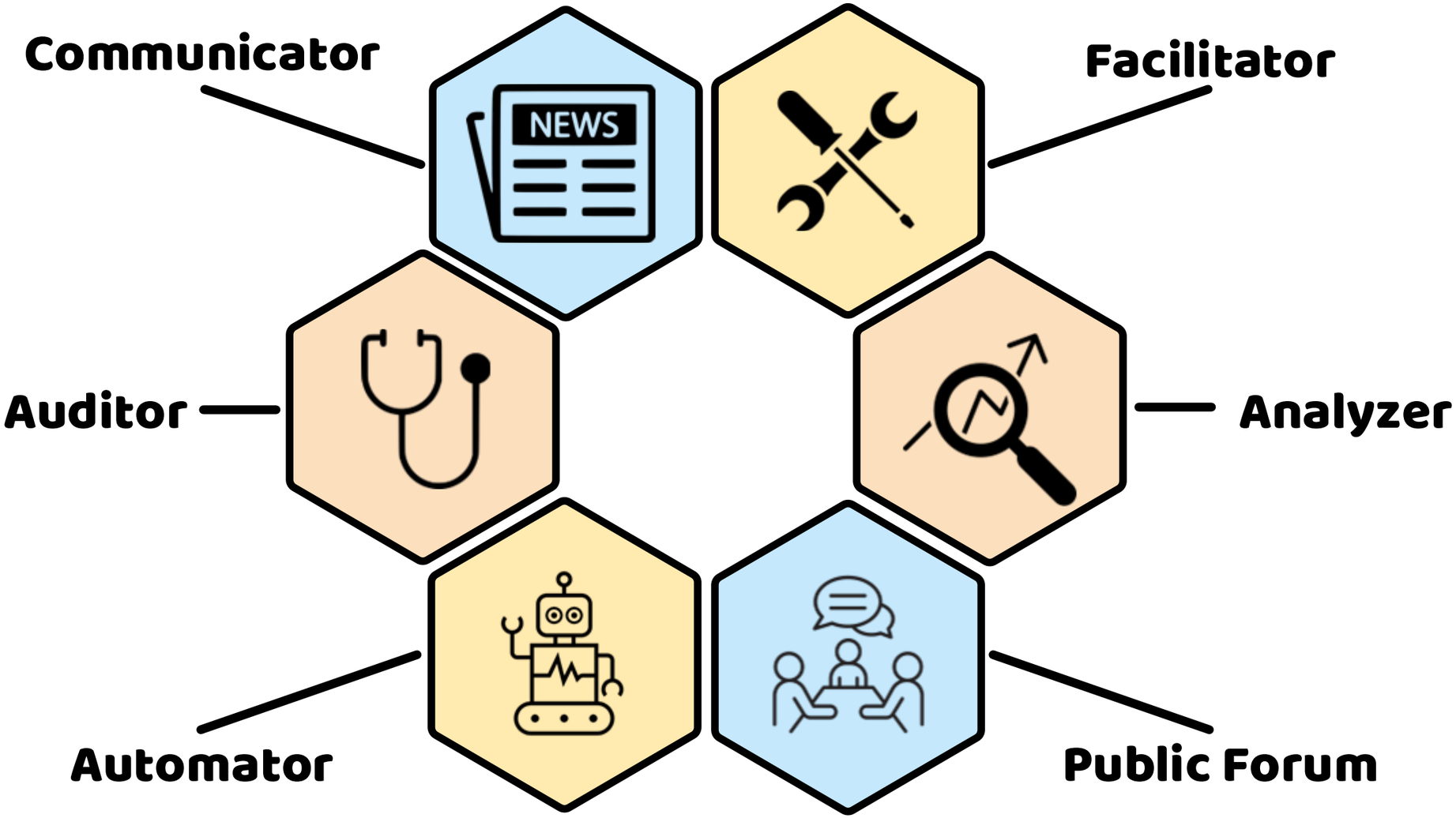
Journalistic Transformations and Challenges

Transformations	Challenges
New Media & Interactive Journalism	• Requires skills not traditionally acquired by journalists • Audience may not utilize interactivity
Participatory Environment	• Weakens journalism's gatekeeping role • Lowers barriers, enabling misinformation • Challenges in fostering coherent online discussion
Loaded & Polluted Infosphere	• Public sphere flooded with skepticism • Audiences struggle to discern trustworthy info
News Personalization	• Filtering and biases can worsen the "filter bubble" effect
Generative AI	???

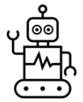
Journalism’s Computational Turn

Forms	Perspectives
Computer-assisted Reporting (CAR)	• Leverages investigative techniques, scientific methods, and data management to enhance **credibility** and **analysis**.
Data Journalism	• Places **data** at the **core of storytelling**, empowering public access, exploration, and participation in data-driven analysis.
Computational Journalism	• **Automates** journalistic tasks • Focuses on **transparency** and **accountability** in journalism's algorithms

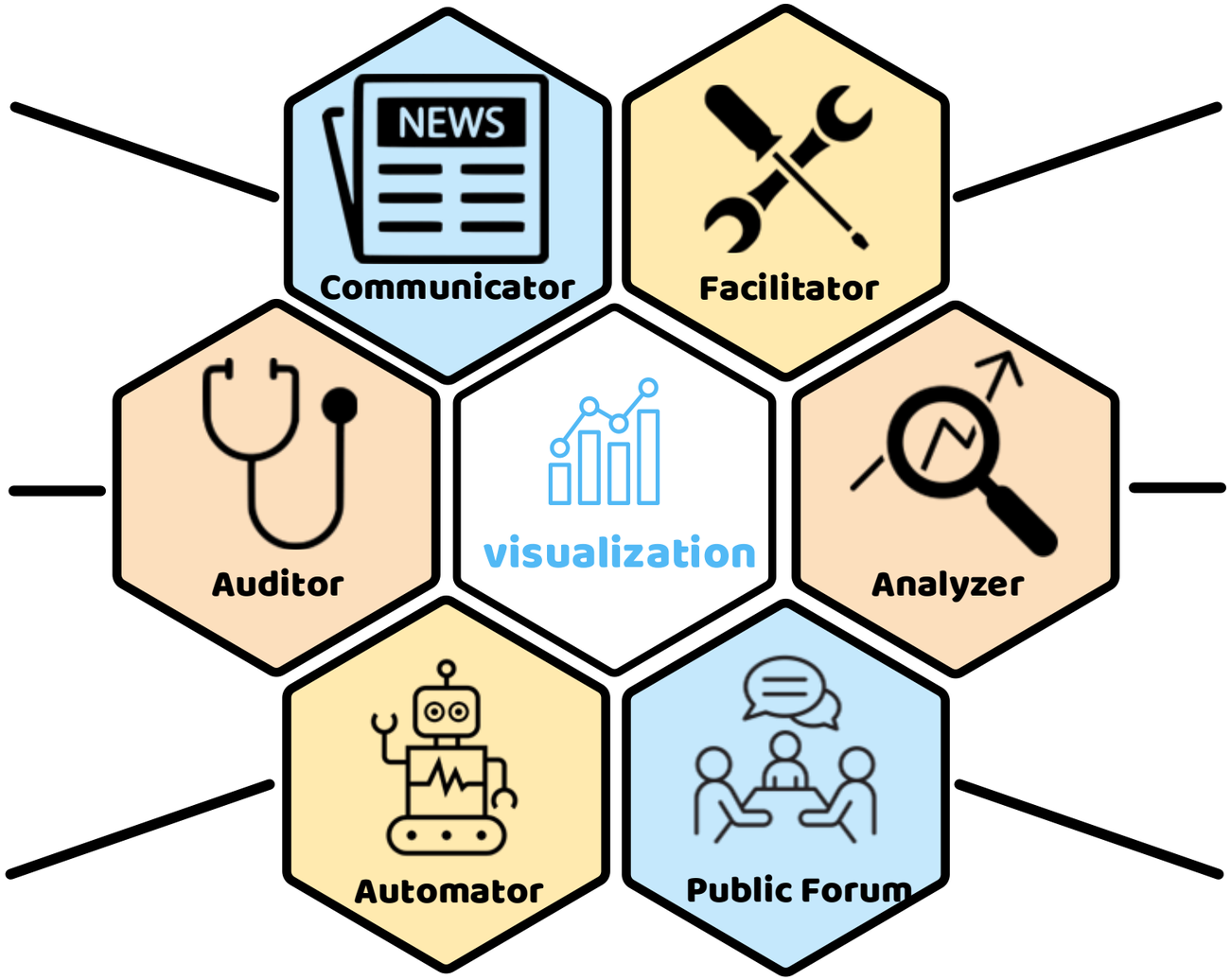
Six Roles of Computing in Journalism



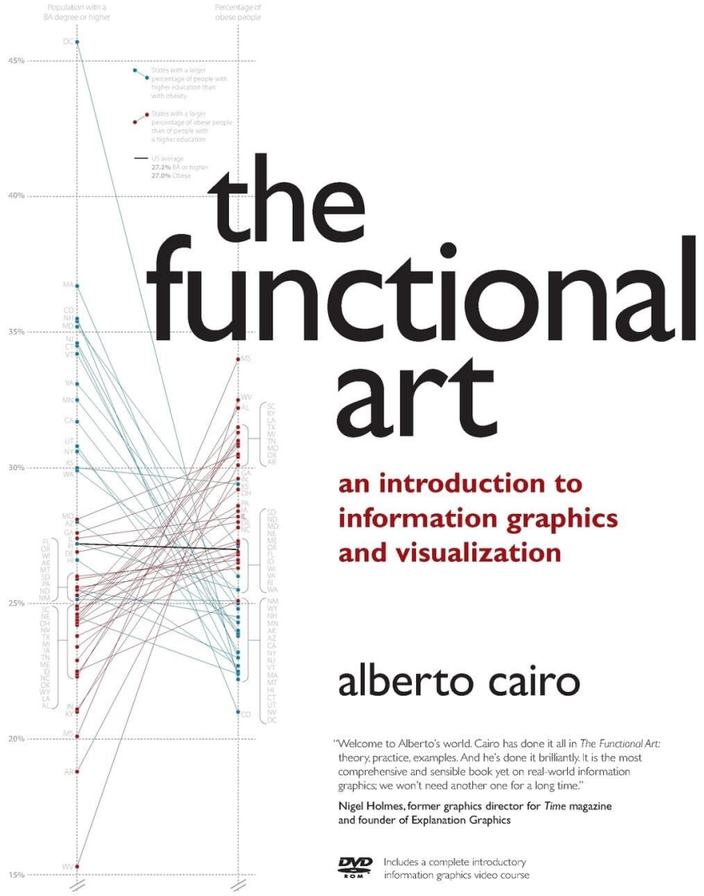
Six Roles of Computing in Journalism

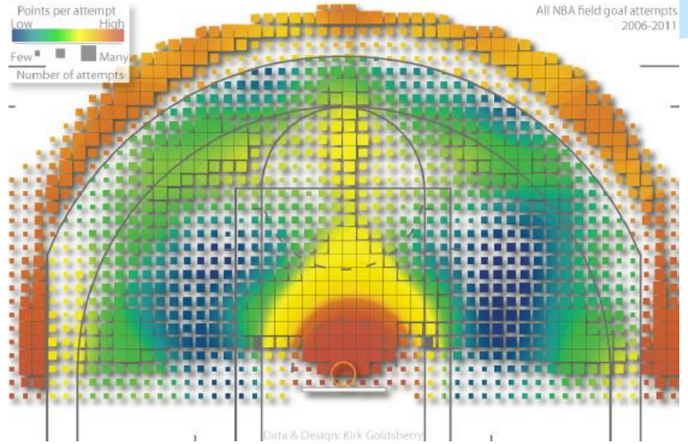
Roles	Definitions
Communicator 	Digital artifacts that communicate insights or narratives to audiences
Facilitator 	Tools that support news production tasks, such as content creation, news gathering, and fact-checking.
Analyzer 	Tools that help analyze information and uncover insights or narratives
Public Forum 	Platforms that foster dialogue and the exchange of ideas among the public.
Automator 	Computational technology (e.g., algorithms, AI) that automates journalistic tasks, such as content creation and news delivery
Auditor 	Technology that uncovers the effect of computing artifacts and audits their accuracy, credibility, and biases.

