ISSN 2186-7437

NII Shonan Meeting Report

No. 212

Colors and Visualization Strategies in Structural and Cellular Biology

Monica Zoppè
Yoshie Kiritani
Daisuke Inoue

May 12–16, 2025

National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-Ku, Tokyo, Japan

Colors and Visualization Strategies in Structural and Cellular Biology

Organizers:
Monica Zoppè (Institute of BioPhysics, Milan, CNR of Italy)
Daisuke Inoue (Kyushu University, Japan)
Yoshie Kiritani (Chiba University, Japan)

May 12–16, 2025

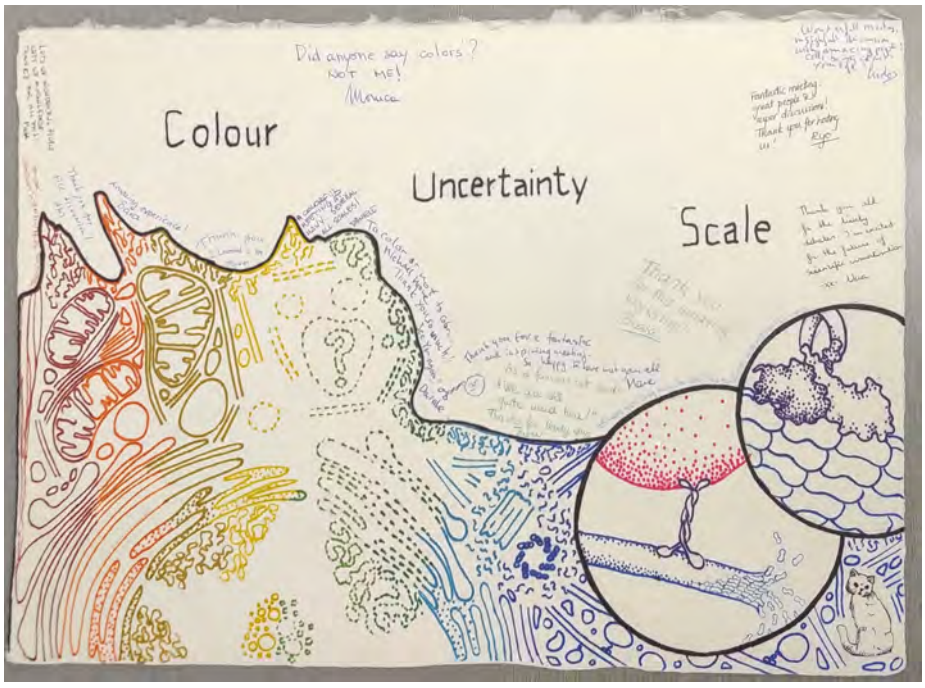


Figure 1: Yosegaki

The art and science of molecular representation has been developed in the last decades to a high degree of sophistication. Several solutions have been proposed to show the invisible; for objects (proteins, cellular parts, small molecules) it is now relatively straightforward to obtain a 3D shape that can be shown in print, video or other 3D interactive ways. Other invisible concepts, such as potentials, pH, motion and environmental conditions, often important for the comprehension of biological phenomena, can be displayed in various ways.

The aim of the workshop is to discuss possible paths that may lead to the harmonization of these representations. Experts from different fields (structural and cellular biology, computer graphics, visual design, education, information design and more) will explore the many options technically available in terms of content communication, ease of use (in the production and in the fruition), and aesthetic quality.

While the potential developments around the subject are much more than it can be discussed in a week, we hope to conclude the workshop with a general agreement on the use of symbols, colors and other graphic means, and with a proposal for developing an implementation avenue, such as a research project for testing and evaluating the proposed ideas.

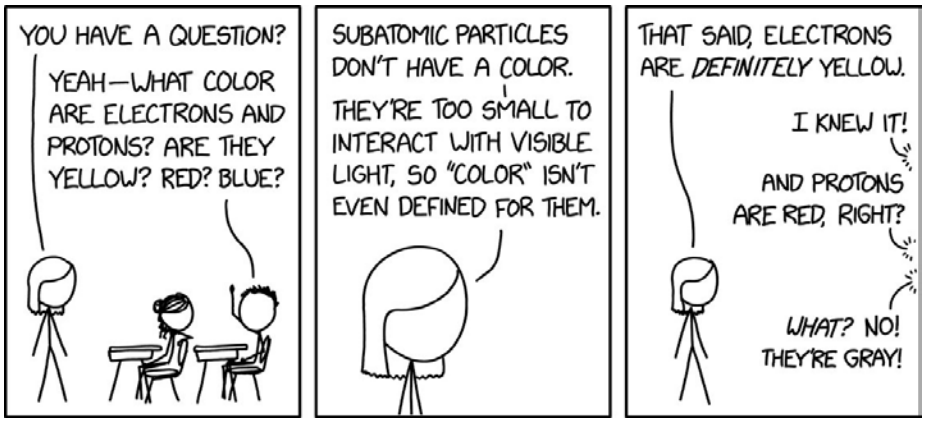


Figure 2: Source: https://xkcd.com/2734/

List of Participants

- Maria Teresa Artese, IMATI - CNR, Italy
- Ludovic Autin, Scripps Research Institute, USA
- Marc Baaden, IBPC - CNRS, France
- Muyuan Chen, Stanford University, USA
- Hsiwen Fan, National Taichung University of Education, Taiwan
- Cinzia Ferrara, Dept. of Architecture, University of Palermo, Italy
- ★ Daisuke Inoue, Kyushu University, Japan
- ★ Yoshie Kiritani, Chiba University, Japan
- Akihiko Konagaya, Keisen University, Chem-Bio Informatics Research Instiute, Molecular Robotics Research Institute, Co., Ltd., Japan
- Barbora Kozlikova, Masaryk University, Czech Republic
- Michael Krone, Stuttgart University of Applied Sciences, Germany
- Beata E. Mierzwa, University of California San Diego & Beata Science Art, USA
- Ryo Mizuta, University of Cambridge, UK
- Paola Perin, University of Pavia, Italy
- Daniele Savasta, Izmir University of Economics, Turkey
- Vera Williams, Institute Pasteur, France
- ★ Monica Zoppè, IBF - CNR, Italy

★ Organizers

Meeting Schedule

Check-in Day: May 11 (Sun)

- Welcome Banquet

Day1: May 12 (Mon)

- Presentation by all participants
- **Cell and molecular graphics / representation**
  - Monica Zoppè. *Introduction*
  - Marc Baaden. *Grounding visualization standards in structural and cell biology*
  - Muyuan Chen. *The game engine based cell rendering work: to bridge the gap between scientific visualization and artistic illustration, enhancing our understanding of complex phenomena*
- **Molecular visualization**
  - Barbora Kozlikova. *Visualization of 3D chromatin fiber*
  - Ludovic Autin. *The Mesoscale Challenge*
  - Beata Mierzwa. *Visual Metaphor and Artistic Representations in Science Communication*
  - Michael Krone. *To Color or not to Color? Uncertainty-Aware Visualization of Biomolecular Structures*

Day2: May 13 (Tue)

- **Color perception, cognition, and theory for visualization**
  - Yoshie Kiritani. *Handling precautions of colors from psychological perspectives*
  - Hsiwen Fan. *The Function and Impact of Colour in Data Visualization*
  - Maria Teresa Artese. *Color theory and tools for color coding*
  - Ryo Mizuta. *The Use of Illustrative Schematics in Interdisciplinary Academic Publishing*
  - Cinzia Ferrara & Daniele Savasta. *'Colors are words': categorical color schemes and scientific visualization*
- Group Photo Shooting
- **Visualization tools**
  - Daisuke Inoue. *MR-based visualization of 3D protein models*
  - Paola Perin. *Make 3D molecular worlds interactive: an IA guide?*
  - Akihiko Konagaya. *The Effect of Molecular Coloring in Comprehensive Real- time Interactive Simulation (CRIS)*

Day3: May 14 (Wed)

- Vera Williams. *Making Science Tangible*
- Monica Zoppè. *A modest proposal – introduction to group discussions*
- Discussions in groups
- Excursion and Main Banquet

Day4: May 15 (Thu)

- Talks and Discussions
- Groups

Day5: May 16 (Fri)

- Talks and Discussions
- Wrap up

Overview of the meeting

Background and introduction

Images of biological entities, from single molecules to whole cells, are increasingly used both in research and in media aimed at the public. Although the objects of such images are inherently colorless, in most cases colors are introduced, either as markers of specific features, or with purely aesthetic intent.

