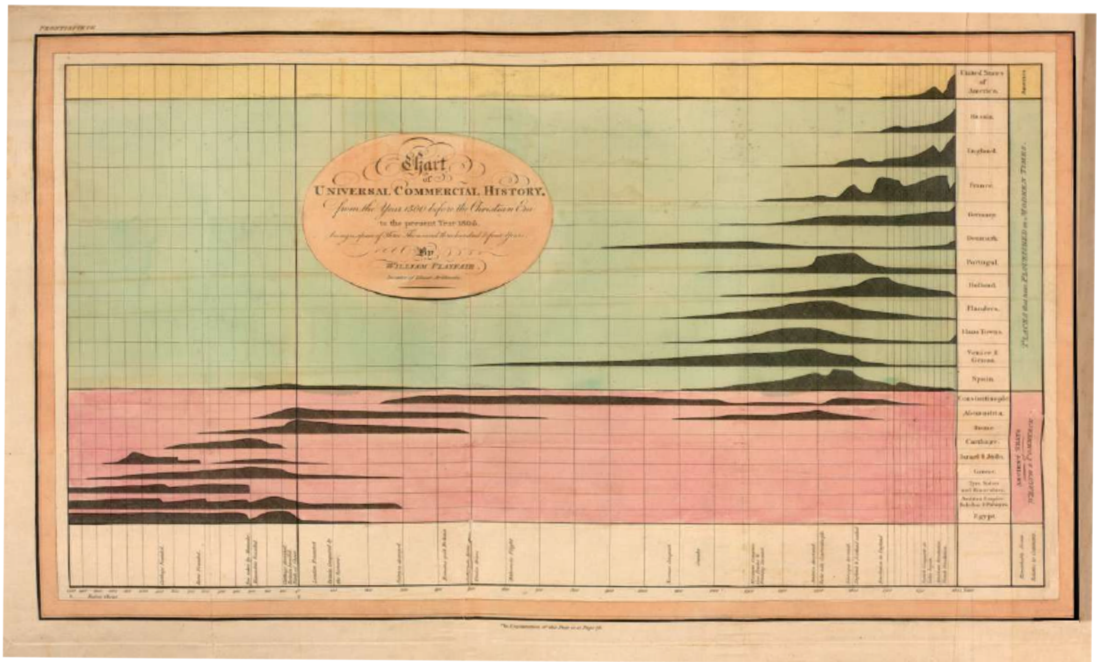
What is information visualization?



In 1805, William Playfair writes the following about the chart above:

> *The Chart, No. 1, representing the rise and fall of all nations or countries, that have been particularly distinguished for wealth or power, is the first of the sort that ever was engraved, and has, therefore, not yet met with public approbation.*[1]

About which Lev Manovich, contemporary new media Professor, commented:

> *It seems that 300 years after William Playfair's amazement at the cognitive power of information visualisation, others are finally acknowledging the projective power of his method.*[2]

[1] Playfair, W., 1805. *An Inquiry Into the Permanent Causes of the Decline and Fall of Powerful and Wealthy Nations*. London: Greenland and Norris.

[2] Manovich, L., 2011. What is visualisation? Visual Studies, 26(1), pp. 36-49.

Most definitions of information visualization by computer science researchers encompass the usage of computational and interactive visual representations and interfaces.

> *The use of computer-supported, interactive, visual representations of abstract data to amplify cognition.*[3]
>
> *Information visualization is the communication of abstract data through the use of interactive visual interfaces.*[4]
>
> *Information visualization utilizes computer graphics and interaction to assist humans in solving problems.*[5]

These definitions fail to accommodate all kinds of information visualization projects being created today, as they not always make intensive usage of interaction or do not solve specific problems. Information visualization nowadays is scattered between scientific and communicational driven approaches. For some authors, the classic definitions of information visualization refer instead to **scientific visualization** and it should be distinguished from the information visualization of today.

According to Manovich[6], scientific visualization and information visualization are distinct and came from different cultures: science and design. Scientific visualization developed in the 1980s along with the emergence of 3d computer graphics[7]. Contemporary information visualization started its development in the 1990s with the availability of 2d graphics software in personal computers[8]. The development of information visualization accelerated in the 2000s, with designers adopting high-level programming languages and APIs (e.g. Processing, D3).

---

[3] Card, S.K., Mackinlay, J. and Shneiderman, B., 1999. *Readings in information visualization: using vision to think.* San Francisco: Morgan Kaufmann Publishers Inc.

[4] Keim, D.A. et al., 2006. Challenges in Visual Data Analysis. In *Proceedings of the IEEE Conference on Information Visualization*. Baltimore: IEEE, pp. 9-16.

[5] Purchase, H.C. et al., 2008. Theoretical Foundations of Information Visualization. *In Information Visualization: Human centered issues and perspectives.* LNCS 4950. Berlin: Springer, pp. 46-64.

[6] Manovich, L., 2011. What is visualisation? Visual Studies, 26(1), pp. 36-49.

[7] Friendly, M., 2008. A Brief History of Data Visualization. In *Handbook of Data Visualization*. Berlin: Springer, pp. 15-56.

[8] Manovick, *loc. cit.*

From scientific visualization to information visualization

The term *information visualization* was probably first coined in 1989 by Robertson, Card, and Mackinlay in a paper named “The Cognitive Coprocessor Architecture for Interactive User Interfaces.” In that paper, they describe an application, called *Information Visualization*, which uses 2D and 3D animation to explore information and its structure. Their research described the automatic generation of graphical presentations of relational information (e.g. bar charts, scatter plots, node and link diagrams). They generated a wide variety of graphical presentations and evaluation criteria to identify effective presentations. After working on 2D static presentations they focused on novel 3D computer graphics and interactive animation, naming the research area **information visualization**. Some of these authors later defined information visualization as being:

> *The use of computer-supported, interactive, visual representations of abstract data do amplify cognition.*[9]

[from the Oxford Dictionary of English]

data |ˈdeɪtə|

1 facts and statistics collected together for reference or analysis;*

information |ɪnfəˈmeɪʃ(ə)n|

1 facts provided or learned about something or someone;*

**In Latin, data is the plural of datum and, historically and in specialized scientific fields, it is also treated as a plural in English, taking a plural verb, as in the data were collected and classified. In modern non-scientific use, however, it is generally not treated as a plural. Instead, it is treated as a mass noun, similar to a word like information, which takes a singular verb. Sentences such as data was collected over a number of years are now widely accepted in standard English.*

[9] Card, S., J. Mackinlay, and B. Shneiderman, Information visualization, in *Readings in information visualization: using vision to think*. 1999.