How can **Visualization** support or be strengthened by these roles?

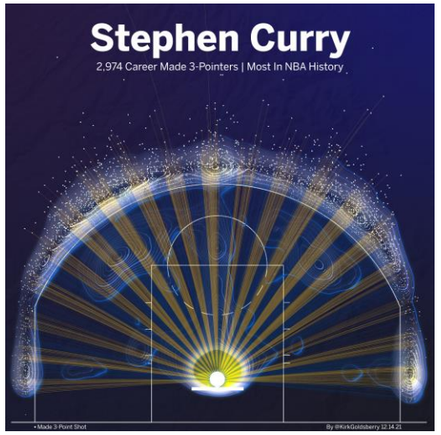


What **values** does **Visualization** provide in this context?

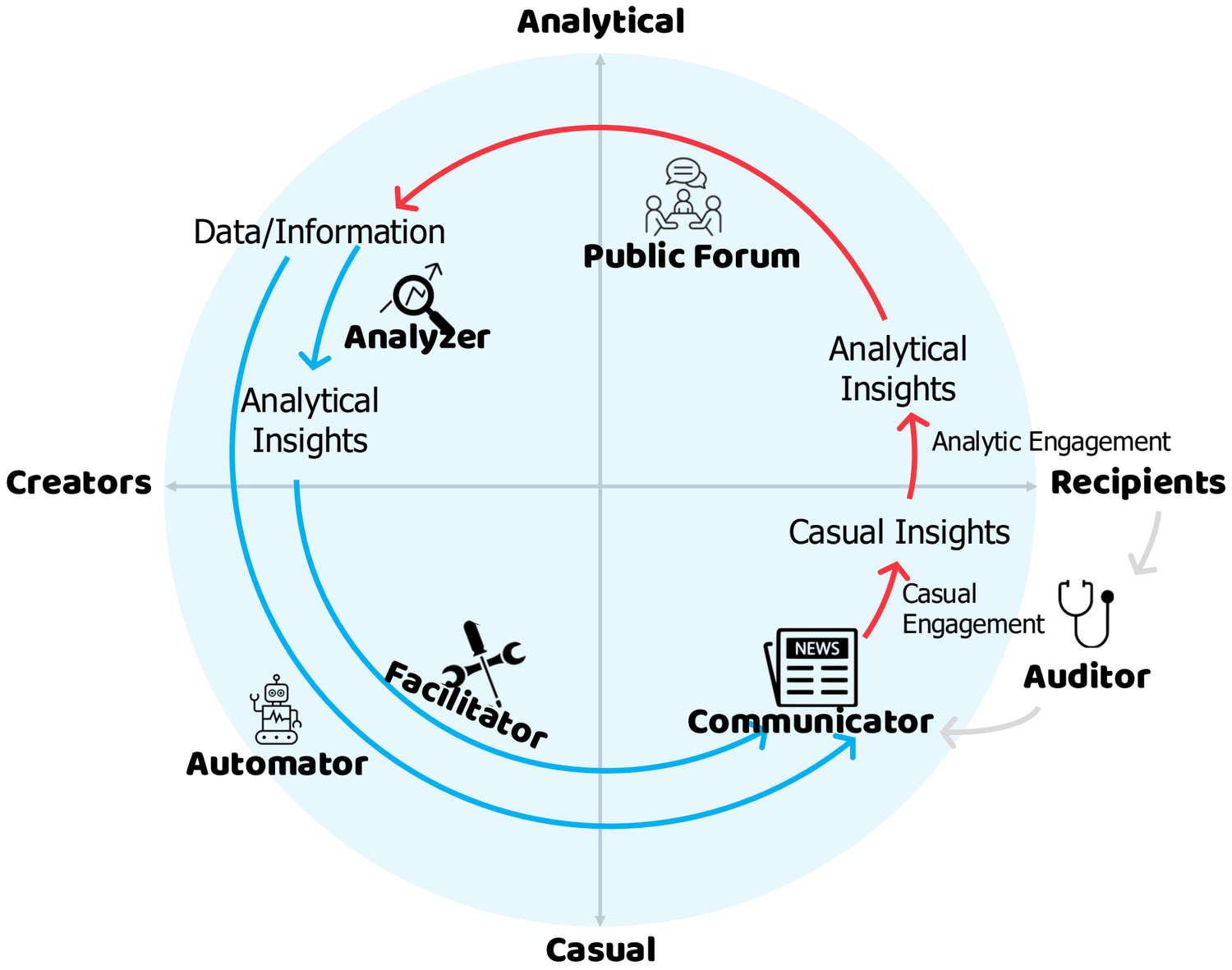


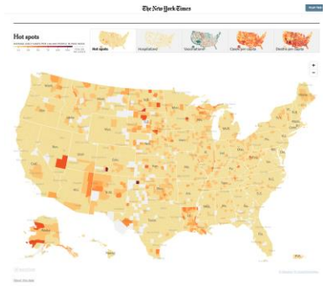


CourtVision: NBA Visual and Spatial Analytics (2012)
Kirk Goldsberry



Kirk Goldsberry (2022)

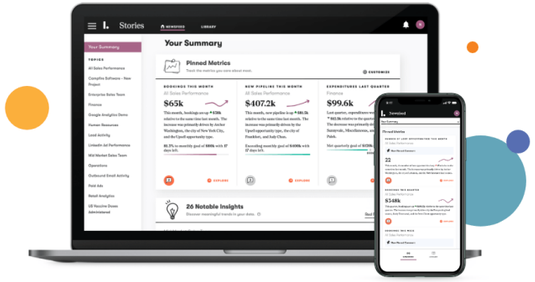

Analytical
Data/Information
Public Forum
Analyzer
Analytical Insights
Analytical Insights
Analytic Engagement
Creators
Recipients
Casual Insights
Casual Engagement
Auditor
NEWS
Facilitator
Communicator
Automator
Casual

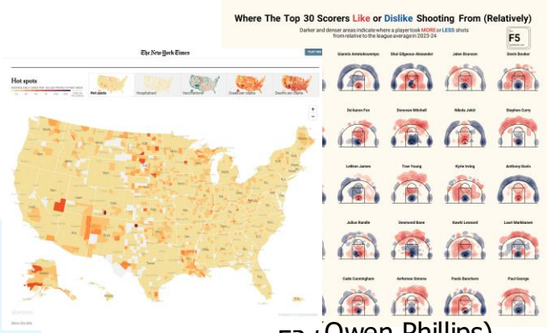


New York Times Covid Dashboard

FiveThirty

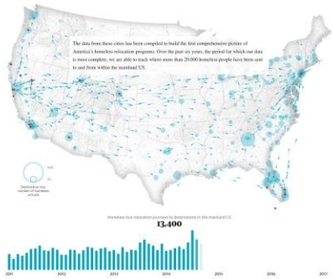
narrative science



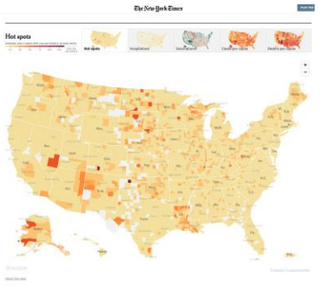


New York Times Covid Dashboard

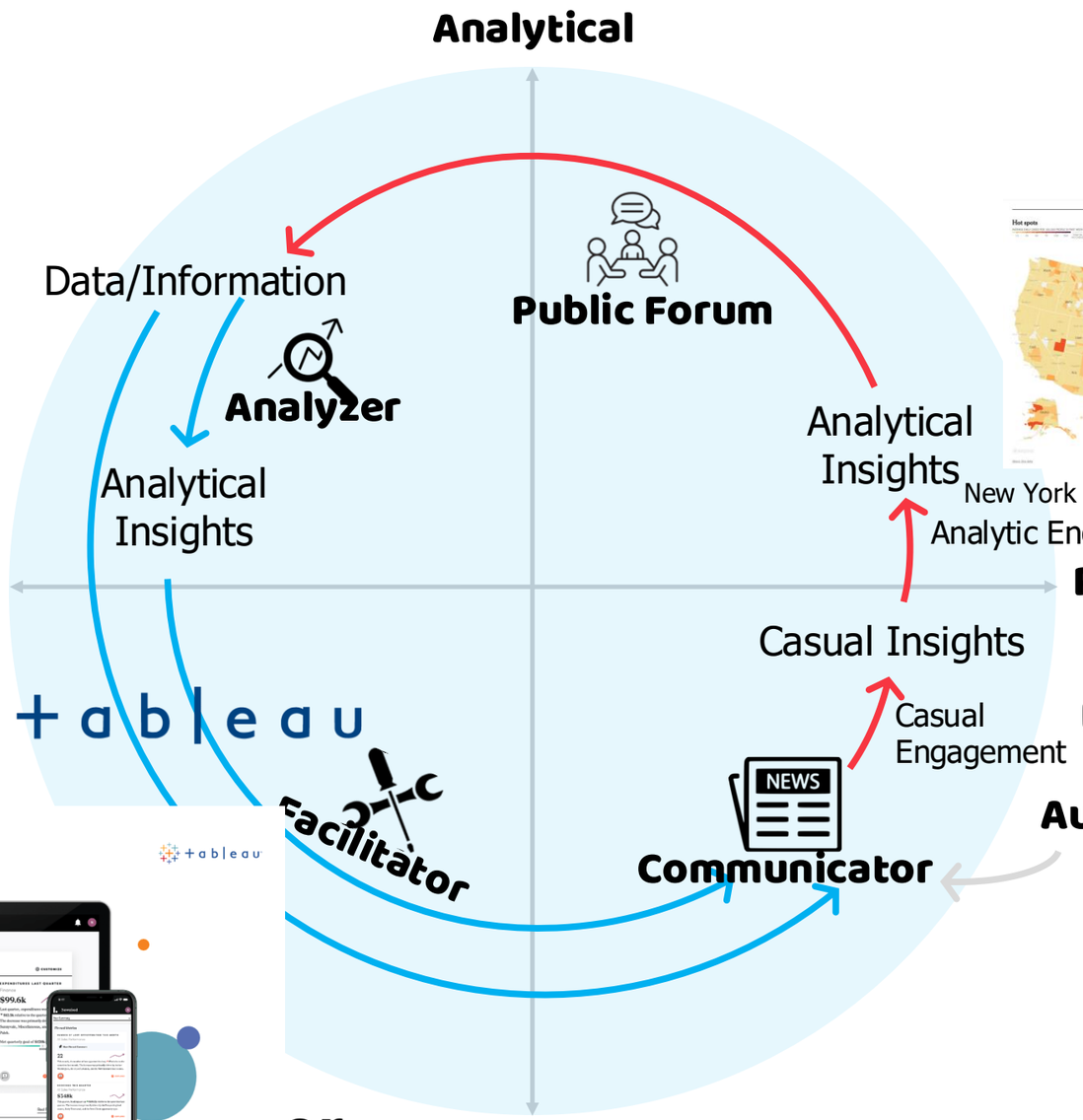
F5 (Owen Phillips)