The results are often pleasing, sometimes spectacular. However, we wonder if colors could not be also a channel to deliver additional information to viewers, if they were used in a consistent manner, while still obtaining beautiful images. The community concerned with the notion of 'molecular graphics' includes professionals from a variety of fields, including cell and structural biology, bioinformatics, physics, computer graphics, and art. During the Design for Bioinformatics [23] workshop in 2022, it emerged that developing a set of guidelines for the use of color, as proposed in [30] and [24], could help both creators and viewers.

Accordingly, we convened a special workshop that gathered a very diverse group of experts to focus on the exchange of expertise and ideas and on the collaboration towards the common aim of introducing meaningful colors and other visual signs in the representation of small scale biology. The first workshop on Color and Visual Strategies in Structural and Cellular Biology was hosted by the Lorentz Center in July 2024, and resulted in the initial steps of such endeavour.

The Shonan meeting presented here is the second one, after the Lorentz workshop with the same title as the one held in Leiden. The list of participants, the results and an outline of the discussions of the Lorentz workshop are added as an appendix to this report (see Appendix A).

We report here the major steps taken during the first meeting and some conclusions, as introductory background for the present workshop.

Similarly to the Shonan meeting, the initial step has been a self-introduction by all participants (see Appendix A.1), which highlighted the wide range of disciplines represented, and revealed the necessity to make an effort to find a 'common language' to be able to understand each other (see Appendix A.3).

The discussions covered many different topics, and revolved around issues of colors in general and in scientific representations (Appendix A.4), in relation to the specific content to be represented, including issues of properties, forces, scale, etc..(A.5, A.6), We also considered the distinction among Data and Interpretation of data, and explored the tools to be used in various situations also considering the intended audience, as reported in A.7.

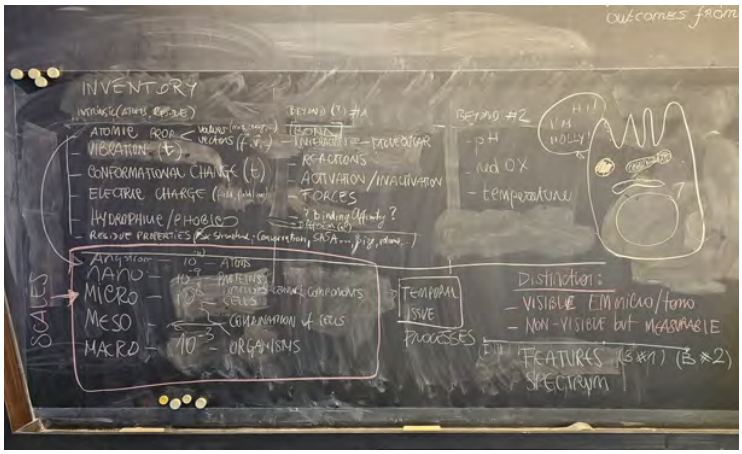


Figure 3: WorkInProgress@Lorentz

	NATURE OF DATA (VALUES)					TIME DEPENDENCY	MEASURABILITY / VISIBILITY
	categorical value (label)	spatial value (position/orientation)	scalar value	vector value	(tensor value)	0-2	0-5
RINSIC							
tion		x		x		1	5
ge			x			1	2
ation				x		2	4
s			x			0	0
d / velocity				x		2	4
				x		1	0
ophilic / hydrophobic	x	-	-	-	-	0	0
LECTIVE (*EXTRINSIC)							
ormational change				x		2	5
ds	x			x		0	2
action		x		x		1	2
tion						2	0
ation/inactivation						1	0?
VIRONMENTAL*)							
			x			1	4
x			x			1	4
perature			x			1	4

ATOMS (Å) | MOLECULE (Å) | MACROMOLECULE (Å) | MACROMOLECULAR ASSEMBLY (nm) | NUCLEUS (µm)

Figure 4: Tentative inventory of visualizable items

During the workshop we made an attempt at building an inventory, or 'catalogue raisonné' of everything we wish to represent. In summary, the Lorentz meeting primary outcome was a thorough dissection of the many interrelated issues that, in different situations, concur to the complexity of the work underlying the representations of cellular and structural biology.

This led the group to assess a wide range of topics, which were discussed at length, revealing in the end more problems than solutions, but that also helped each participant to focus on specific subjects.

On the basis of the work at Lorentz in July 2024, we organized the Shonan meeting in May 2025. Five of the participants attended both workshops, ensuring continuity.

Overview of Talks

Grounding visualization standards in structural and cell biology

Marc Baaden, Université Paris Cité, CNRS, Laboratoire de Biochimie Théorique, 13 rue Pierre et Marie Curie, 75005, Paris, France

Visualisation in structural and cell biology faces a fundamental challenge: the creation of uniform, consistent standards from the atomic level upwards. While visualisation tools have made dramatic progress, the development of uniform conventions remains elusive. This contribution focuses on the development of visualisation standards from the atomic level upwards and explores the complexity of their coherent translation to molecular, cellular and tissue representations. Current visualisation practises have evolved along different historical paths, leading to conflicting conventions in different subfields. For example, the CPK colour scheme for atoms is widely used in education but is rarely used for research purposes as it competes with numerous alternative schemes that are optimised for specific research questions. These inconsistencies make interdisciplinary communication and data integration difficult. The challenge goes beyond technical considerations, as cultural biases significantly influence colour perception and interpretation, and different scientific communities develop different visual languages. Personal aesthetic preferences further complicate standardisation efforts, as researchers often change visualisation parameters to emphasise the features they consider important. Furthermore, accessibility considerations for colour-blind researchers continue to be implemented inconsistently. Rather than proposing rigid conventions, this work motivates the search for a framework for developing more comprehensive visualisation standards that account for this complexity while providing flexibility for different contexts and needs.

Introduction

The visualisation of scientific data is becoming increasingly sophisticated, but there is a lack of coherent standards for colour representation at different scales in this field. This contribution presents ideas and considerations ranging from practical implementation challenges to philosophical questions about the nature of scientific visualisation standards. Using the visualisation of proteins as a representative case study — as proteins can be considered the "hydrogen atom of molecular graphics"— problems are highlighted that apply to other biomolecules such as DNA, RNA, glycans and lipids. The fundamental topic addressed here is not only of a technical nature, but touches on deeper questions of scientific communication, accessibility and intercultural understanding in an increasingly globalised research environment.

The Challenge of Inconsistent Standards

Even at the most basic level of atomic representation, there are significant inconsistencies between visualisation programmes. An examination of the colouring of the periodic table on different software platforms (Fig. 5) shows that PyMol,

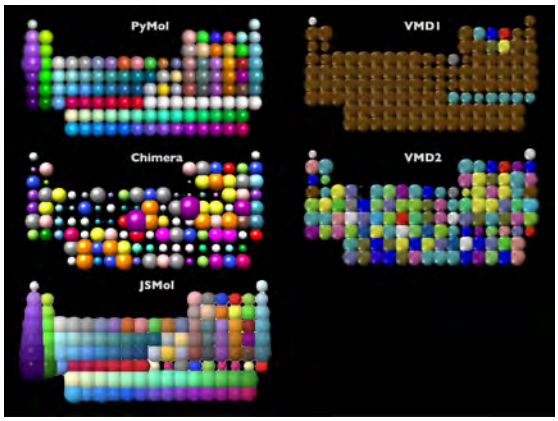


Figure 5: Periodic table colors from different software packages.