The same authors also distinguish *scientific visualization* from *information visualization*. Scientific visualization uses statistical data or data that were measured by physical instruments. On the other hand, information visualization uses abstract data, for example, financial data, business information, collections of documents, or information that does not have an obvious spatial mapping. Nonetheless the concept of cognition amplification via the extraction of relevant knowledge is ubiquitous in the definition of both.

It is a historical fact that the most prominent steps in enhancing and improving visualization experiences were made with the aid of graphics and more specifically, computer graphics. Computer graphics precipitated the evolution of a specific area of computer graphics into a discipline of its own. The data being treated has a scientific nature and is the substratum for the established field of *scientific visualization*.

Scientific visualization evolved from cartography and static diagrams to incorporate large datasets and represent the data in dynamic ways mostly with the aid of computer graphics. Its purpose is scientific insight, its data is often physically based and its techniques constitute a battery of tools that employ graphical representations that are nowadays utilized in information visualization.

Nowadays information visualization is seen as an interdisciplinary field that emerged from an evolution that passed through different fields – from the earliest and most basic representations of information, passing by Playfair’s infographics and through the techniques used in statistics and data mining that handle enormous quantities of data, and the representation tools developed within scientific visualization.

[from the Oxford Dictionary of English]

visualization
| ˌviZH(o͞o)ələˈzāSH(ə)n, ˌviZH(o͞o)əˌlīˈzāSH(ə)n |
noun

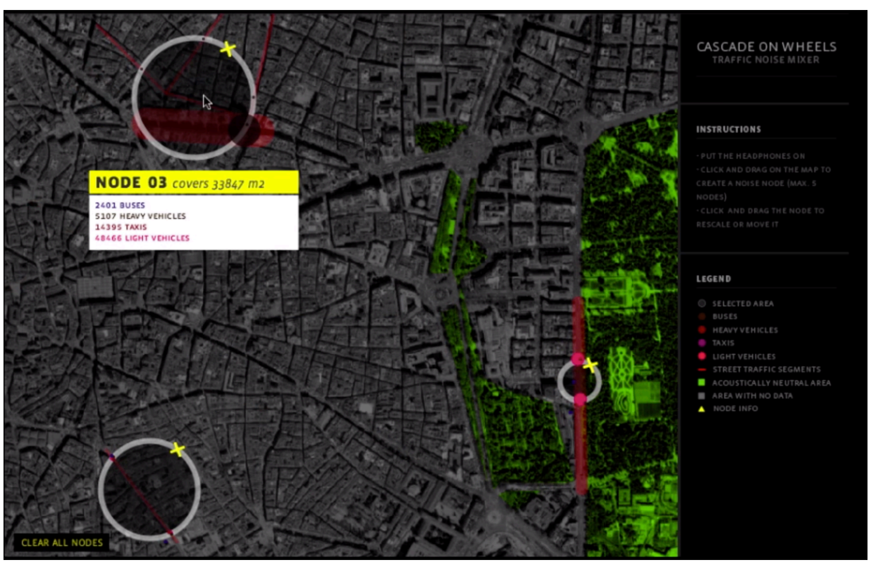
1. the representation of an object, situation, or set of information as a chart or other image: *video systems allow visualization of the entire gastrointestinal tract.*
    - a chart or other image that is created as a visual representation of an object, situation, or set of information: *3D visualizations for architectural design*.
2. the formation of a mental image of something: *the story uses descriptive language to aid visualization | visualization is a helpful technique for relieving stress |* *[count noun]* *: a powerful visualization of a future dystopia.*

But we shall start by dissecting the term *information visualization*, by first looking at the word *visualization*. **Visualization** is broadly defined as the formation a mental image of something and the act or process of interpreting in visual terms, or of putting into visible form. Considering the second definition in the Oxford Dictionary of English for visualization, one may also say that visualization is:

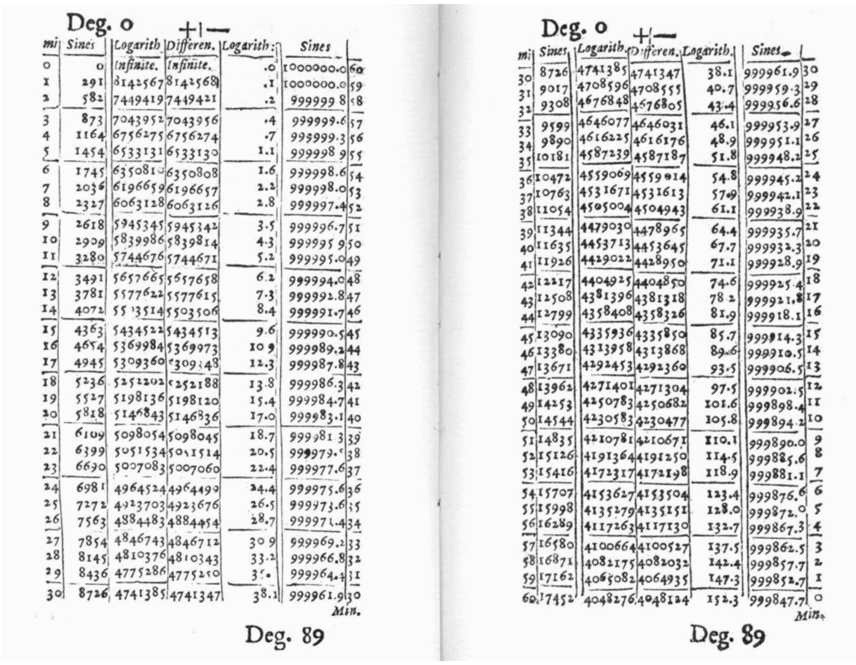
> *The formation in the mind of the image of an abstract concept.*