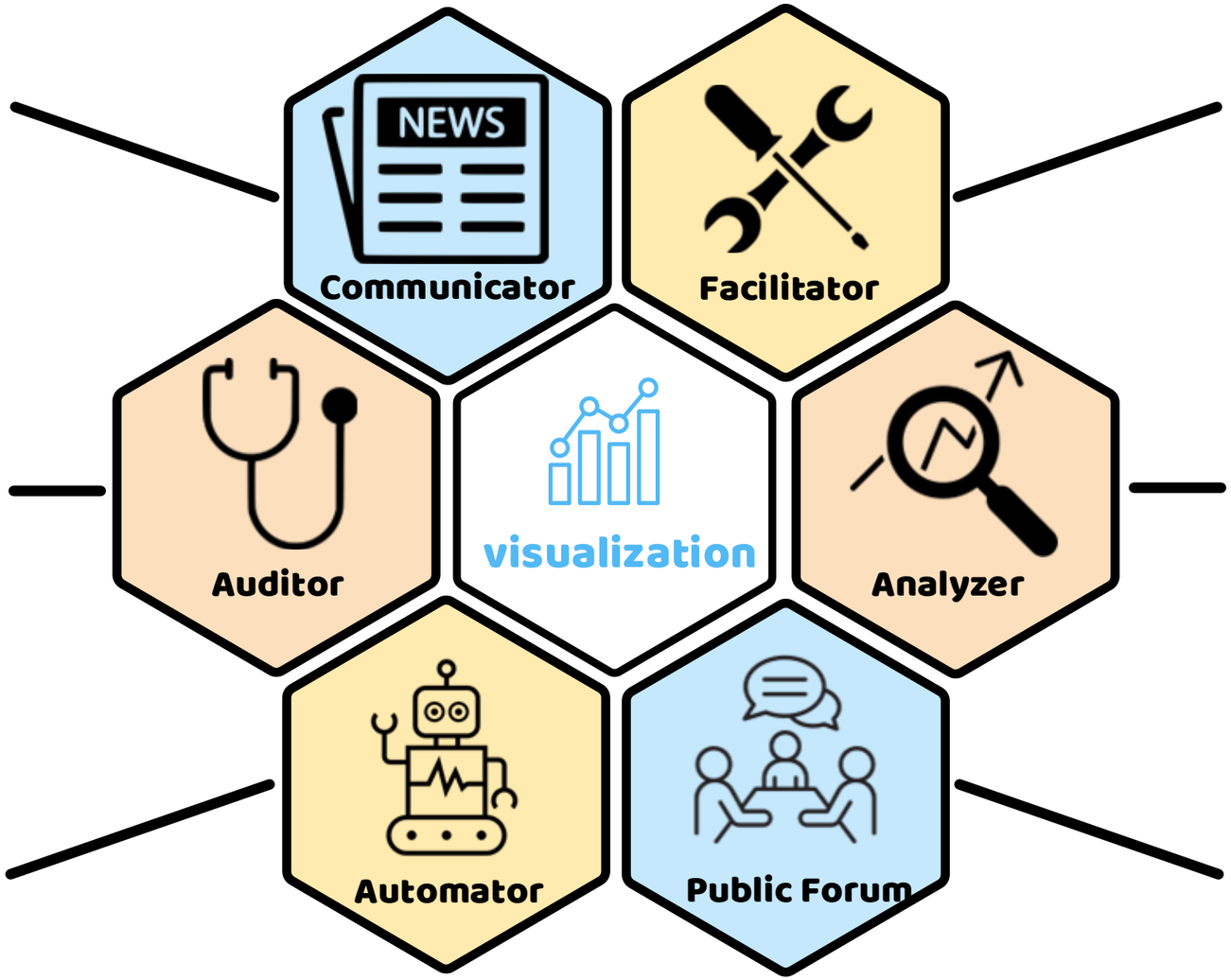
The Guardian (Nadieh Bremer et al.)



New York Times Covid Dashboard

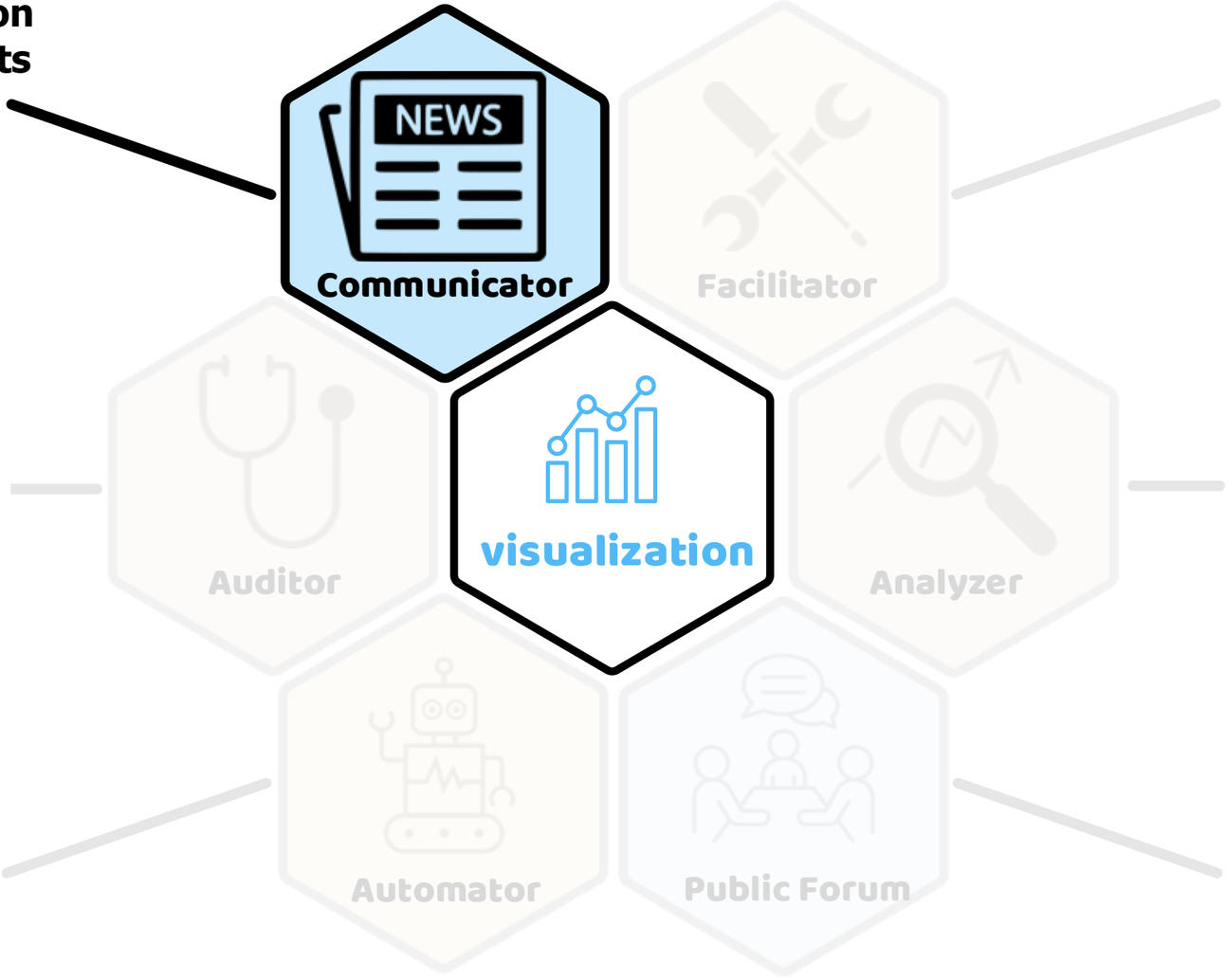


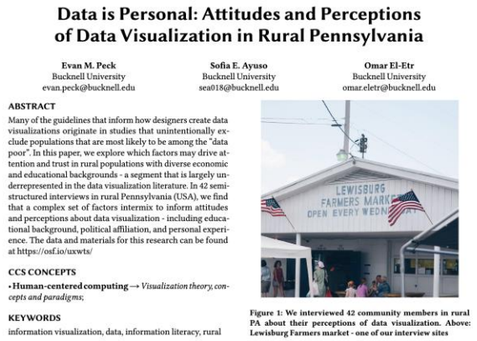
Agendas for **Visualization Research**



Agendas for **Visualization Research**

- **The design of visualization artifacts and their impacts on broader audiences**





Data is Personal: Attitudes and Perceptions of Data Visualization in Rural Pennsylvania

Evan M. Peck, Bucknell University, evan.peck@bucknell.edu
Sofia E. Ayuso, Bucknell University, sea018@bucknell.edu
Omar El-Etr, Bucknell University, omar.eletr@bucknell.edu

ABSTRACT

Many of the guidelines that inform how designers create data visualizations originate in studies that unintentionally exclude populations that are most likely to be among the "data poor". In this paper, we explore which factors may drive attention and trust in rural populations with diverse economic and educational backgrounds - a segment that is largely underrepresented in the data visualization literature. In 42 semi-structured interviews in rural Pennsylvania (USA), we find that a complex set of factors intermix to inform attitudes and perceptions about data visualization - including educational background, political affiliation, and personal experience. The data and materials for this research can be found at https://osf.io/uxwts/

CCS CONCEPTS

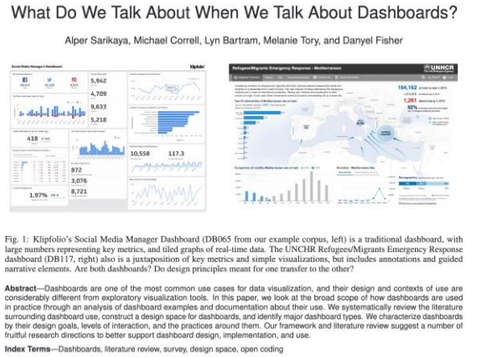
• **Human-centered computing** → *Visualization theory, concepts and paradigms;*

KEYWORDS

information visualization, data, information literacy, rural

Figure 1: We interviewed 42 community members in rural PA about their perceptions of data visualization. Above: Lewisburg Farmers market - one of our interview sites

Peck et al. 2019



What Do We Talk About When We Talk About Dashboards?