Chimera, JSMol and VMD only colour four atoms consistently: Hydrogen, Oxygen, Nitrogen and Sulphur. This observation is consistent with the concerns raised by Brecher [3] , who pointed out the discrepancy between the total number of elements and those typically used in biological contexts, primarily carbon, hydrogen, nitrogen and oxygen, along with chloride and various cations. The historical roots of today's colour schemes can be traced back to an 1865 lecture by A. W. Hofmann, who used croquet balls for molecular demonstrations — a choice that continues to influence modern visualisation conventions [50].

The need for colour coding becomes clear when looking for example at raw molecular dynamics simulation data, which consists only of the positions and velocities of atoms over time. Without appropriate abstractions and colour schemes, such data remains uninformative — a fact easily demonstrated by representing the atoms in molecular dynamics data as indistinguishable grey dots. While there are common representation standards for the different molecular abstraction levels, no colour standards exist yet. This discrepancy is already problematic for representations as common as the cartoon representation, where the colours vary greatly depending on the software platform. It becomes critical when examining colour scales for protein property visualisation, e.g. when depicting electrostatic surfaces, where identical colours can represent opposite potentials in different software packages or communities.

A similar variety can be observed in other property-based colouring schemes for features such as polarity, charge, electronegativity or hydrophilic/hydrophobic nature of residues. For hydrophobicity alone, there are several published scales, including the Wimley-White scale and various alternatives documented in specialised resources (see Experimentally Determined Hydrophobicity Scales), each of which can potentially be assigned to different colour schemes. This proliferation creates a landscape where scientists working on related problems can literally see their data in completely different colours, making communication and collaboration difficult.

The diversity of colour schemes is not the result of random confusion, but of several interrelated factors. Historical factors play an important role, as scientific fields have developed separately and have each developed their own colour standards, often based on the preferences of early pioneers or influential

textbooks. This leads to friction when disciplines overlap or collaborate — a physicist may associate red with positive charge, while a chemist may use blue for the same concept. Technical limitations, particularly old systems such as RasMol's 8-bit colour systems, which are limited to 256 colours, still affect practise today. The differences between screen and print media and the various colour models such as CMYK and RGB make things even more complicated.

Cultural influences and unconscious biases in various scientific circles contribute significantly to the discrepancies. Colour choices often reflect design preferences that focus on visual impact rather than purely functional considerations, with colours having symbolic meanings that vary from culture to culture. For example, red represents oxygen in the CPK scheme, luck in Chinese culture and danger in security contexts. Similarly, green can symbolise nature in some contexts, while in others it stands for toxicity. Although these cultural associations are often unconscious, they influence the researchers' interpretation and choice of colours for their visualisations.

Physiological factors pose an additional challenge, as around one in twelve men suffers from colour vision deficiency. When colour serves as the sole carrier of meaning, important information can become inaccessible to the researchers concerned. Human preferences, including artistic, aesthetic and habitual choices made by individual researchers, further complicate matters. Researchers often change visualisation parameters to highlight features they consider important, resulting in personalised colour schemes that may not be easily understood by a wider audience.

Beyond Simple Standardization

The first impulse could be to establish universal colour standards for atoms and molecules to ensure consistency, clarity and easier interpretation in different disciplines. However, this approach proves insufficient, as a one-size-fits-all solution does not do justice to the different objectives, target groups and established traditions. Forcing a standardised palette risks losing important local, technical and cultural variations that serve specific purposes in their respective contexts.

A more promising approach takes inspiration from keyboard layouts that have been customised for different countries and languages. Just as we use QWERTY, AZERTY and other regional keyboard layouts, scientific visualisation could benefit from interchangeable colour palettes that are adapted to regional contexts and take into account different cultural symbolisms and conventions, individual needs that consider contrast requirements and colour vision deficiencies, and disciplinary conventions that respect established practises in chemistry, physics and biology.

This concept extends to reader-driven colour selection, where scientific software and publications should support dynamic colour palette changes. Each figure would include comprehensive metadata indicating the identification of the base palette, the available alternative palettes and a clear mapping between data elements and colours. With this approach, the burden of standardisation lies with journal publishers and editors, who should require authors to provide colour-accurate images. The proposed workflow requires authors to submit their papers with defined palettes selected from a standardised set, allowing readers to switch between palettes according to their needs, accessibility requirements

or cultural preferences.

The interconnectedness of biological scales requires a coherent transfer of colour conventions from atoms to residues, chains, molecules, cells, organs and organisms. Atoms are the smallest scientific objects that are routinely visualised with complex colour requirements. They are ubiquitous in molecular models, crystal structures and simulations and require consistent, three-dimensional, information-rich visual representation. While smaller objects such as quarks and gluons also exist, they are less frequently visualised and their colours often have a symbolic rather than literal meaning, such as the "colour charge" in quantum chromodynamics.

Visual continuity across scales is crucial — if blue represents positive charge at the atomic level, this association should persist across all relevant scales to avoid confusion and maintain semantic consistency. This challenge of coherence applies not only to biology, but to all fields where visualisation spans multiple levels of organisation. The symbolic meaning assigned to colours must be consistent across all scales, starting at the basic atomic level, to ensure that researchers can move seamlessly between different levels of detail without losing conceptual coherence.

A Proposed Tool: Colors of Science

To systematically address these challenges, we propose the development of a "Colours of Science" catalogue — a living, structured reference system for scientific colour standards. This database would serve as a comprehensive repository containing colour sets for atoms, residues, molecules and other scientific objects, standard palettes of common software packages and published colour schemes, unique identifiers similar to DOIs for each colour set, comprehensive metadata including version information, search capabilities by element, object, software or source, side-by-side comparison capabilities, and integration support for image creation and software development.

The proposed system would organise the information around three core components: Colour sets which represent collections of colour mappings, such as RasMol standards, Jmol standards, or schemes published in the literature; scientific objects, including elements, residues, molecules, cell types, and other entities that require colour representation; and colour entries which establish specific connections between objects and colours within defined colour sets. This structure allows researchers to add new colour sets, define new scientific objects and create mappings between objects and colours, while supporting collaborative input, filtering, comparison, palette export and integration with existing visualisation tools.

An illustrative example would be to verify the representation of hydrogen atoms in different colour sets. The RasMol software - as many packages in the field - traditionally uses white for hydrogen as part of its classic scheme, while the Jmol software uses light grey as part of its open source standard. In individual publications, hydrogen may sometimes be defined as blue depending on the conventions defined by the author. This systematic approach allows each atom, residue or concept to be examined across multiple colour sets, facilitating informed decision making and promoting consistency where appropriate, while retaining flexibility where it is required.

The Colours of Science catalogue is intended as a community-oriented refer-

ence tool that can be easily integrated into visualisation tools and publications using an API-based interface. The system aims to promote accessibility, consistency and colour literacy in science, with potential expansion beyond the molecular sciences to other fields that require complex colour coding. Success depends on broad community acceptance and participation, which will require collaboration with software developers, journal publishers, professional organisations and individual researchers. The system must offer clear added value to users while minimising barriers to participation.

Publishers could play a critical role in driving adoption, and journals should establish requirements for colour images, including mandatory colour palette specification for all colour-coded images, support for alternative colour versions, clear documentation of colour to data mapping, and accessibility verification. The initial implementation could use existing database technologies that support visual interfaces and collaborative input and require storage of colour information, metadata management, and search and export capabilities that are compatible with common visualisation software.

Conclusions

The challenge of visualisation standards in structural and cell biology goes far beyond technical considerations and includes cultural, physiological and practical factors that resist simple approaches to standardisation. Rather than imposing rigid, uniform conventions, the scientific community needs a flexible framework that recognises diversity while promoting coherence and accessibility. The proposed Colours of Science catalogue represents one approach to systematically address these challenges by creating a comprehensive, searchable repository of colour standards with robust metadata and comparison capabilities, enabling researchers to make informed decisions about the use of colour while being aware of the alternatives and their implications.

Future work should focus on developing prototype implementations of the Colours of Science concept, engaging stakeholders from all relevant scientific communities, building partnerships with journal publishers and software developers, creating guidelines for the accessibility of colours in scientific visualisation, and extending the framework beyond the molecular sciences to other fields. The ultimate goal is not uniformity, but informed diversity — a visualisation ecosystem in which colour choices are consciously made, documented and adapted to the needs of different audiences, while maintaining semantic coherence across scales and contexts.