In spite of the usage of the term *visual*, according to all those definitions, one does not necessarily has to be involved in a visual experience in order to visualize something. This means that other sensorial experiences, such as hearing and touching can theoretically be used in order to convey information. Additionally, visualization doesn't necessarily involve the use of any tool or technology, since it solely refers to an internal cognitive activity in which humans engage.

One example of data sonificiation, and how to use sound to convey information is the Traffic Mixer of the Cascade on Wheels Project.



https://vimeo.com/14853139



Deg. 0 +|—

mi	Sines	Logarith	Differen.	Logarith:	Sines	
0	0	Infinite.	Infinite.	.0	1000000.0	60
1	291	8142567	8142568	.1	1000000.0	59
2	582	7449419	7449421	.2	999999.8	58
3	873	7043952	7043956	.4	999999.6	57
4	1164	6756275	6756274	.7	999999.3	56
5	1454	6533131	6533130	1.1	999998.9	55
6	1745	6350810	6350808	1.6	999998.6	54
7	2036	6196659	6196657	2.2	999998.0	53
8	2327	6063128	6063126	2.8	999997.4	52
9	2618	5945345	5945342	3.5	999996.7	51
10	2909	5839986	5839814	4.3	999995.9	50
11	3280	5744676	5744671	5.2	999995.0	49
12	3491	5657665	5657658	6.2	999994.0	48
13	3781	5577622	5577615	7.3	999992.8	47
14	4072	55[illegible]3514	5503506	8.4	999991.7	46
15	4363	5434522	5434513	9.6	999990.5	45
16	4654	5369984	5369973	10.9	999989.2	44
17	4945	5309360	[illegible]309[illegible]48	12.3	999987.8	43
18	5236	5252202	[illegible]252188	13.8	999986.3	42
19	5527	5198136	5198120	15.4	999984.7	41
20	5818	5146843	5146836	17.0	999983.1	40
21	6109	5098054	5098045	18.7	999981.3	39
22	6399	5051534	50[illegible]1514	20.5	999979.[illegible]	38
23	6690	5007083	5007060	22.4	999977.6	37
24	6981	4964524	4964499	24.4	999975.6	36
25	7272	4923703	4923676	26.5	999973.6	35
26	7563	4884483	4884454	28.7	99997[illegible].4	34
27	7854	4846743	4846712	30.9	999969.2	33
28	8145	4810376	4810343	33.2	999966.8	32
29	8436	4775286	4775250	3[illegible]	999964.4	31
30	8726	4741385	4741347	38.1	999961.9	30
					Min.	

Deg. 89

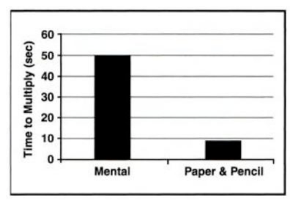
Deg. 0 +|—

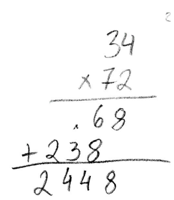
mi	Sines	Logarith.	Differen.	Logarith.	Sines	
30	8726	4741385	4741347	38.1	999961.9	30
31	9017	4708596	4708555	40.7	999959.3	29
32	9308	4676848	4676805	43.4	999956.6	28
33	9599	4646077	4646031	46.1	999953.9	27
34	9890	4616225	4616176	48.9	999951.1	26
35	10181	4587239	4587187	51.8	999948.2	25
36	10472	4559069	4559014	54.8	999945.2	24
37	10763	4531671	4531613	57.9	999942.1	23
38	11054	4505004	4504943	61.1	999938.9	22
39	11344	4479030	4478965	64.4	999935.7	21
40	11635	4453713	4453645	67.7	999932.3	20
41	11926	4429022	4428950	71.1	999928.9	19
42	12217	4404925	4404850	74.6	999925.4	18
43	12508	4381396	4381318	78.2	999921.8	17
44	12799	4358408	4358326	81.9	999918.1	16
45	13090	4335936	4335850	85.7	999914.3	15
46	13380	4313958	4313868	89.6	999910.5	14
47	13671	4292453	4292360	93.5	999906.5	13
48	13962	4271401	4271304	97.5	999902.5	12
49	14253	4250783	4250682	101.6	999898.4	11
50	14544	4230583	4230477	105.8	999894.2	10
51	14835	4210781	4210671	110.1	999890.0	9
52	15126	4191364	4191250	114.5	999885.6	8
53	15416	4172317	4172198	118.9	999881.1	7
54	15707	4153627	4153504	123.4	999876.6	6
55	15998	4135279	4135151	128.0	999872.0	5
56	16289	4117263	4117130	132.7	999867.3	4
57	16580	4100664	4100527	137.5	999862.5	3
58	16871	4082175	4082032	142.4	999857.7	2
59	17162	4065082	4064935	147.3	999852.7	1
60	17452	4048276	4048124	152.3	999847.7	0
					Min.	

Deg. 89

John Mapier, 1614 —
Table of logarithms

Card[10], through and an informal experiment, suggested that the time to multiply two numbers mentally is five times superior than the time to do it with the aid of paper and pencil. Such a difference exists because the multiplication is being externalized to a visual representation, holding partial results outside the mind and extending the person's working memory. This means that an internal memory task was converted to an external task that consists of manual writing and visual search. In this simple example there are two aspects of special relevancy. First, the use of written partial and total results, are a visual aid to the understanding of the classic multiplication process. Second, the manual writing represents a tool for aiding cognition. Therefore one can state that the aid of the paper and pencil results in a more effective visualization of the final result of the multiplication, supported by an adequate tool and a visual representation.