Alper Sarikaya, Michael Correll, Lyn Bartram, Melanie Tory, and Danyel Fisher

Fig. 1: Klipfolio's Social Media Manager Dashboard (DB065 from our example corpus, left) is a traditional dashboard, with large numbers representing key metrics, and tiled graphs of real-time data. The UNCHR Refugees/Migrants Emergency Response dashboard (DB117, right) also is a juxtaposition of key metrics and simple visualizations, but includes annotations and guided narrative elements. Are both dashboards? Do design principles meant for one transfer to the other?

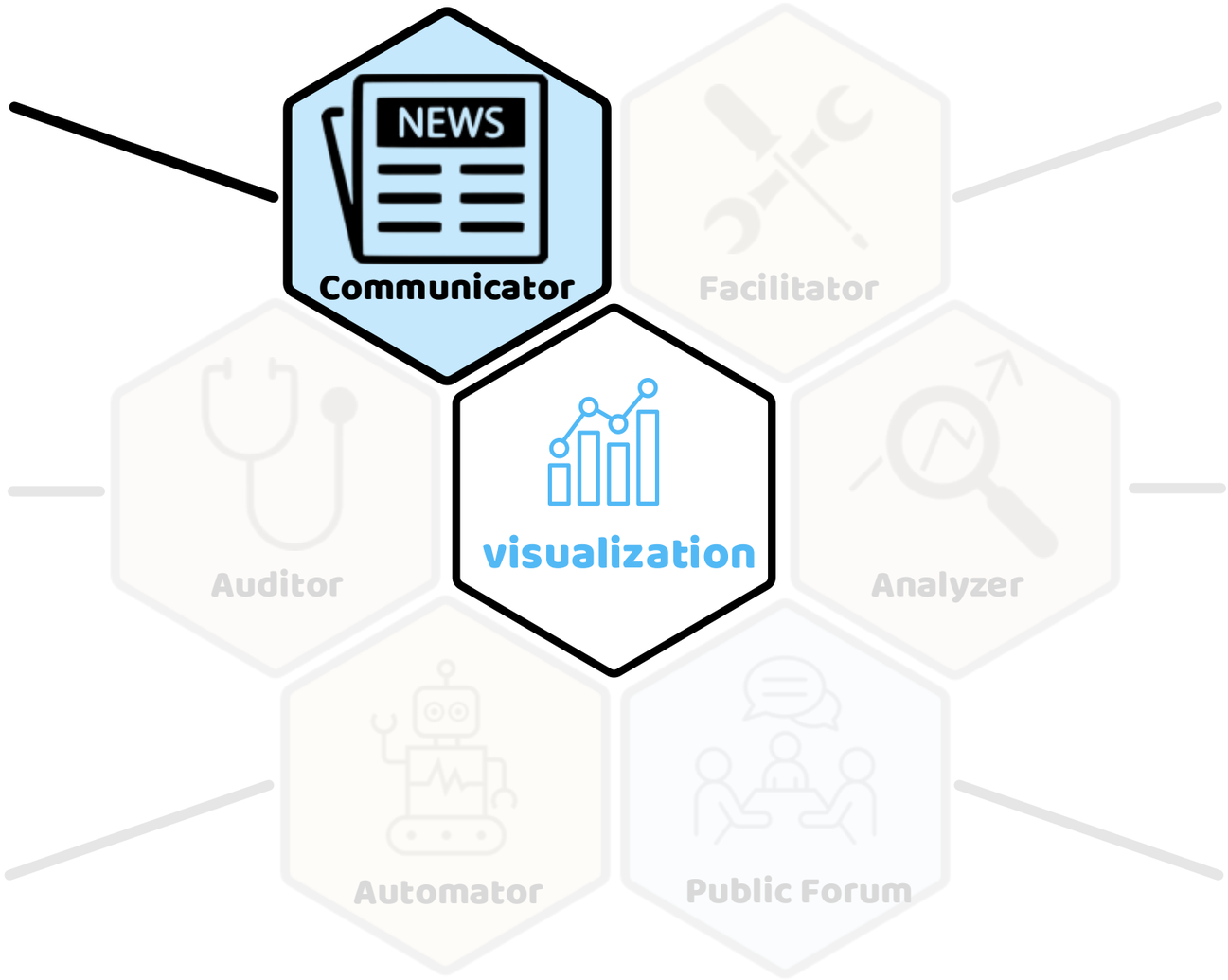
Abstract—Dashboards are one of the most common use cases for data visualization, and their design and contexts of use are considerably different from exploratory visualization tools. In this paper, we look at the broad scope of how dashboards are used in practice through an analysis of dashboard examples and documentation about their use. We systematically review the literature surrounding dashboard use, construct a design space for dashboards, and identify major dashboard types. We characterize dashboards by their design goals, levels of interaction, and the practices around them. Our framework and literature review suggest a number of fruitful research directions to better support dashboard design, implementation, and use.

Index Terms—Dashboards, literature review, survey, design space, open coding

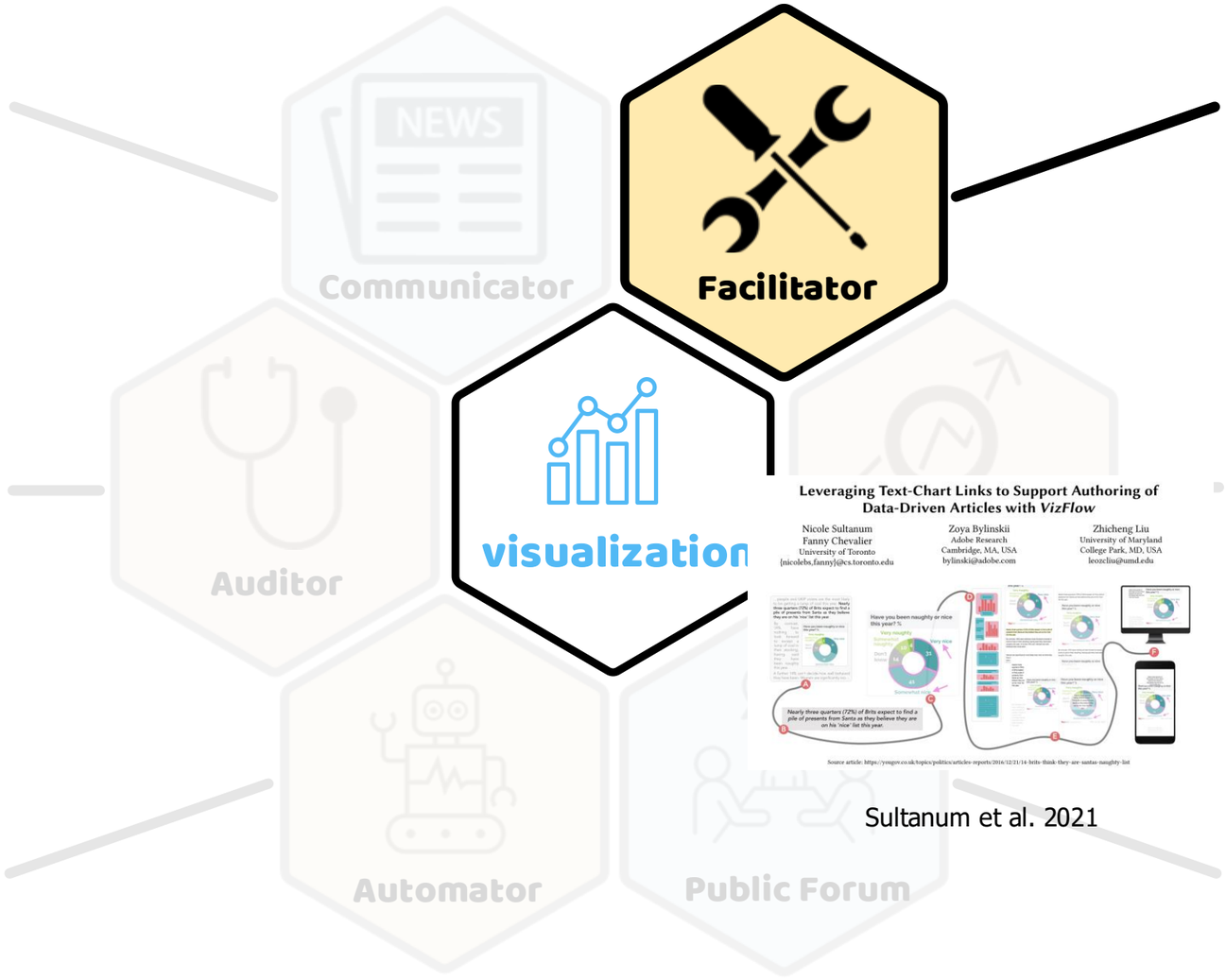
Sarikaya et al. 2019

Agendas for **Visualization Research**

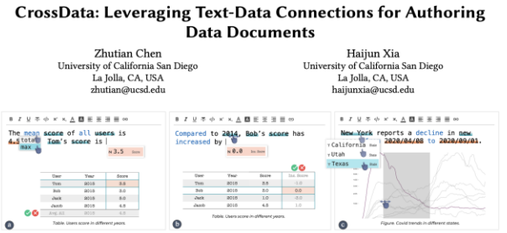
- **Understand casual engagement and analytical engagement**



Agendas for **Visualization Research**



- **Facilitate Data-driven Story Authoring**



CrossData: Leveraging Text-Data Connections for Authoring Data Documents

Zhutian Chen
University of California San Diego
La Jolla, CA, USA
zhutian@ucsd.edu

Haijun Xia
University of California San Diego
La Jolla, CA, USA
haijunxia@ucsd.edu

Figure 1: CrossData leverages text-data connections to enable users to efficiently retrieve (a), compute (b), interactively explore data (a, b, c), and adjust tables (a, b) and charts (c) during their writing processes, while also automatically maintaining data consistency between their text, data, tables, and charts.