This work is a first step towards a more thoughtful approach to scientific visualisation standards that acknowledges complexity while working towards solutions that serve the communication and accessibility needs of the broader scientific community. The ideas presented here have been developed based on discussions at previous workshops and benefit from genuine interest and feedback from the visualisation community. The focus on protein visualisation serves as an illustrative example of broader challenges that extend to the visualisation of other biomolecules and scientific objects at different scales, ultimately contributing to a more inclusive and effective scientific communication landscape.

Make 3D molecular worlds interactive: an IA guide?

Paola Perin, Department of Brain and Behaviour Science, University of Pavia, Italy

One of the critical issues in the exploration of the molecular world is that an unsupervised approach, although allowing maximal freedom, often reduces the learning value of the experience, especially for less expert users. A study comparing student awareness of scales before and after the vision of the documentary "Powers of ten" [53] found out that the effects of watching it were limited for unfamiliar regions of the scale spectrum (such as DNA size)[9]. The problems in learning to navigate an unfamiliar world (like the cell at molecular level) come from the presence of a myriad of different objects (which saturate our working memory, impairing retention of information) and from lack of understanding of their function (which impairs our active interaction, our manipulative "playing" with the structures). Building interactive learning scenarios is very complex and time-consuming but allows for an unrivalled exploration and understanding of the molecular world, as shown in [15]. The rise of generative AI cannot be ignored in this regard- in particular, AI chatbots could be used as interactive commentaries on the scenarios, with a language tailored to the user. However, although these tools are known to be very impactful on learning [7] they may also introduce illusions of depth, breadth and objectivity [17] which could negatively affect the deeper forms of learning, such as those requiring active exploration of the environment and manipulation of the objects therein.

Chromatin visualization: gaps in representation

Barbora Kozlikova, Masaryk University, Brno, Czech Republic

Chromatin fiber is inherently a multiscale structure that can be visually inspected on several levels of detail. Although in some of the levels, we can be inspired by the protein visualization principles, there are other levels that completely lack any standard for representation, including color usage. In this talk, I primarily focused on presenting the current state of the existing approaches to 1D, 2D, 3D, and 4D (including dynamic behavior) of chromatin visualization. I presented several visual abstractions that we recently designed to communicate different features of 3D chromatin spatial arrangement (e.g., hierarchical nature of subcompartments or spatial relationships between subcompartments – see Figure 6) and its dynamic behavior (e.g., animation or aggregated view using volumetric representations).

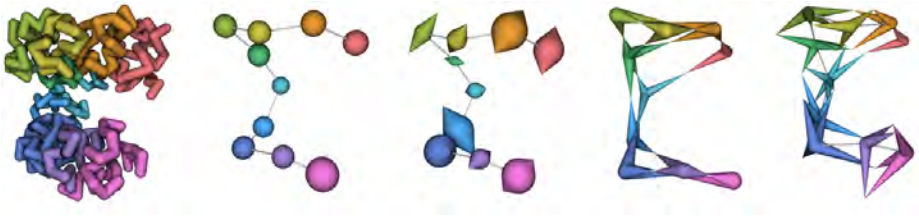


Figure 6: Visual abstractions of chromatin subcompartments [20].

I also presented our recent approach to visual exploration of 3D chromatin structure in a virtual environment. Finally, I outline the major challenges and gaps in both obtaining the data on different resolutions and their appropriate visual representation.

The Mesoscale Challenge

Ludovic Autin, Scripps Research, USA

Current science and technology have reached a point where building detailed models of cellular environments at the molecular and atomic scales, visualizing them, and interacting with them is feasible but very challenging. In this talk, Ludovic showcased several innovative approaches he has developed—such as cellPACK, cellPaint, cellPaint-VR, and Mesoscale Explorer—to tackle these challenges head-on. Ludovic's work as an Institute Investigator at the Scripps Research Institute intended to bridge the gap between scientific visualization and artistic illustration, enhancing our understanding of complex phenomena like viral self-assembly, intricate DNA structures, and crowded cellular environments. By leading the development of advanced software frameworks for multiscale molecular modeling and visualization, Ludovic has helped model diverse molecular landscapes—from tiny exosomes and viruses to bacteria and cellular organelles such as the insulin secretory granule and Mycoplasma genitalium.

Visual Metaphor and Artistic Representations in Science Communication

Beata E. Mierzwa, University of California San Diego, USA; Beata Science Art, USA

Visuals are an integral part of both research and communication, and using images that combine biological data with visual metaphor creates unique and effective ways to communicate scientific information both inside and outside the scientific community. Alongside my research on cell division, I explore approaches to visual science communication with hand-drawn illustrations, science fashion, and interactive media. Conveying complex biological concepts using visual metaphor, abstract imagery, and real scientific data places abstract concepts into a familiar context, making visuals engaging across various levels of expertise. This form of visualization provides a powerful tool to facilitate communication between researchers from diverse fields, promote creativity in scientific research, as well as spark fascination in our next generation of scientists. In this talk, I will share my experiences in using artistic representations to communicate complex ideas in intuitive ways and explore ideas to initiate a broader conversation on the benefits and limitations of visualization standards in this context.

To Color or not to Color? Uncertainty-Aware Visualization of Biomolecular Structures

Michael Krone, Stuttgart University of Applied Sciences, Germany

Molecular structure visualization is a fundamental aid for understanding complex biological processes. While molecular visualization has greatly improved over the last five decades, inherent uncertainties—such as molecular flexibility, inaccuracies in atomic positions, or variability in secondary structure classifications—are often still not shown, thus negatively impacting the accuracy of the visualizations. Uncertainty-aware visualization (UAV) integrates such uncertainties into established molecular representations to improve data interpretation and decision-making. While color is one the most widely used visual channels to express uncertainty, it often leads to a loss of information in molecular visualization, since color is usually already used to express other molecular properties. In my talk, I will review existing UAV approaches for biomolecular structures and discuss our recent contributions to this field. Figure 7 shows examples of uncertainty-aware visualizations for protein structures.[62]

Figure 7: Examples for uncertainty-aware visualizations of protein structures. Color (left) is a common visual channel to express uncertainty [21], yet color is typically used to express other properties in molecular graphics. Therefore, other channels like transparency and geometric distortion [52] (center) or illustrative methods like contour lines [25] (right) have been investigated.

Handling precautions of colors from psychological perspectives

Yoshie Kiritani, Chiba University, Japan

This talk presented psychological insights into how we use colors. Two key points were discussed: (1) increasing the lightness contrast between the figure and the background, and (2) considering the cognitive aspects of color. Each point included three topics. For the first point: the functions of color, color illusions, and color vision deficiency. For the second point: color illusions, emotional or symbolic meanings of color, and again, the functions of color. The talk concluded by emphasizing the importance of psychological experimentation.

The primary function of color in terms of lightness contrast is visibility—that is, how easily an object can be noticed when one is actively searching for it. Increasing the contrast between figure and background enhances visibility. Among the elements of color, lightness plays the most significant role. Differences in chroma or hue are less influential. In discussing visibility, the Leibmann effect is especially important. This visual illusion demonstrates a problematic color pairing—such as red and green—that causes a flickering effect and visual discomfort. To avoid this, increasing the lightness difference between areas can reduce the illusion's negative effects. This approach also helps address color vision deficiencies. People with protanopia or deuteranopia struggle to distinguish red and green, but enhancing the lightness contrast between these colors can improve legibility and perception.

The second key point emphasizes avoiding miscommunication through color by considering its cognitive dimensions. The Stroop effect is a classic example of how color and language interact cognitively: when the meaning of a color word interferes with naming the color it's printed in. Symbolic and emotional associations with color must also be considered. At the symbolic level, color meanings vary by culture and age. A Japanese example involves long-running children's TV shows featuring color-coded heroes. These characters demonstrate how colors can symbolize personality traits—not just appearance but also inner qualities. However, these meanings are fluid and can shift depending on time and cultural context. These hero series also highlight another function of color: distinguishability—the ability to tell elements apart. This connects to the limits of human short-term memory. When the number of elements is small, recognition works more efficiently. Finally, it's important to remember that the pairing of color and element is arbitrary.