[10] Card, S., J. Mackinlay, and B. Shneiderman, Information visualization, in *Readings in information visualization: using vision to think.* 1999.

According to Norman, the human intelligence is superb at inventing procedures and objects that overcome its own limits. Historically, humans have increased their memory, thought and reasoning through the invention of external aids, capable of enhance cognitive abilities.

> *It is things that make us smart.*[11]

One of the most primitive and peremptory external aid is graphics. And graphics serve two purposes:

1. communicate an idea;
2. create or discover the idea itself (***or new ideas***).

Graphics are an extremely important way to communicating quantitative information. Graphs are used extensively to convey information because they do so effectively. Quantitative patterns and relationships in data are easily revealed by graphs because of the enormous power the **cognitive-visual system**[12] to perceive geometrical patterns. This system can summarize vast amounts of quantitative information on a graph, perceiving distinctive features, or focusing on specific details.

A critical view on a classical definition

The classic definition on information visualization by Card should be revisited from a critical perspective.

> *The use of computer-supported, interactive, visual representations of data to amplify cognition.*

Being visualization a cognitive activity, it can be facilitated by graphical external representations from which an internal mental representation is constructed. Consequently the definition of visualization is also stated in the context of the type of external representations and the specific tools used. In this case the specific tools are computers and the external representations are graphics with perhaps some kind of physical interaction. While it can be reasonable to associate visualization with computers in consideration to its modern form, imposing visual representations to data portrayal is rather restrictive. As it was mentioned before, they too can

[11] Norman, D.A., *Things That Make Us Smart: Defending Human Attributes In The Age Of The Machine.* 1993: Perseus Books.

[12] William S. Cleveland & Robert McGill (1984) Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods, Journal of the American Statistical Association, 79:387, 531-554, DOI: 10.1080/01621459.1984.10478080

make use of other sensory experiences. The same goes for the interactive restriction in the definition, as it is not necessary in order to have an effective visualization. Truth be told, that the cognitive component has always been the backbone of all of Card's definitions, as he says:

> *The purpose of visualization is insight, not pictures.*[13]

Scientific visualization within information visualization

Card always contrasted his definition of information visualization to his own definition of scientific visualization, as it was:

> *Use of interactive visual representations of scientific data, typically physically based, to amplify cognition.*

In 2008, Card redefined information visualization in a less restrictive way:

> *Information visualization is a set of technologies that use visual computing to amplify human cognition with abstract information.*[14]

Naturally, the approach to information visualization should be less restrictive and broader, as it doesn't have to necessarily involve computing or visual representations. A less restrictive definition comes from Chen in 2005:

> *Visual representations of the semantics, or meaning, of information. In contrast to scientific visualization, information visualization typically deals with nonnumeric, nonspatial, and high-dimensional data.*[15]

---

[13] *Ibid*.

[14] Card, S., The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications. In *Information visualization*, ed. A. Sears and J.A. Jacko. 2008: Law- rence Erlbaum Assoc Inc.

[15] Chen, C., Top 10 unsolved information visualization problems. *Computer Graphics and Applications*, IEEE, 2005. 25(4): p. 12-16.

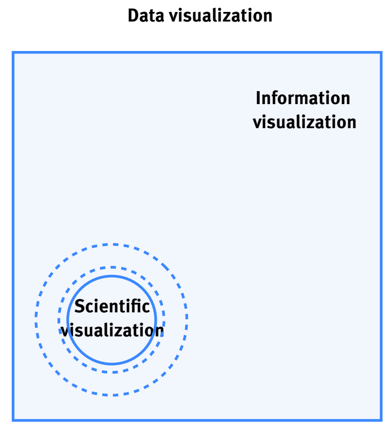
Data visualization is usually categorized between information visualization and scientific visualization. Such boundaries are fuzzy. Scientific visualization has a more restricted audience—the scientific community. But nowadays information visualization can be understand as also embracing scientific visualization if the following points are considered:

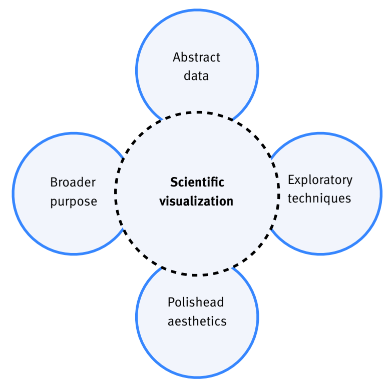
1. Deals not only with physical related data, but also with abstract data that does not have a straightforward mapping to a graphic space. The datasets are usually high dimensional and comprise large quantities of data.

2. Information visualization has a broader purpose than scientific visualization. It deals too with scientific analysis but not exclusively. The artifacts produced in the context of information visualization often deal with questions outside academic research. Everyday data is processed in order to produce high level insights that synthesize relevant aspects. It is also important to note that this relevancy is targeted to a much broader audience than within scientific visualization. It is possible to get meaningful insights about questions that concern a large audience without the deep knowledge in of these specific areas (in contrast with scientific visualization).

3. Information visualization characterizes itself for often making use of less orthodox representation techniques, investing in other ways of perception besides the visual (e.g. data sonification, data physicalization). It also enhances a high degree of interactivity (e.g. invisible interfaces, tangible interfaces, multitouch) and large installations such as immersive environments. Computer graphics are pushed beyond their possibilities with animation techniques. Storytelling is often used in information visualization.

4. With the broader audience comes bigger concerns with the aesthetics of the artifacts. Information visualization gathers enthusiasts from artistic or design related fields, where there is an interest in transmitting a message in an effective way. Beyond that, information visualization cares about usability issues from the human-computer interaction discipline and incorporates several aspects of traditional graphic design, when it comes to the building of visual artifacts or to the whole design process.