Chen and Xia. 2022

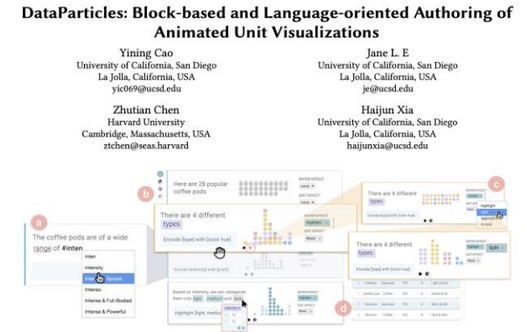
Leveraging Text-Chart Links to Support Authoring of Data-Driven Articles with *VizFlow*

Nicole Sultanum, Fanny Chevalier
University of Toronto
{nicolebs,fanny}@cs.toronto.edu

Zoya Bylinskii
Adobe Research
Cambridge, MA, USA
bylinski@adobe.com

Zhicheng Liu
University of Maryland
College Park, MD, USA
leozcliu@umd.edu

Sultanum et al. 2021



DataParticles: Block-based and Language-oriented Authoring of Animated Unit Visualizations

Yining Cao
University of California, San Diego
La Jolla, California, USA
yic069@ucsd.edu

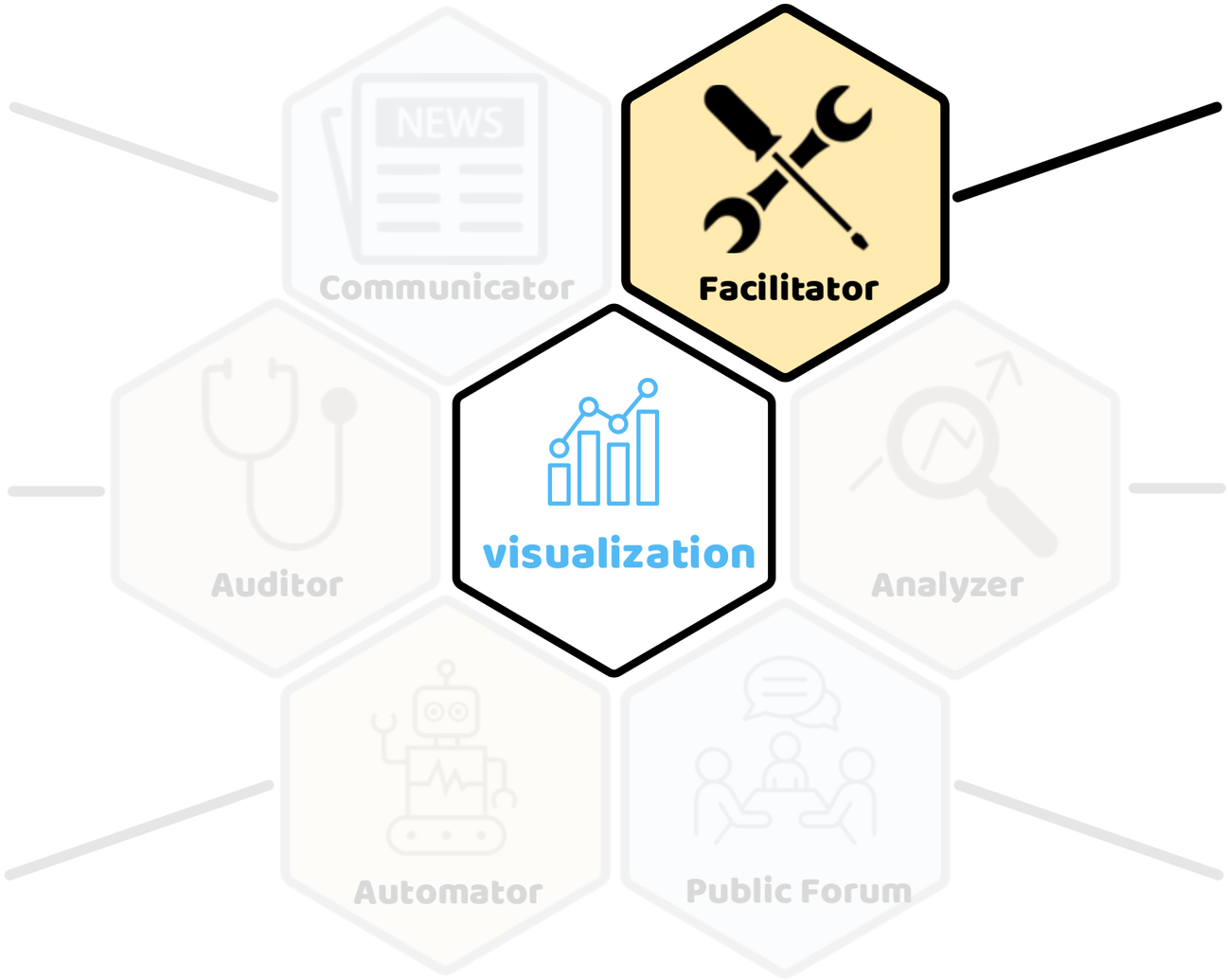
Jane L. E
University of California, San Diego
La Jolla, California, USA
je@ucsd.edu

Zhutian Chen
Harvard University
Cambridge, Massachusetts, USA
ztchen@seas.harvard

Haijun Xia
University of California, San Diego
La Jolla, California, USA
haijunxia@ucsd.edu

Cao et al. 2023

Agendas for **Visualization Research**



- **Bridge the gap between exploration and presentation**

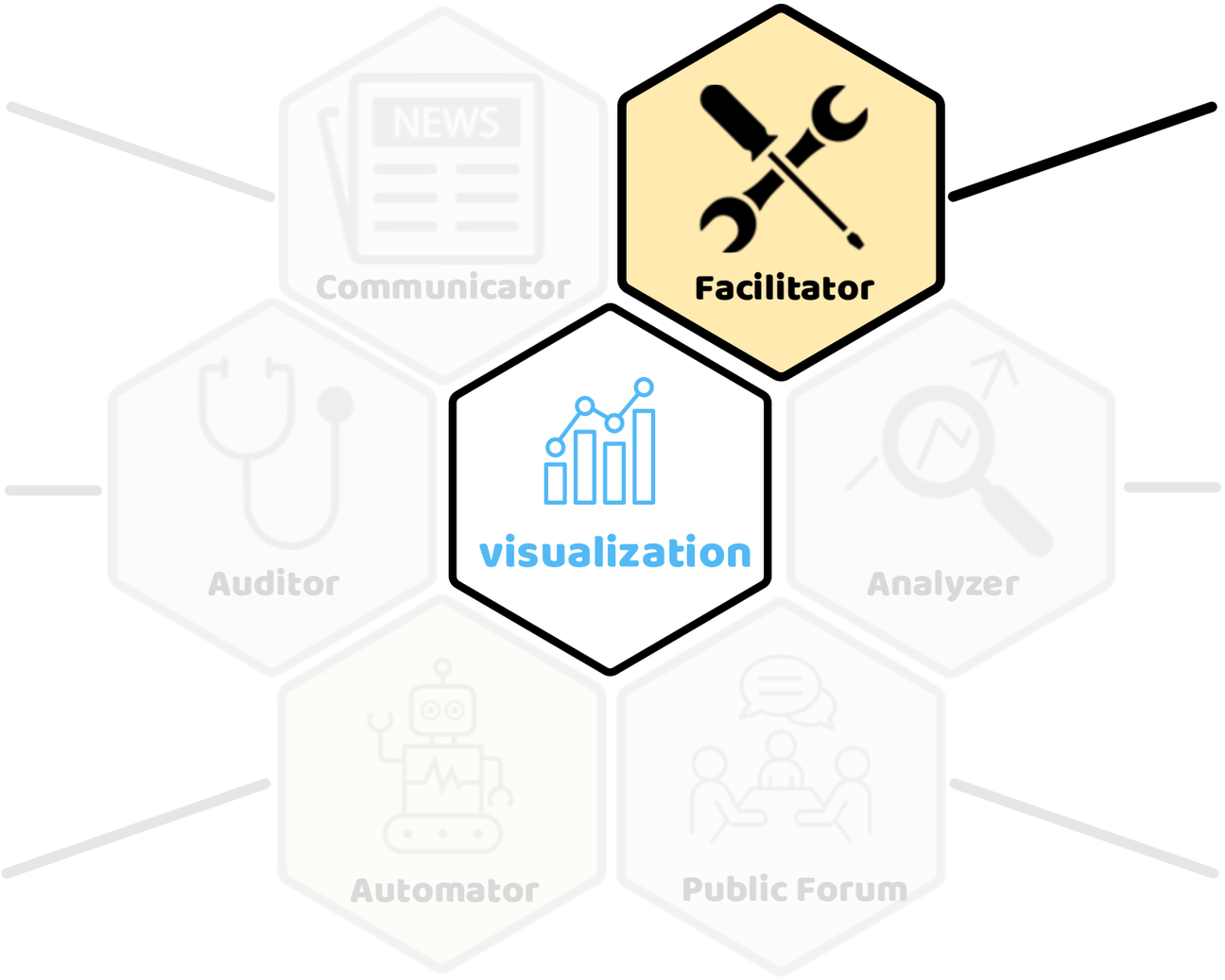
Supporting Story Synthesis: Bridging the Gap between Visual Analytics and Storytelling

Siming Chen, Jie Li, Gennady Andrienko, Natalia Andrienko, Yun Wang, Phong H. Nguyen, and Cagatay Turkay

Abstract—Visual analytics usually deals with complex data and uses sophisticated algorithmic, visual, and interactive techniques supporting the analysis. Findings and results of the analysis often need to be communicated to an audience that lacks visual analytics expertise. This requires analysis outcomes to be presented in simpler ways than that are typically used in visual analytics systems. However, not only analytical visualizations may be too complex for target audiences but also the information that needs to be presented. Analysis results may consist of multiple components, which may involve multiple heterogeneous facets. Hence, there exists a gap on the path from obtaining analysis findings to communicating them, within which two main challenges lie: information complexity and display complexity. We address this problem by proposing a general framework where data analysis and result presentation are linked by story synthesis, in which the analyst creates and organises story contents. Unlike previous research, where analytic findings are represented by stored display states, we treat findings as data constructs. We focus on selecting, assembling and organizing findings for further presentation rather than on tracking analysis history and enabling dual (i.e., explorative and communicative) use of data displays. In story synthesis, findings are selected, assembled, and arranged in meaningful layouts that take into account the structure of information and inherent properties of its components. We propose a workflow for applying the proposed conceptual framework in designing visual analytics systems and demonstrate the generality of the approach by applying it to two diverse domains, social media and movement analysis.