The Function and Impact of Color in Data Visualization

Hsiwen Fan, National Taichung University of Education, Taiwan

When discussing data visualization, colour is not merely a decorative element, but also a powerful tool for conveying information. The role of color extends far beyond aesthetics; it significantly influences the emotional response and cognitive interpretation of the audience. Designers need to understand how color is used to highlight information, create visual hierarchy, and direct the viewer's attention to the most important content. At the same time, color choices should strike a balance between functionality and emotion, ensuring that the data is not only clear and understandable but also evokes an emotional connection. Furthermore, designers must consider color accessibility, ensuring that visuals are effective for all viewers while avoiding potential cultural misinterpretations. In conclusion, color is essential for data visualization, and designers must understand its functions to create visuals that are both visually appealing and effectively convey the intended information.

Color theory and tools for color coding

Maria Teresa Artese, Institute for Applied Mathematics and Information Technologies, Milan. CNR of Italy

Data visualization is the process of transforming data into information. Effective visualization helps users to understand how data works and provides them with valuable insights. The use of color to encode information improves the observer's understanding and ability to remember the information depicted by images. The presentation of data in graphical form is achieved by the use of visual variables, that are processed immediately at the sensory level of the human eye, allowing the perception process of recognizing, organizing and interpreting sensory information. As three of these variables relate to color, allowing for associative, selective and ordered perception, it is necessary to have knowledges of color reference systems, color perception mechanisms and emotional reactions.

The sensation of color is produced in the brain in response to light received by the eye. In fact, light is the only stimulus for vision. Human color vision is trichromatic: the retina has three types of color-sensitive photoreceptor that perceive long, medium, and short wavelenghts, corresponding to red, green, and blue light. Starting in 1929 with Wright and Guild's studies, color matching experimentation led to the development of various color space models that describe human perception. Each model has different characteristics and variables. Several stages were involved in the development of the HCL (hue, chroma, luminance) color space, which is a mathematical transformation of the CIELUV and CIELAB spaces, both of which are perceptually uniform and device-independent. This makes it possible to compute the color differences using the Euclidean distance.

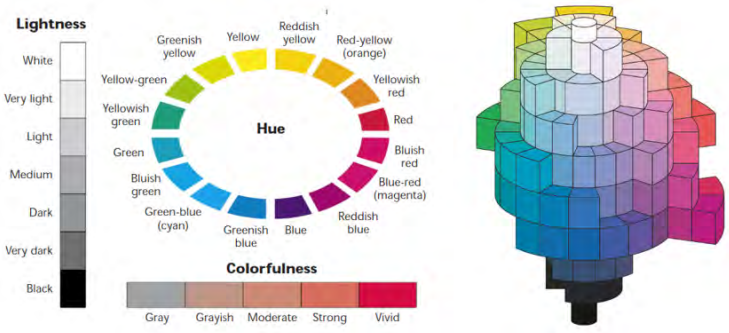


Figure 8: HCL Color Space.

The HCL color model is particularly useful for specifying individual colors and color palettes as the three axes closely match those of human vision. Such color models provide the basis of color coding, which is the process of mapping information space onto color space.

The process is ruled by six principles: the chosen colors should be discriminable, detectable, the steps should be perceptually equal, the code should be relevant to the user, consistent and aesthetically pleasing [44].

It is important to understand what type of information we have to code:

- when information can be arranged in a numerical series and is defined by its quantity or amount, such as population, sea levels, or temperatures, the best choice is to use small differences in hue, or a saturation or luminance code. These scales emphasize the relationship between the different levels of information.
- when information is categorized by quality or nature, with distinct groups defining categories such as the names of countries or species of plants, it is preferable to use significant differences in hue to emphasize this categorization, applying the principles of color harmony selection, as shown in Figure 9, from Ms. Weeya's Art Class.

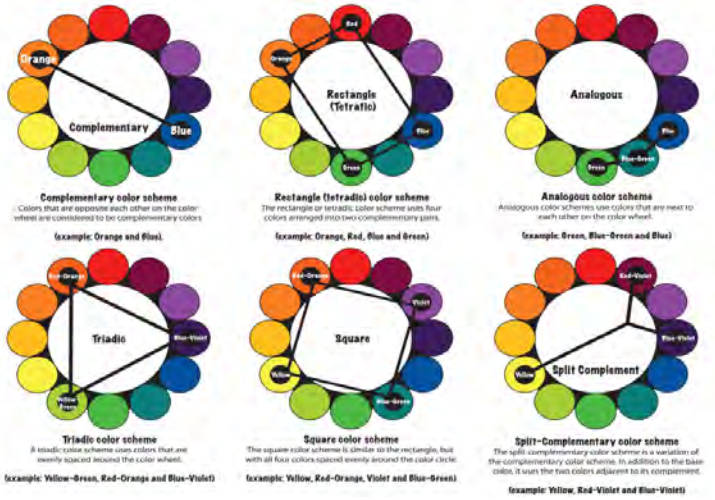


Figure 9: Color harmony principles.

Among the tools available for selecting colors for color coding we can mention Color Brewer 2.0, which is mainly used in cartography field, and Adobe Color, which follows the rules of color harmony selection. Both tools are interactive and can be used to produce and test different types of color codes. They also have options to ensure color-blind and print-safe color palettes.

‘Colors are words’: categorical color schemes and scientific visualization

Cinzia Ferrara and Daniele Savasta, University of Palermo, Italy and Izmir University of Economics, Turkey

Among the visual variables available to information designers to represent data, color is uniquely complex. It enables almost immediate comprehension by the readers while carrying rich cultural references. Like other visual variables, the key challenge lies in balancing aesthetic appeal with clear communication. Our attempt is to suggest a color scheme for molecular visualization based on the experience in the design of categorical color schemes for statistical representation.

Authors like Edward Tufte [46] recommend creating balanced palettes with distinguishable and harmonious colors, particularly by maintaining low saturation levels. The information design community has recently seen the emergence of several categorical color schemes, driven largely by the growth of visualization tools such as D3.js and Tableau. These statistical visualization tools have conditioned the development of reliable, ready-to-use categorical color schemes, helping designers create readable visualizations that align with the recommendations of leading scholars. Designers can either use these proven color palettes directly for quick results or use them as a foundation to develop their own schemes.

We can initially identify different strategies that designers adopt to select a palette for their projects. Common practices like deducing the palette from the existing context through relief and sampling techniques can lead to coherent or opportunely discordant palettes. Furthermore, the presence of accent colors can reshape the perception of spaces and their mood or perceived size.

The question that emerges though is "how we might adopt colors in a context that does not possess inherently color or in which color cannot be perceived nor being meaningful?" Researchers and visualization experts are then free to choose, and without established best practices within the software and scientific community, there has been a wide range of variations and interpretations of the standard, often leading to poor aesthetic results and difficult communication. Taking Observable10 [58] and Tableau10 [59] as the foundation, respectively, of the color scheme of D3.js and Tableau, we explore creating a color scheme for molecular visualization that follows expert recommendations while providing the necessary flexibility and extensibility. Our challenge extends beyond creating an original standard - we must expand beyond the ten core colors of categorical schemes to develop strategies that are both memorable and semantically meaningful within the scientific community. In essence, we need a new color vocabulary to tell a new story.

“For me, a color is like a word. Colors are words. Stories can be told with colors, and colors tell stories” [43].

MR-based visualization of 3D protein models

Daisuke Inoue, Kyushu University, Japan

Understanding the complex three-dimensional (3D) structures of proteins is fundamental to structural biology, but making these structures comprehensible to non-experts requires innovative approaches. In this study, we propose a novel method using Mixed Reality (MR) technology to project 3D protein models into the real world, enhancing accessibility and engagement. Multiple proteins are rendered simultaneously, with colors assigned based on the emotional or functional "feelings" evoked by specific hues. Drawing inspiration from Japanese popular culture, such as Power Rangers and Sailor Moon, where character colors reflect personality or roles, we apply a similar rule to proteins: for example, dynamic proteins like kinesin are assigned reddish tones to signify motion and energy. This intuitive color-coding strategy, combined with MR and VR, enables real-time interaction with molecular structures, offering significant potential for education and research. MR and VR technologies bridge the gap between abstract data and tangible visualization, making complex biomolecular concepts accessible to diverse audiences (Figure 10).