With this point of view it is possible to outline and envision information visualization through various perspectives capable of embracing a greater set of information visualization artifacts. Therefore, information visualization is understand as a subject that takes care on synthesizing large quantities of data with the purpose to enhance cognition and typically conveying it with a great regard for aesthetics. The importance of this view is that there are no restrictions imposed by the medium, techniques or tools used. This conceptual and general view enables to decipher information visualization trough other, more concrete, and contemporary perspectives.

A more contemporary view

Information visualization is an **interdisciplinary** field that intersects *graphic design*, *human-computer interaction*, *computer graphics* and *statistics*.[16] The graphic portrayal of quantitative information has deep roots in the early history of the earliest map-making and visual depiction, but also in the modern era of thematic cartography, statistics and statistical graphics. Other advancements contributed to the widespread use of data visualization such as technologies for drawing and reproducing images, advances in mathematics and statistics, and new developments in data collection. Since its inception, information visualization was tightly connected with the emergence of new computer graphic techniques in the 80s. The *democratization of large amounts of data*, the *increase of software availability* and the evolution of more *intuitive paradigms of visual programming* has broadened the field with the inclusion of practitioners from other non-scientific areas and new audiences. **The inclusion of practitioners with backgrounds in graphic design, communication and storytelling defines contemporary information visualization.**

The interdisciplinary nature of information visualization attracted researchers, intellectuals, curious and enthusiasts from a broad range of disciplines and culminated a boom of works in the field in the beginning of the 21st century. Many of them came from a nonscientific background and frequently with high sensibility towards aesthetics (e.g. graphic design, art). Sometimes aesthetics' concerns were over glorified to the point of constituting the sole purpose of the visualization. Those types of works should constitute themselves as **information aesthetics**, where the beauty

[16] Friendly, Michael. A Brief History of Data Visualization. In *Springer Handbooks Comp.Statistics*, 15–56. Springer Berlin Heidelberg, 2008.

of information is glorified in detriment of the extraction of relevant knowledge.

In 2009, Manuel Lima wrote an **Information Visualization Manifesto**[17] that aims to abruptly distinguish information visualization from what he calls *information art* or *information aesthetics*. The emerged problem consists of a series of data graphics that claim themselves as information visualization artifacts, following similar aesthetic approaches but without providing meaningful insights (i.e. they existing for the sake of the colloquial eye-candy). Therefore, Lima enumerates a list of ten directions, methodologies or techniques, that should be adopted in information visualization projects.

1. Form follows function – form follows revelation.
2. Start with a question.
3. Interactivity is the key.
4. Cite your source.
5. The power of narrative – using narratives.
6. Do not glorify aesthetics.
7. Look for relevancy.
8. Embrace time – envisioning data through time.
9. Aspire for knowledge.
10. Avoid gratuitous visualizations.

These directives provide a framework that targets projects to the production of information visualization artifacts, but they also establish solid boundaries in the field. Beyond that, it sets a tone towards the glorification of aesthetics. That tone seems to impose a prohibition towards overly expressive artifacts.

Lima's tone constitutes an analogy with Tufte's point of view. Edward Tufte is a scholar well known for introducing a series of principles to graphical excellence and integrity. Amongst other rules of thumb, he writes ***above all else show the data***[18], defending that within reason, non-data ink or redundant ink should be erased as a way to maximize the data-ink ratio. Tufte introduced the term "chartjunk" in his first book, referring to unnecessary graphical elements that add no value to data graphics:

[17] http://www.visualcomplexity.com/vc/blog/?p=644

[18] Tufte, E R. *The visual display of quantitative information*. Vol. 7. Graphics press Cheshire, CT. Graphics press Cheshire, CT, 1983.

> *The interior decoration of graphics generates a lot of ink that does not tell the viewer anything new... Regardless of its cause, it is all non-data-ink or redundant data-ink, and it is often chartjunk.*[19]

He adds that often, chartjunk does not involve artistic considerations, and it is simply conventional paraphernalia that is added routinely to data graphics. He states that good data graphics are dependent on good datasets, and that overly simple data does not benefit from the addition of decorative elements.

> *Chartjunk can turn bores into disasters, but it can never rescue a thin data set. The best designs... are intriguing and curiosity-provoking, drawing the viewer into the wonder of the data, sometimes by narrative power, sometimes by immense detail, and sometimes by elegant presentation of simple but interesting data.*[20]

His remarks emerged in a context when many statistic and information graphics, aimed either at the scientific community or the general public, frequently incorporated decorative elements without any semantic value, such as adding an unnecessary third dimension to a chart, recurring to heavy patterns that can cause moiré vibrations and making use of unnecessarily heavy ink.[21]

Ellen Lupton, a well known design educator addressed Tufte's vision, presenting a counter-perspective.

> *Tufte (...) argues that a chart or diagram should employ no metaphoric distractions or excessive flourishes (...) Tufte's purist point of view is profound and compelling, but it may be overly restrictive. Information graphics have a role to play in the realm of expressive and editorial graphics... They can be clean and reductive or richly expressive, creating evocative pictures that reveal surprising relationships*

---

[19] *Ibid*.

[20] Tufte, E R. *The visual display of quantitative information*. Vol. 7. Graphics press Cheshire, CT. Graphics press Cheshire, CT, 1983.

[21] Examples of these will be discussed further on in the course.