Chen et al. 2020

Agendas for **Visualization Research**



- **Facilitate data-driven storytelling in new media (e.g., Video, immersive, live streaming)**

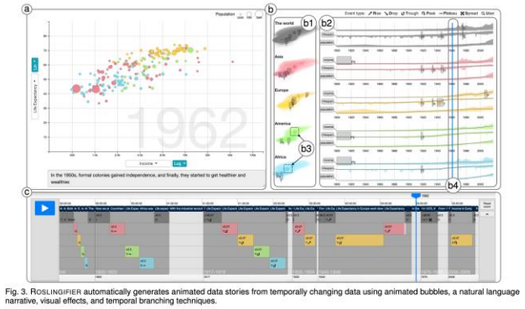
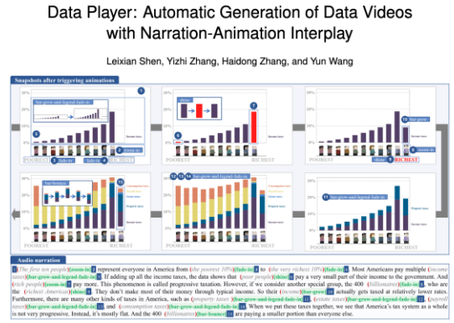


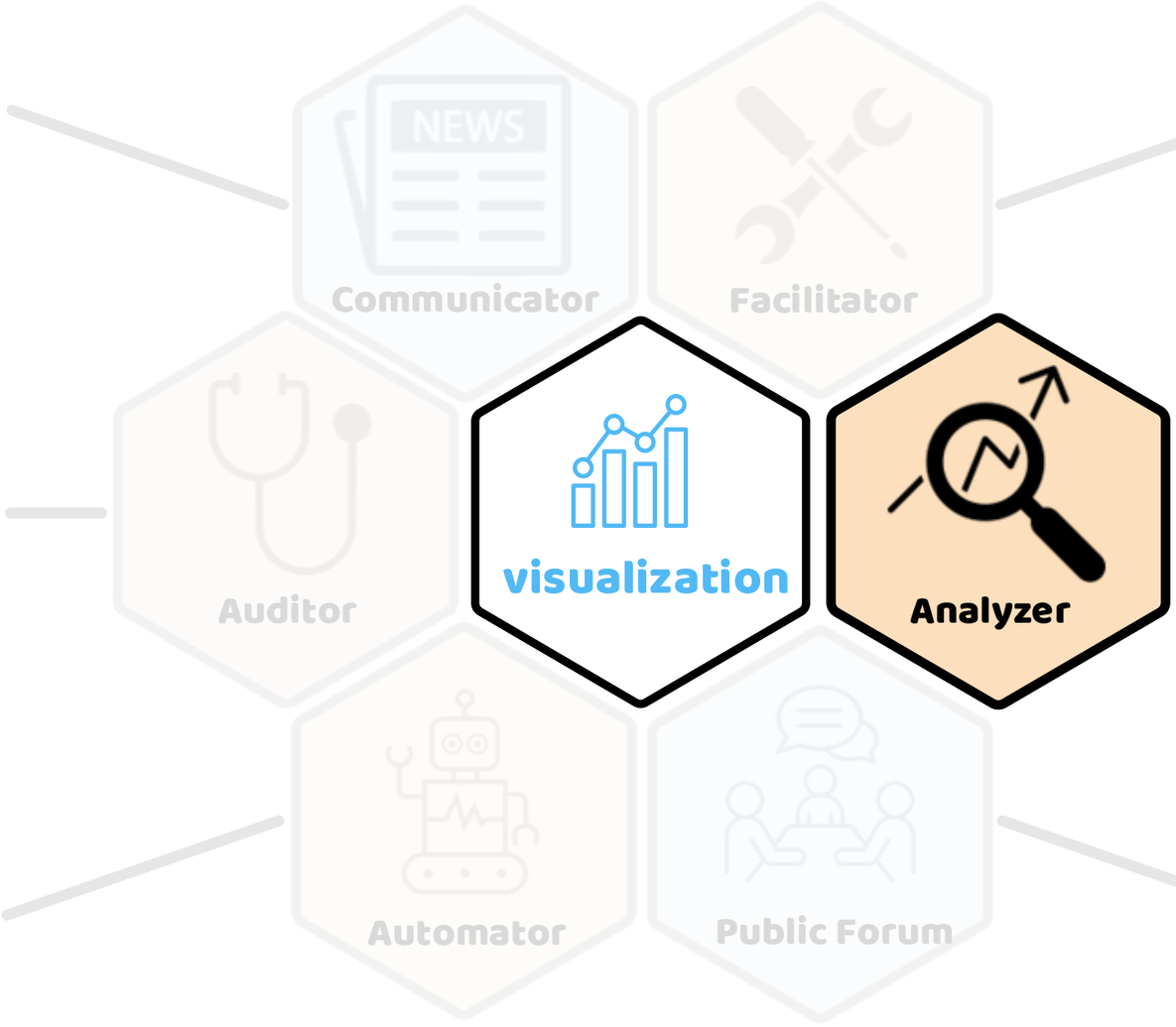
Fig. 3. ROSLINGIFIER automatically generates animated data stories from temporally changing data using animated bubbles, a natural language narrative, visual effects, and temporal branching techniques.

Roslingifier (Shin et al. 2021)

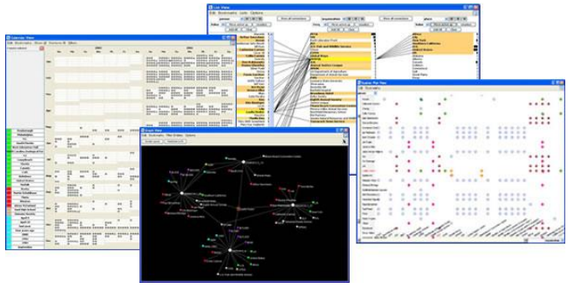


Data Player (Shen et al. 2023)

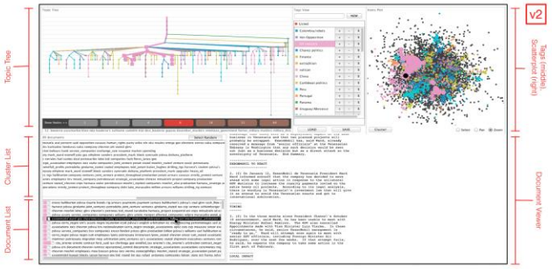
Agendas for **Visualization Research**



- **Support document investigation through visual analytics**

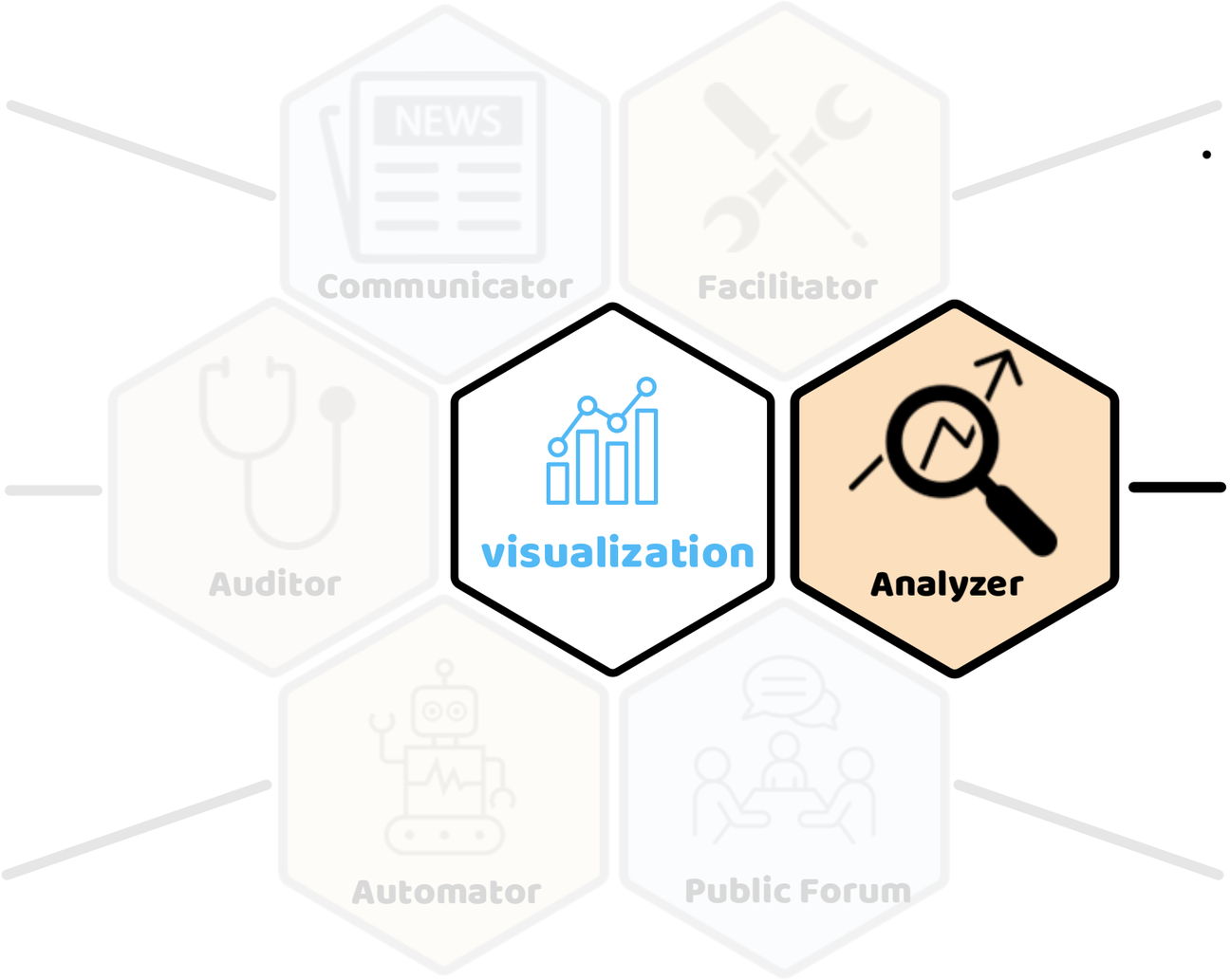


Jigsaw (Stasko et al. 2007)



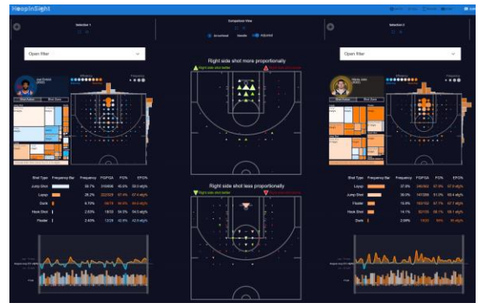
Overview (Brehmer et al. 2014)