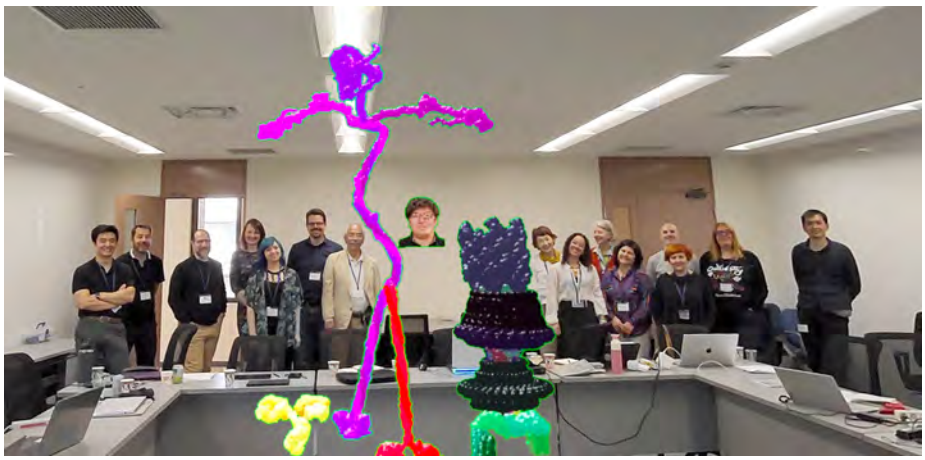


Figure 10: A group photo with virtual proteins taken by MR system and the passthrough camera of Meta Quest3

At the conference, participants had an experience this MR-based visualization through hands-on interaction with actual devices, demonstrating its practical application in scientific communication.

The game engine based cell rendering work: to bridge the gap between scientific visualization and artistic illustration

Muyuan Chen, Stanford University, USA

While the development of imaging and simulation techniques makes it possible to identify the various proteins and RNAs inside cells and determine their structures at atomic resolution, it remains challenging to render the complex cellular environment at the atomic level. One of the main hurdles in cell visualization is to simultaneously display the millions of molecules in real time. Even with the acceleration of modern GPUs, rendering the atomic model of every protein in a cell still requires enormous computational resources.

Here (Figure 11, from ref. [4]) using Unreal Engine 5 (UE5), a video game engine, we demonstrate the capability of rendering massive biological macromolecules at the atomic level within their native environment. The Nanite system featured in UE5 renders objects at different levels of detail automatically based on the distance of each object from the viewer, drastically reducing the number of triangles needed to display at any given moment. Using the game engine, we rendered a salmonella minicell, invaded by bacteriophages, at 3 Å resolution. The scene includes more than 40,000 objects, composed of around 1.5 billion triangles, which can be rendered in real-time with a consumer GPU.

In addition to generating static images and videos, the game engine based system also allows users to navigate inside a cell, both on a 2D screen and in 3D in a virtual reality (VR) system, and explore the diversity of macromolecules interactively.

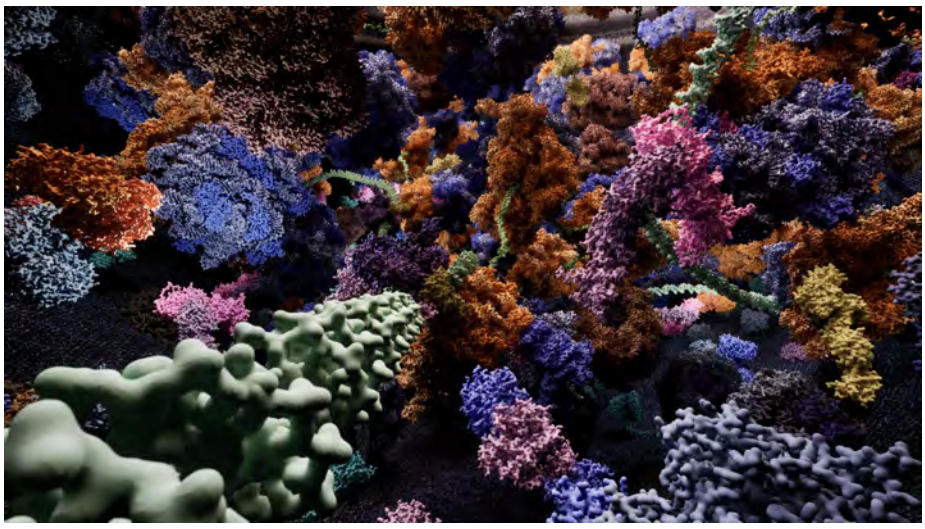


Figure 11: Salmonella minicell rendered with UE5

The Effect of Molecular Coloring in Comprehensive Real-time Interactive Simulation (CRIS)

Akihiko Konagaya, Keisen University, Chem-Bio Informatics Research Institute, Molecular Robotics Research Institute, Co., Ltd.

Molecular simulation is a well-established methodology for understanding molecular interaction mechanisms at the atomic level. However, conventional molecular dynamics simulations often focus on small-scale interactions, such as protein-ligand binding, primarily due to the computational challenges associated with calculating femtosecond time steps. This is essential to ensure consistency with experimental data, including binding free energy and diffusion coefficients. To overcome these limitations and expand our imagination, we have developed the Comprehensive Real-time Interaction Simulation (CRIS) system, which enables the observation of large-scale molecular interactions in a virtual reality (VR) environment [11]. In this presentation, I will highlight the significance of molecular coloring in visual inspections, which aids users in intuitively understanding the dynamics and behaviors occurring during the simulation as shown in the Figure 12.

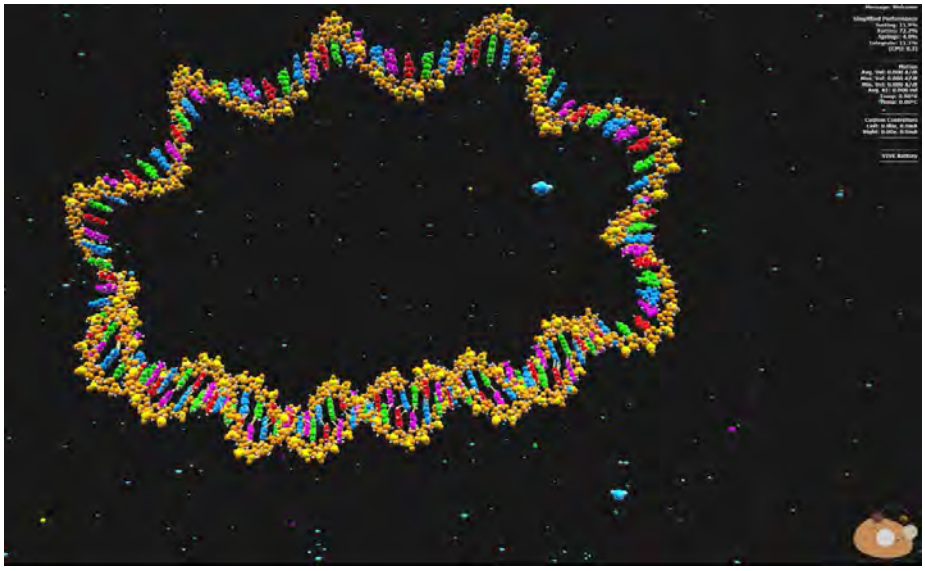


Figure 12: A ring DNA on VR

The Use of Illustrative Schematics in Interdisciplinary Academic Publishing

Ryo Mizuta, University of Cambridge

This talk explores the evolving role of illustrative schematics in academic publishing, with particular emphasis on their use in interdisciplinary research. Schematics have undergone a significant transformation over the recent decades, from sparse and monochrome line drawings to complex, highly stylised visuals, including photorealistic 3D renderings. I present a survey of recent publications in interdisciplinary journals that particularly highlights a growing trend toward 3D schematic use in high-impact outlets. Such a shift in visual styles may be a natural consequence of changes in publication standards, advances in digital illustration tools, and shifting expectations around scientific communication. However, it equally raises concerns about the suggestive power of visuals in shaping research perception and inflating scientific credibility, especially among non-specialist technical audiences in interdisciplinary research spaces. I therefore advocate for rethinking the common distinction between "illustration" and "visualisation," proposing that schematics be treated as data representations subject to the same scrutiny as empirical results. More comprehensive training in visual communication, covering not just tools but design ethics and strategy, are also urgently needed to support responsible schematic use in this evolving landscape of publications. These considerations are especially important in interdisciplinary science, where clear and accurate visual communication plays a critical role in bridging diverse expert domains.