> *and impress the eye with the sublime density and grandeur of a body of data.* [22]

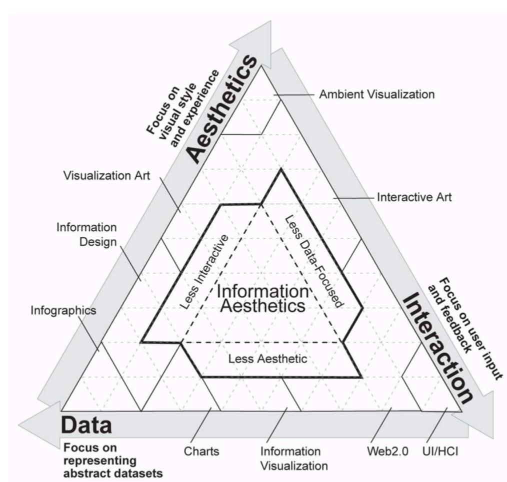
As much as the extraction of relevant knowledge is the purpose of information visualization, the aesthetic concern with the final artifacts should not be perceived as a restriction. Therefore, information design and visualization see itself as associated with a graphic design, following a specific function and but doing so with great attention to graphic detail. This results that formally the practicing of graphic design and expressivity, together with affirmative aesthetics' concerns can be embraced in the process of information visualization.

Nowadays the term *information visualization* is often loosely applied to all sorts of artifacts that spans through related fields that in rigor do not comprise the set of characteristics already defined. Terms in vogue are:
— infographics
— data visualization
— data graphics
— information art
— information design
— data visualization
— information aesthetics
— visualization art

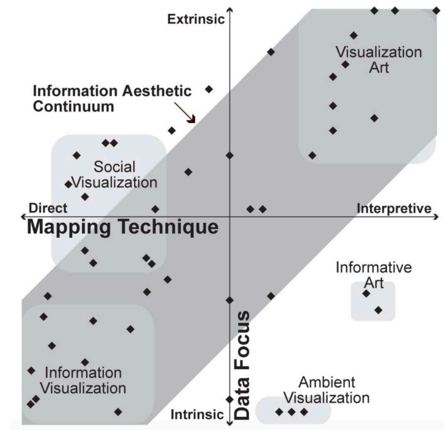
The most peremptory overlapping between concepts and definitions is between information visualization and information aesthetics. Information aesthetics is defined by Andrea Lau and Andrew Moere as focusing on three main issues:

> *Representing data, providing an interactive interface, and using visual appeal to engage the user.*[23]

Lau proposed a model is based on analyzing existing visualization techniques by their interpretative intent and data mapping inspiration. It reveals information aesthetics as the conceptual link between information visualization and visualization art, and includes the fields of social and



[22] Lupton, E. and Phillips, J.C., 2008. *Graphic Design The New Basics*. New York: Princeton Architectural Press.
[23] Lau, A. and A.V. Moere, Towards a Model of Information Aesthetics in Information Visualization, in *11th International Conference Information Visualization*, IEEE, Editor. 2007, IEEE Computer Society.



ambient visualization. As several information visualization artifacts fall within several characteristics described in Tau's model: data, aesthetics and interaction. Accordingly to the model, information aesthetics uses more interactive approaches than visualization art and places more emphasis on visual style than information visualization. Therefore the model defines that one should not impose strict boundaries on information aesthetics, but should envision it as a variable on a contiguous scale, that in spite of being autonomous, is a result of different focusing degrees across the three fields. One should also notice that the model defines the focusing of different fields, but it never states for example that a high focusing degree on data means exclusively that aesthetics are not present. It just states that little concern was given to aesthetics.

Design + visualization

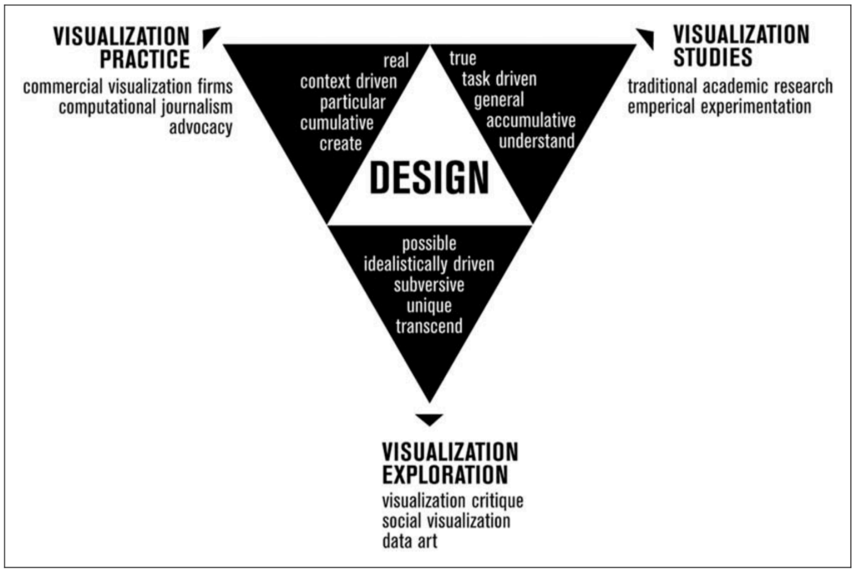
For Vande Moere and Purchase[24], visualization has traditionally focused on supporting expert users in exploring and analyzing complex data as efficiently and effectively as possible, portraying data in a scientific and neutral way. In recent years, information visualization has moved from an advanced research topic to mainstream adoption, encompassing lay users as opposed to a traditional audience of scientists and analysts. For example, Heer et al.[25] classify a few of many information visualization projects across this spectrum. There are new users utilizing visualization tools, from expert developers to savvy designers and novice consumers. Additionally, there are new goals in information visualization detached from traditional interactive exploratory analysis. These new goals are communication-driven and do not demand as much effort from its users.

Considering a more general characterization of information visualization, superseding the duality between science and communication, Vande Moere and Purchase place[26] visualization on three major axes: **visualization studies**, **visualization practice**, and **visualization exploration**. Visualization studies have their ground in academic research and computer science, using systemic inquiry and searching for empirically proven insights. Visual-

[24] Vande Moere, A.V. and Purchase, H., 2011. On the role of design in information visualization. *Information Visualization*, 10(4), pp. 356-371.