Agendas for **Visualization Research**



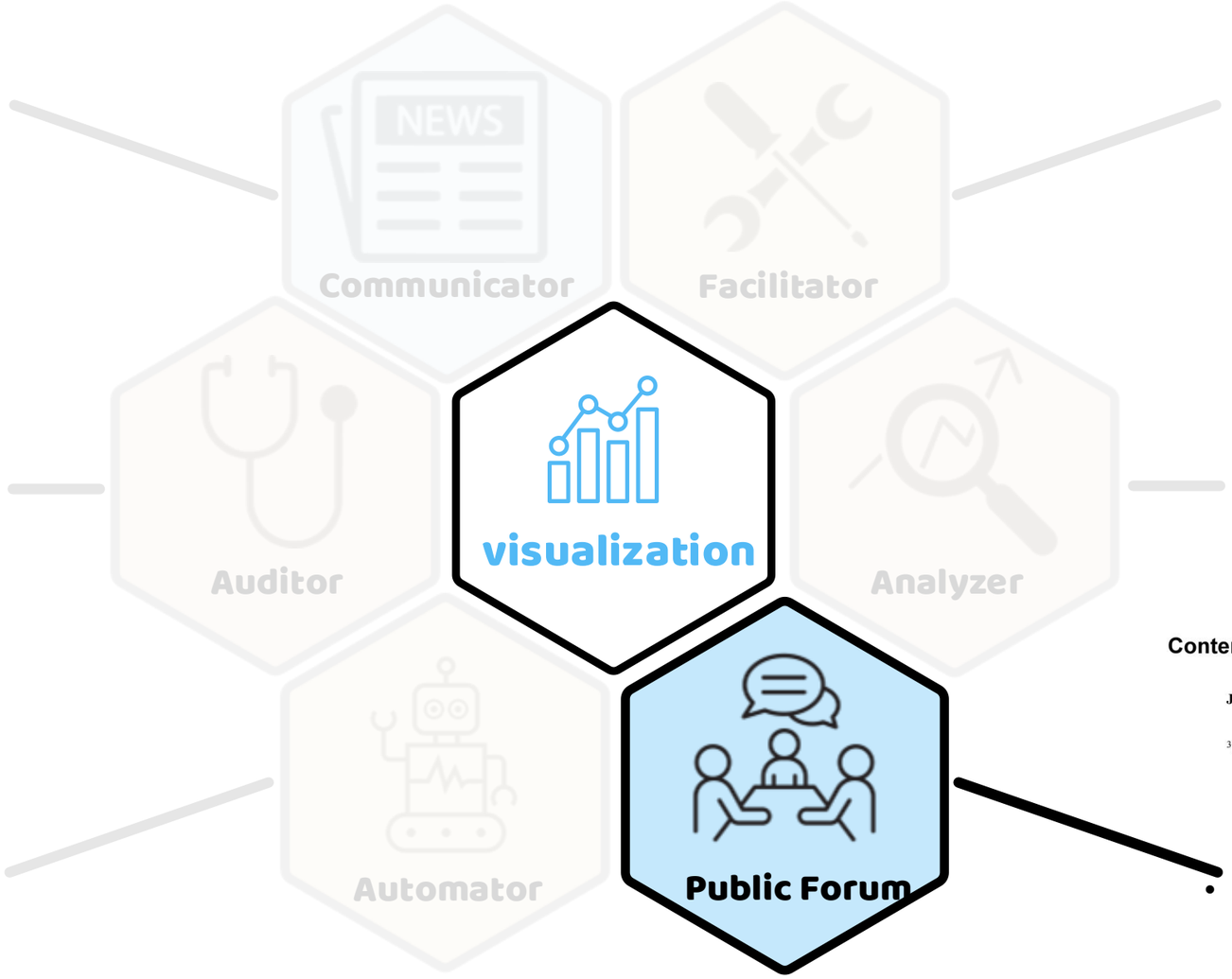
- **Support Journalists' Analytical Tasks through Visual Analytics**

GameViz (Fu and Stasko 2022)



HoopInSight (Fu and Stasko 2023)

Agendas for **Visualization Research**



Content, Context, and Critique: Commenting on a Data Visualization Blog

Jessica Hullman[1], Nicholas Diakopoulos[2], Elaheh Momeni[3], Eytan Adar[4]
[1]University of Washington, Seattle, WA, USA. jhullman@uw.edu
[2]University of Maryland, College Park, MD, USA. nad@umd.edu
[3]University of Vienna, Vienna, Austria. elaheh.momeni.roochi@univie.ac.at
[4]University of Michigan, Ann Arbor, MI USA. eadar@umich.edu

Hullman et al. (2015)

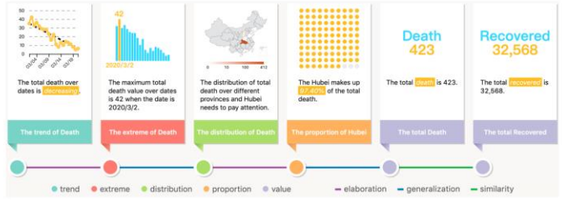
- **Understand, Enable, and facilitate data-driven commenting**

Agendas for **Visualization Research**

DataShot (Wang et al. 2019)

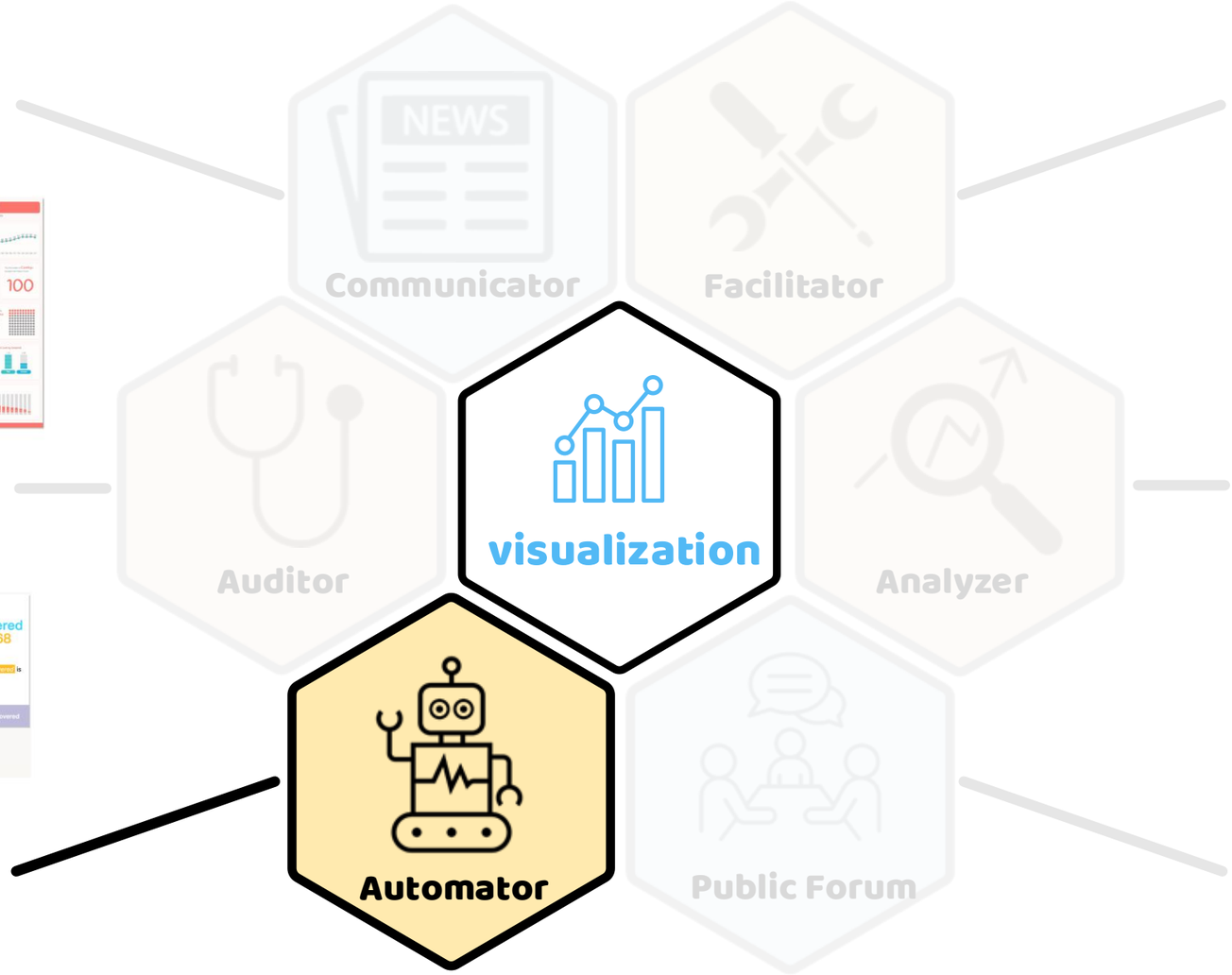
Calliope: Automatic Visual Data Story Generation from a Spreadsheet

Danqing Shi, Xinyue Xu, Fuling Sun, Yang Shi and Nan Cao



Calliope (Shi et al. 2020)