Making Science Tangible

Vera Williams, Institute Pasteur, France.

In this talk I briefly highlight some of the understated tools that I use when creating scientific visualisations. I used "Treponema through the ages" as an introductory piece [62]. This work illustrates some of the methods I use to make science accessible to the general public: storytelling, and mindful use of values, colours and textures. I highlight these tools as they are crucial in my endeavor to communicate about the alien world of molecular biology.

As a scientific visualiser, I see myself as the bridge between the scientific world and the general public. Having close collaborations with researchers is essential to the process as this facilitates accurate depiction of cutting-edge science. When translating the textual descriptions of dynamic molecular events for example, I need access to experts on the subject to ensure that the movements are translated correctly. Such visualisations can then be used to accompany publications for example. In this context the visualisations should be understood as hypotheses but sometimes an issue arises; once something is visualised, it intuitively seems to represents absolute truth, rather than truth in a specific frame of reference. One underused tool to remedy this issue is the tool of storytelling. Storytelling can be applied in animations and illustrations to facilitate a clear frame of reference from which to understand the content in its context as demonstrated in the video. Additionally, storytelling is a very efficient tool to make science more understandable for a general audience in that it can help a lay-person descend into the depth of the molecular world by the guiding hand of a story.

I addressed visual tools as well. I discussed the importance of applying values effectively to separate back-, mid-, and foreground and individual entities in a visualisation. I also demonstrated that the effects of hue might actually distract from creating effective images. We attribute a lot of power to using different hues. However, if we don't pay attention to a colour's value and solely focus on the hue we can actually create suboptimal views due to the inherent differences in value of each colour. Figure 13 demonstrates the ambiguity of hue. The image of the strawberries demonstrates that while the actual hue of the strawberries is teal, our brain recognises the objects as strawberries and we observe them as being red. When creating visualisations we need to be aware of the psychological effects that colours have to apply them effectively.

Besides shifting the focus from hue to values, I aimed to highlight the effects of textures as well, as this is another great yet underutilised means of conveying information in any type of visualisation. Textures can intuitively demonstrate the difference between natural vs. artificial models, healthy vs. unhealthy molecules, or even known vs. unknown structures [61]. This can for example be done by using metallic textures for misfolded proteins in contrast to diffuse textures of regular proteins in molecular visualisations as demonstrated in Figure 14. The contrast between inorganic looking textures in an organic looking setting automatically evokes a sense of wrongness about the molecules.

Lastly, I discussed the versatility of textures in 2D illustrations where it can be used to communicate the 3D shape of a molecule. In this example, stippling is used to create a sense of depth [63] but the applications and techniques are endless.

Figure 13: Ambiguous effect of hue

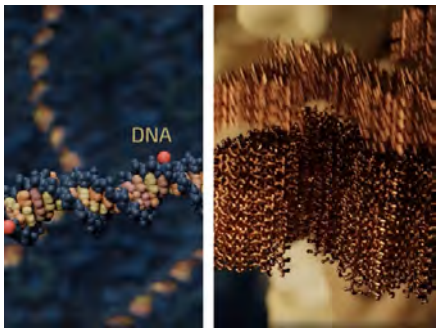


Figure 14: Demonstration of healthy vs. unhealthy molecules by assigning inorganic vs. organic looking textures

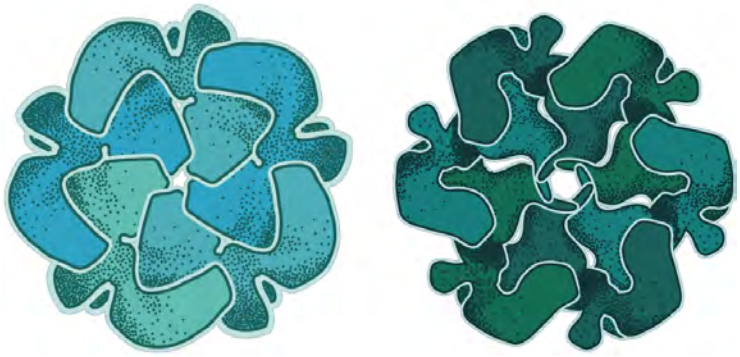


Figure 15: Using stippling to create depth in 3D

In all, this presentation aimed to highlight the use of storytelling and the artistic tools of hue, value, and texture to overcome the hurdle of communicating the alien world of molecular biology and to create effective scientific visualisations. I hope this work demonstrated the versatility of these tools and sparks an interest in exploring their applications.

A modest draft proposal – introduction to discussion in groups

Monica Zoppè, Institute of BioPhysics, Milan. CNR of Italy

The proposal is presented as a basis aimed at defining a path that should eventually lead to reaching a consensus on representing structural and cellular biology. It is meant to serve as ground for discussion, and can undergo serial modifications up to a complete turnaround: I present it because it is usually easier to work on a draft than to start from blank page.

I start by drawing a parallel between the world we know by direct experience (real world for short) and the cellular world, that we know through experimental evidence. I propose that we try to reach a common view of what the cellular world would look like if we were able to directly experience it. After (and if) we reach such a common view, then it should become easier to think about representations.

We know the real world not just through sight: many senses are involved, tactile experience, hearing, and more, including inner feelings (e.g. for temperature) and transmitted knowledge. It is this prior knowledge that allows us to interpret representations that go from verbal descriptions to realistic photographs, to sketches and to symbolic, as illustrated in Fig 16, extracted from *Understanding Comics* [41].

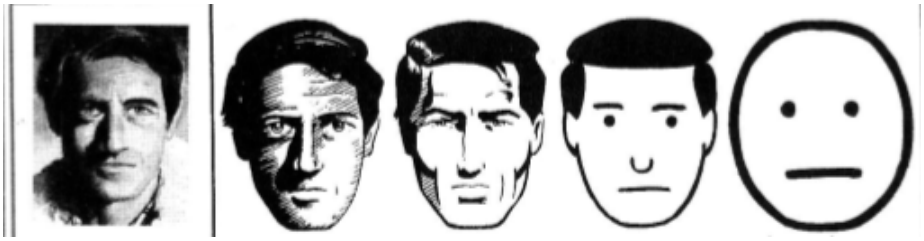


Figure 16: Our knowledge of real faces enables us to recognize one even when synthesized as a cricle, two dots and a line

Unfortunately we don't enjoy the same level of knowledge and experience about the cell world. Nevertheless we do have a vast range of information, which might suffice to build a mental reconstruction of the cells, at different scales. Once one has built a mental representation, then they will be able to set a visual representation (described in words, painted, printed, animated, 2D, 3D and in VR).

This is why I would ask all participants to NOT consider representation, for now, but to concentrate on our own imaginary mental vision.

As we have heard during the talks in the last two days, scaling is a major obstacle, and I suggest that, as a tool to overcame this problem, we adopt the Ten Million Times scale [29]: in this scale, a cell can be measured in km, and the smallest objects (atoms and water) would be in the range of mm. In between, proteins are in the cm range, larger complexes like ribosomes are in the tens of cm, membranes would be 6 to 8 cm thick, and larger objects, like the nucleus and other organelles, would have the size of buildings.