[25] Heer, J. et al., 2008. Creation and Collaboration: Engaging New Audiences for Information Visualization. *Information Visualization*, 4950, pp. 92-133.

[26] Moere, *loc. cit.*



ization practice uses the body of knowledge produced in visualization studies to produce visualizations. Visualization practice is seen by Vande Moere and Purchase as a design practice in the sense that it intends to create workable solutions rather than accumulate knowledge. Visualization as a practice is found in data analysis firms, design studios and individual freelancers. Visualization exploration has more idealistic and visionary goals, often advancing not the techniques themselves, but how these techniques are contextually used and intertwined with their communicational purposes. From this point of view, visualization exploration also covers more artistic approaches to visualization. For example, Viegas and Wattenberg[27] have used the term **artistic data visualization** to refer to several data visualizations used for mass communication that employ vigorous perspectives of their authors. This contrasts with scientific visualization where the importance of an authorial role is disregarded, and visualizations are for the most part seen as neutral. In scientific visualization, *developers* build visualizations, and in exploratory visualizations *authors* build their visualizations.

> *The value of the artworks rests on the fact that their creators recognize the power of visualization to express a point of view. By contrast, traditional analytic visualization tools have sought to minimize distortions, since these may interfere with dispassionate analysis. Is it possible that*

[27] Viegas, F. and Wattenberg, M., 2007. Artistic data visualization: Beyond visual analytics. *Online Communities and Social Computing*, LCNS 4564, pp. 182-191.

> *this focus on minimizing "point of view" is misguided? For one thing, it is generally impossible to create a visualization that is truly neutral, just as it is impossible to create a flat map of the Earth's surface without distorting distances. (...) Perhaps instead of seeking simply to minimize the intrusion of point- of-view, a more realistic attitude for a designer of a visualization should be, as with traditional maps, to choose which perspective is the right one for a given analytic task.*[28]

This exploratory side of visualization is on par with **casual information visualization**, defined as an umbrella term by Pousman et al.[29] in order to extend the traditional view on information visualization. This complementary view, unifies ambient, social, and artistic visualization domains. According to the authors, there are four main differences that characterize casual information visualization, that range from the characterization of the target audience, to the objectives of the authors and the contexts of the visualization:

— The target audience is broad and includes a wide spectrum of users, from experts to novices. The target audience might not be trained in analytical thinking or interpreting visualizations.
— The usage patterns of the users can include momentary, repeatable, or contemplative usages. A contemplative usage pattern would be, for example, a long observation of an intricate visualization in an art gallery.
— The users are more tightly coupled with the data, since they can connect with the messages that the visualization delivers. The data and the message, instead of being solely work-related, have personal relevance, such as providing a provocative view on society, and fostering reflexion by each individual in society, being a way of imagining one's identity.
— The authors of casual information visualization are not strictly interested in providing analytical insights, but instead focus on providing awareness, social, and reflective insights. Such may range from raising consciousness to a particular situation, to providing a different perspective on one's role in society.

---

[28] Viegas, F. and Wattenberg, M., 2007. Artistic data visualization: Beyond visual analytics. *Online Communities and Social Computing*, LCNS 4564, pp. 182-191.

[29] Pousman, Z., Stasko, J.T. and Mateas, M., 2007. Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6), pp. 1145-1152.

Casual information visualization has an increased focus on activities that are less task-driven, datasets that are personally meaningful, and visualizations that are built for wide audiences. According to the authors, this view points to new questions in visualization, such as:

> *How can we design systems whose highest aims are not focused on productivity, but instead on notions of usefulness, enjoyment, and reflection?* [30]

Vande Moere and Purchase's triadic view on visualization is not discrete, but is instead continuously defined on a triangle with vertices on practice, studies, and exploration. In the middle ground of these approaches to visualization is **design**. All of these three approaches contribute with their own unique perspectives on information visualization. The view of Vande Moere and Purchase is that most innovation can occur in the design domain, when the three domains overlap. They then characterize the role of design in information visualization:

— Design in visualization is not an *add-on* reasoning that can be added after the decisions involved in making the visualization. That is to say that designing a visualization does not only refer to the *visual mapping stage*, or the *representation* stage in the information visualization process; instead, it should be present in every step, from data acquisition, data processing, up to refinement and interaction. Furthermore, the design process does not have to be streamlined – it may be iterative or non-linear.

— Design in visualization is **not a "black box"** and can be disseminated. Design is a well-reasoned activity that involves making conscious decisions that are consistent with each other, which have a conceptual ground and can be objectively reproduced. The rationale should be explicit and not hidden by appropriately justifying visual encoding choices, reflecting on aesthetic considerations, and logically defining the metaphors being used. These rationales contribute to a body of generalizable knowledge, and to other cases beyond the one described. As noted by Vande Moere and Purchase, researchers in design-oriented approaches to visualization, have the scientific responsibility to be as articulate as possible about their rationale in their design processes.

30 Pousman, Z., Stasko, J.T. and Mateas, M., 2007. Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6), pp. 1145-1152.

— Design in visualization **is not art** since it still has a functional purpose that is the clear communication of a data-driven message. As Fallman notes[31], design should not be thought as positioned on a continuum between science and art. He emphasizes that design builds on its own body of knowledge, being a “tradition guiding action and thought, which spans across many disciplines.” [32]