- **Automate visual data story creation**

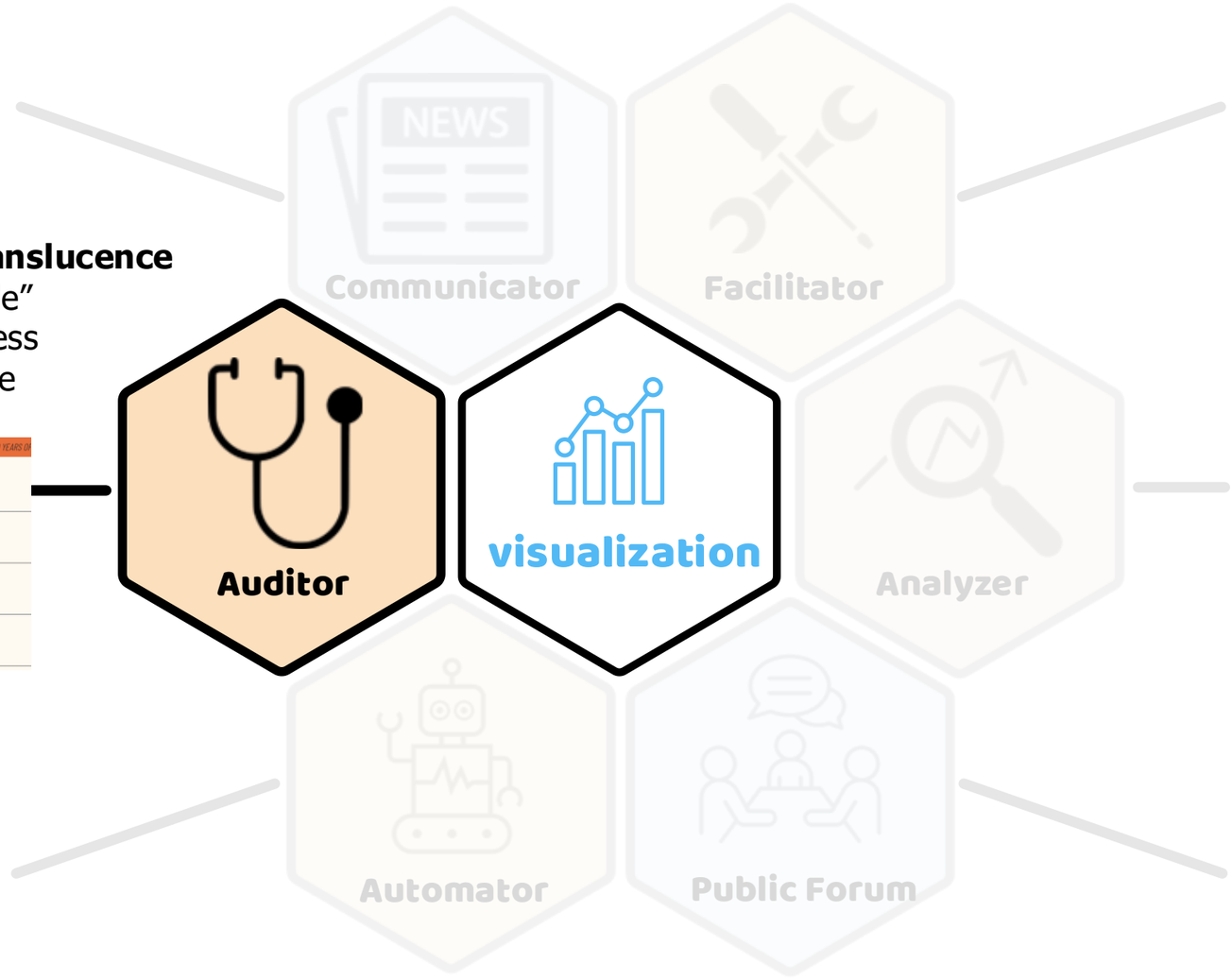


Agendas for **Visualization Research**

Visualization for Journalistic Translucence

- Make news personalization "visible"
- Raise news consumption awareness
- Hold news production accountable

The Pudding (2022)



Agendas for **Visualization Research**

Combat Data-driven Misinformation

- Detect visual deception
- Detect vis-text misalignment
- Communicate data evidence

Annotating Line Charts for Addressing Deception

Arlen Fan
Arizona State University
Tempe, Arizona, USA
arlen.fan@asu.edu

Yuxin Ma*
Department of Computer Science and Engineering, Southern University of Science and Technology
Shenzhen, China
mayx@sustech.edu.cn

Michelle Mancenido
Arizona State University
Phoenix, Arizona, USA
mickey.mancenido@asu.edu

Ross Maciejewski
Arizona State University
Tempe, Arizona, USA
rmacieje@asu.edu

Fan et al. (2022)

EMPHASISCHECKER: A Tool for Guiding Chart and Caption Emphasis

Dae Hyun Kim, Seulgi Choi, Juho Kim, Vidya Setlur, and Maneesh Agrawala

Abstract— Recent work has shown that when both the chart and caption emphasize the same aspects of the data, readers tend to remember the doubly-emphasized features as takeaways; when there is a mismatch, readers rely on the chart to form takeaways and can miss information in the caption text. Through a survey of 280 chart-caption pairs in real-world sources (e.g., news media, poll reports, government reports, academic articles, and Tableau Public), we find that captions often do not emphasize the same information in practice, which could limit how effectively readers take away the authors' intended messages. Motivated by the survey findings, we present EMPHASISCHECKER, an interactive tool that highlights visually prominent chart features as well as the features emphasized by the caption text along with any mismatches in the emphasis. The tool implements a time-series prominent feature detector based on the Ramer-Douglas-Peucker algorithm and a text reference extractor that identifies time references and data descriptions in the caption and matches them with chart data. This information enables authors to compare features emphasized by these two modalities, quickly see mismatches, and make necessary revisions. A user study confirms that our tool is both useful and easy to use when authoring charts and captions.

Index Terms—Chart and text takeaways, visual prominence, authoring, captions

EmphasisChecker (Kim et al. 2023)

"The Data Says Otherwise" – Towards Automated Fact-checking and Communication of Data Claims

Yu Fu
fuyu@gatech.edu
Georgia Tech
Atlanta, United States

Shunan Guo
sguo@adobe.com
Adobe Research
San Jose, United States

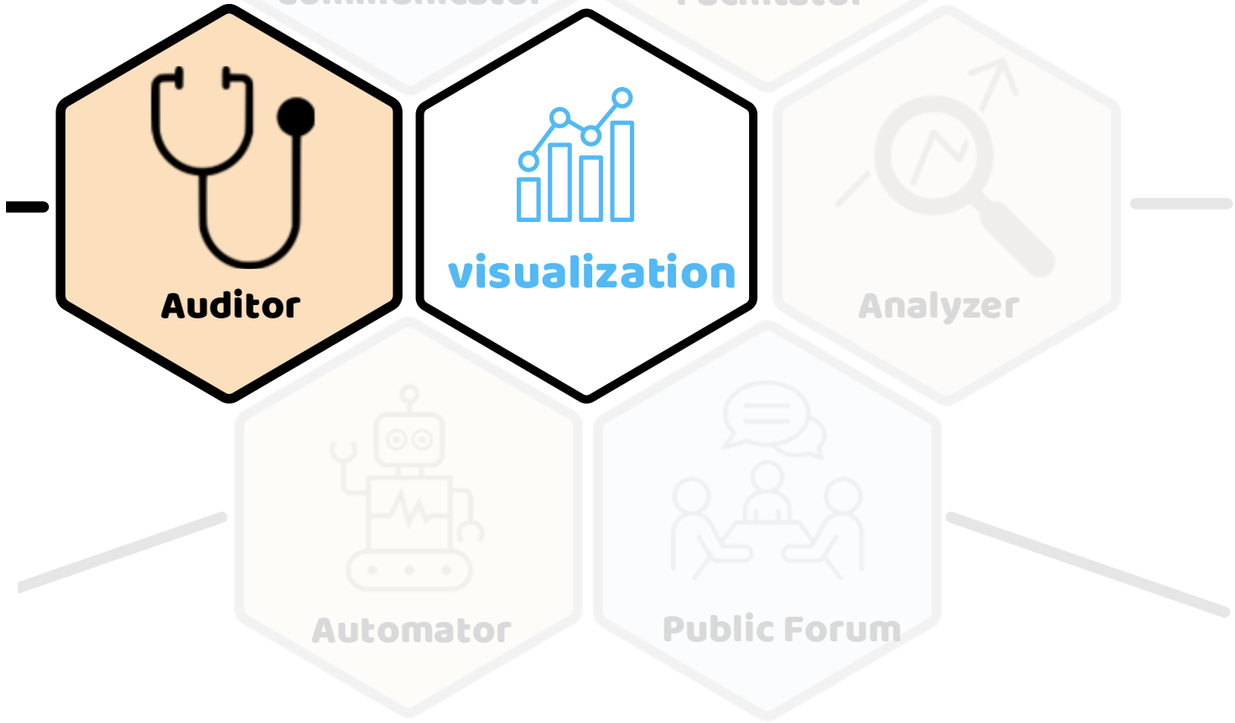
Victor S.Bursztyn
soaresbu@adobe.com
Adobe Research
San Jose, United States

Jane Hoffswell
jhoffs@adobe.com
Adobe Research
San Jose, United States

Ryan Rossi
ryrossi@adobe.com
Adobe Research
San Jose, United States

John Stasko
john.stasko@cc.gatech.edu
Georgia Tech
Atlanta, GA, United States

Aletheia (Fu et al. 2024)



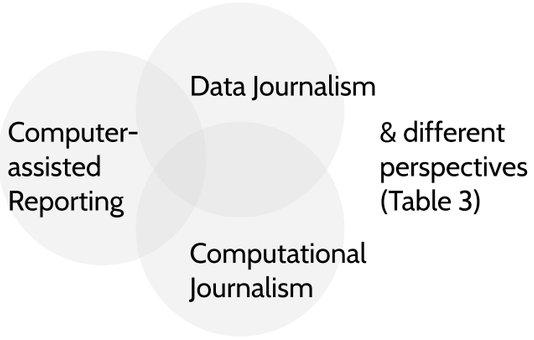
JOURNALISM

§3.1 Salient Transformations

1. Interactive Journalism
2. Participatory Environment
3. Loaded and Polluted Infosphere
4. News Personalization Algorithms

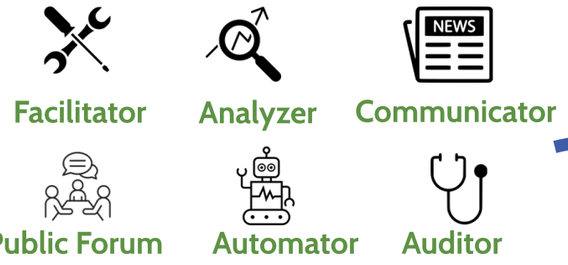
& Emerging Challenges (Table 2)

§3.2 Computational Practices



& different perspectives (Table 3)

Six Roles Of Computing In Journalism

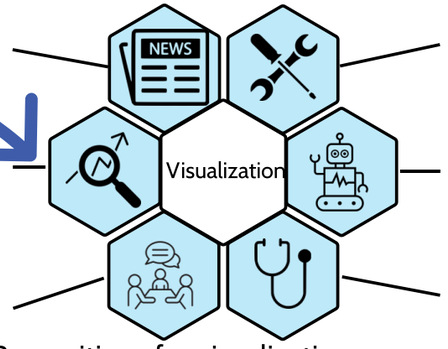


§5 Seven Research Topics And Agendas

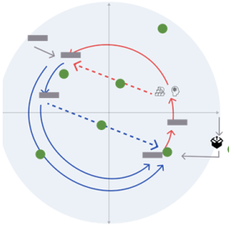
1. Facilitating Data Communication
2. Understand Visualization Dashboards on News Websites
3. Integrate Interactive Visualization with New Media
4. Support Journalists' Analytical Tasks through Visual Analytics
5. Visualization for Journalistic Translucence
6. Combating Misinformation
7. Automated Visual Stories and Insights

VISUALIZATION

§4 Contextualizing The Value Of Visualization



Propositions for visualization research to assist journalism(Figure 2)



Ecological model for mapping computing's roles in contemporary journalism (Figure 3)

Thank you!

Yu Fu
Georgia Tech

John Stasko
Georgia Tech