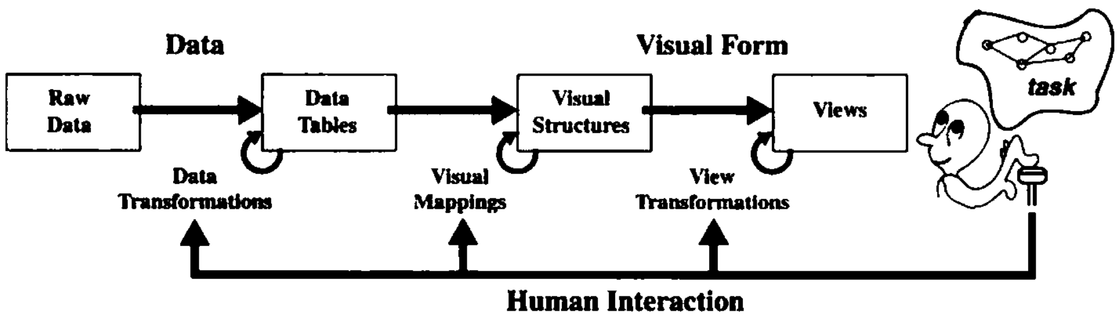
Several visualization applications enjoy wide success and visibility, perhaps due to their design, from the message and the data to the visual representations. There are differences in approach from the traditional visualization community and visualization designers. The later apply design principles that equate both form and function to achieve subjective qualities such as clarity, appeal, and elegant aesthetics. The former applies visual solutions that are empirically proven to be task-effective, hence placing an emphasis on usability, regardless if they are intrinsically unappealing. There are clear advantages in crossing those approaches and merging the perspectives and experiences of both communities. Interestingly, merging these different approaches is still unorthodox in some research communities, agreeing with Vande Moere and Purchase when they say that there is a rift between visualization researchers and designers.

[31] Fallman, D., 2003. Design-oriented human-computer interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Ft. Lauderdale, FL: ACM, pp. 225-232.

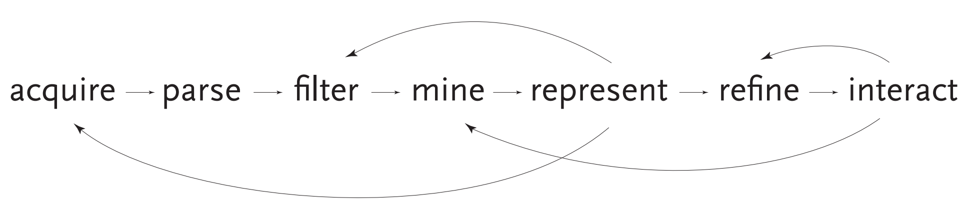
[32] *Ibid.*

A process of many processes

Attempts to describe the process behind practicing information visualization is as early as Card's definition of information visualization[33].



A more recent process is the one envisioned by Ben Fry[34]. Providing meaningful insights about complex data requires interdisciplinary skills from statistics, data mining, graphic design and human-computer interaction. However, these fields historically evolved separated and their specialists often encountered problems in communicating and integrating resolution strategies across fields. This means that, for example, graphic designers know how to choose a proper visual representation but they do not know how to manage large quantities of data. In the same way, data mining techniques deal with complex data but do not address graphical representation problems or interaction issues. Therefore there's the urge to conciliate these fields as parts of a single process. This process should be envisioned as a working and communication methodology in multi disciplinary teams, or define in itself the general competencies of an 'information visualizer'. The original process as laid out by Ben Fry is shown below. This process is going to be described with minor adaptions, as three steps are grouped under a major preprocess step.



---

[33] Card, S.K., Mackinlay, J. and Shneiderman, B., 1999. *Readings in information visualization: using vision to think*. San Francisco: Morgan Kaufmann Publishers Inc.

[34] Fry, B.J., 2004. *Computational information design*. PhD dissertation, Massachusetts Institute of Technology.

1. **Acquire** • The first step concerns about how the data is first retrieved and where does it come form. Several data inputs can be used like analog signals, files on a disk, streams from a network or relational databases. Each source comprises its own set of techniques.
2. **Preprocess** • Treatments given to data.
    1. **Parse** • This step deals with the first treatment given to raw data. Its purpose is to convert data in structures that can be interpreted and used. It usually passes through some pre-filtering such as de-noising procedures, decompress or decrypting data. In the end, data is stored in an abstract structured form, like lists, matrices or graphs.
    2. **Filter** • The filtering step merely consists in selecting a subset of the data with major interest. This also means that non relevant data is discarded.
    3. **Mine** • Mining is the final and most important treatment given to data. It uses techniques from mathematics, statistics, artificial intelligence or advanced data mining. It might consist of simple measures like max, min, median, mean, normalization, variance and standard deviation, but it often involves counting, sorting and searching algorithms, as well as clustering procedures (e.g. self-organizing maps). The purpose of this step is to find mathematical patterns in the data.
3. **Represent** • This step concerns with choosing the first (usually visual) forms of representation. Visual representations can be chosen from a wide variety of possibilities.
4. **Refine** • This stage aims to improve the basic representation already chosen. They concern with the most elementary graphic design principles that enhance visual perception and aesthetic value. Basic principles like contrast, hierarchy, scale, color and rhythm are introduced and refined. In this step it should be paid attention to unnecessary graphic elements, such as Tufte's 'chart junk.'
5. **Interact** • The final step deals with interaction techniques between the user and the data representations or the interactions between the representations of the data themselves. This last kind of interaction is named data self-interaction and comprises computational design algorithms that aim to enhance the visual display of data and avoid obscuring information in visually complex artifacts. Those computational design algorithms can constitute straightforward layout rules (e.g. even distribution, connect elements, grid alignment) or they can invest in physics-based placement, where Newtonian laws are applied to implement attraction or repulsion between elements. In consideration to user

interaction, it controls what data features are visible and deals with classic usability issues (viewing controls, restraint of choices).

It is also interesting to group all of the steps of the information visualization process by field. For instance “acquire and parse" positions itself in *computer scienc*e, “filter and mine” are used in a *mathematical, statistical* and in *data mining* or *machine learning* contexts. "Represent and refine” are closer to *graphic design* and the interact step is about *human-computer interaction* and specific graphic design rules used in the broad context of information visualization. This process constitutes a guideline and rarely is strictly followed. In many projects only some of the steps are used. Beyond that, each step in the process does not have to be static, and does not always move forward. This means that the visualization tools that result from the final interact step may push to modify aspects in previous steps.