Let's now imagine that we land on a small planet, as large as a town or a city (blood cell to neuronal soma) and start exploring: what do we see? We cannot see everything. For example, on the cell surface, there is supposed to be a layer of proteoglycans of unknown structure. This raises the issue of how to represent uncertainties (see discussion). Once we enter the cell, passing through the lipid bilayer, the cytoplasm is so crowded that we are unable to look further than 10 or 20 cm. This raises the issue of how to actually see anything if the place is too full of objects: a suggestion can be to reduce the N of objects to one tenth of the experimental or calculated abundance.

Another important aspect is that the medium in the cell is not air, but water, which is extremely dense, especially at the very small scale of proteins. Wandering about in the cell could be imagined like swimming, rather than walking or flying. The thickness of water can also be exploited as a method to feel distance (see discussion on scales). Once we have solved the issues of crowdedness, uncertainty and scale, we can concentrate on single objects.

Most of us are interested in proteins and protein complexes. We are used to represent them in the ribbon representation, with or without explicit side-chains. However, just like we see the surface of animals and not their skeletons, if we were to wonder about in a cell, we should expect to see proteins' surfaces, another common way of representing them, for which is easily implemented. I suggest that we exploit the molecular surface, and that we try to imagine how surface properties, mainly electrostatic charges and lipophilic potentials could look like. We have no experience in perceiving these features, so this is one of the topics we could discuss. My personal suggestion is that we avoid using colors and rely instead on texture features and special effects (a suggestion not immediately accepted, nor after).

In the cell there are places that are characterized by different environmental conditions: pH, RedOx power, transmembrane potential and others that we would not (yet) know how to visually describe. My suggestion is that we exploit colored light to create different environments: another suggestion that was not immediately accepted.

The issue of Colors is, as expected, very debatable, as it is connected with all the aspects that were exposed in the talks in the previous days: another topic for group discussion.

[Note 1: despite my effort, I did not manage to have the discussion limited to imagination: everyone, including myself, fell back to description from one point of view, as in representation]

[Note 2: a more detailed version of my proposal will be submitted for publication shortly]

Summary of discussions

Group discussion: Rules for Color Representation in Expressing Cell Structure

Participants: Akihiko Konagaya, Cinzia Ferrara, Daisuke Inoue, Hsiwen Fan, Muyuan Chen, Maria Teresa Artese, Yoshie Kiritani

Abstract

The group discussed the roles of colors in cell architecture visualization and concluded that colors can be used as ontology. A color ontology highlights the notions in knowledge as a tree-like structure so that the relation of the notions and the hierarchic structure are clarified. The meaning of cell architecture visualization can be considered as a kind of encoding-decoding communication model between a writer and a reader through a picture with palettes (textures) and a symbolic color such as "a microtubule is green". In the framework, color ontology plays an essential role in reducing the complexity of images with regards to information entropy; high in grayscale, low in black/white, and colored in the middle (Figure 17).

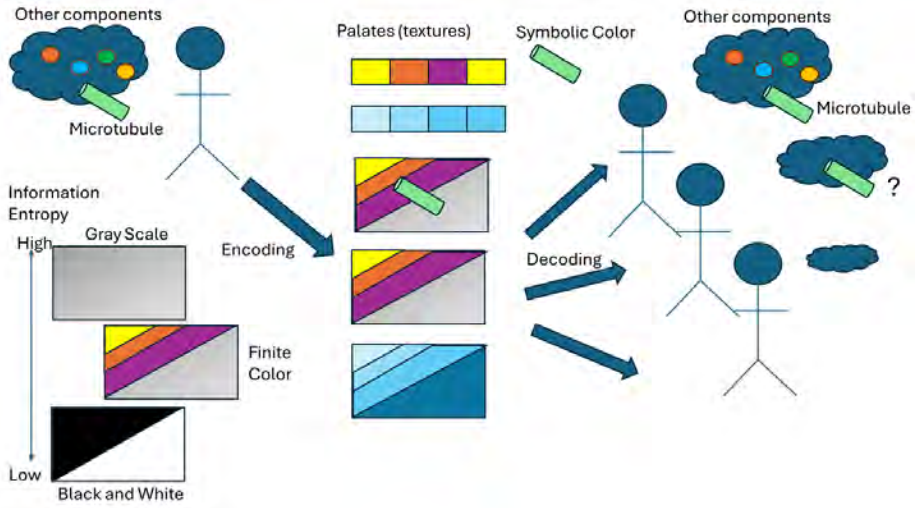


Figure 17: A schematic figure to represent the relation between colors, information entropy, and communication model.

1. Introduction

Cells contain a variety of components, such as proteins, nucleic acids, and organelles, making their structures highly complex. Additionally, the hierarchical organization of living organisms, spanning from molecular to macroscopic scales, further complicates visualization. Color can be a powerful tool for visualizing these intricate biological structures; however, the random selection of colors, color combinations, and the necessity of using multiple colors remain subjects of debate. Our group proposes using information entropy as a criterion to evaluate the number of colors and the significance of using color (e.g., whether

grayscale is sufficient) in the visualization of cellular structures.

Information entropy is a metric that quantifies the uncertainty or diversity of information, aiding in assessing how effectively color choices convey information in visualizations. To implement this, we ontologically classified the hierarchical structure of cells and designed a tree structure. Based on this classification, we conducted a demonstration to color-code the hierarchical structure of the Mycoplasma model by Ludovic Autin and colleagues [54]. Furthermore, using information entropy, we quantified the complexity of the information to be conveyed and designed a prototype program to determine the diversity of colors to be used.

2. Colors as Ontology / Choice of Color Palettes

Firstly, we connected colors with the information entropy for the visualization of micro cellular structures. Two colors, black or white, contain lower information entropy. On the other hand, the grayscale with continuous gradient has higher entropy because of its multiple tones. Finite colors with hue and saturation are located midway in these two situations because of the limitation of color representation. To use colors in this middle class, we introduced the idea of color palettes.

Secondly, we realized the importance of the communication model. A person thinking about cells, for example, in his/her head, is the sender of information. He/She is a specialist, and his/her knowledge is structured. On the other hand, the previous knowledge of people who receive the information is various: the amount and the quality of information obtained depends on their prior knowledge. Thus, which data should be emphasized matters, and the color palettes are useful. Figure 18 represents our color ontology idea of representing prokaryotic cells. The color palettes are used to express the hierarchic structure of prokaryotic cells. According to the level, the saturation of colors in the palettes should be changed.

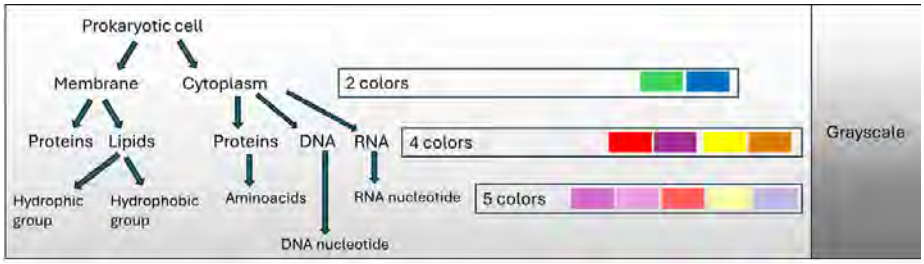


Figure 18: Colors as ontology: a tree structure of prokaryotic cells and an example of used color palettes.

3. Coloring of the Mycoplasma Model Based on Information Hierarchy

We utilized the Mycoplasma model, created by conference participant Ludovic Autin and his colleagues [54], as an example of a prokaryotic organism and performed coloring using the Mesoscale Explorer [55]. Figure 19 is an example of sequential representation using two color palettes applied to the model. In

the figure, (a) and (b) are the original images showing the outside and inside of the whole cell colored as a grayscale color. Figure 19 (c) highlights only the membrane level with pink color, and Figure 19(d) highlights the membrane level and the protein, DNA, and RNA levels. Here, the primary structural components within the cell are colored using a fluorescent color palette based on Pantone colors [56], enabling clearer differentiation of structures.

In Figure 19, high-dimensional structural information is visualized using these fluorescent colors. Conversely, to represent and visualize molecules or their finer structures at lower-dimensional scales, a color palette based on lower chromatic colors can be used to further emphasize the differences in information dimensions. The sample of the color palettes is shown in Figure 20.

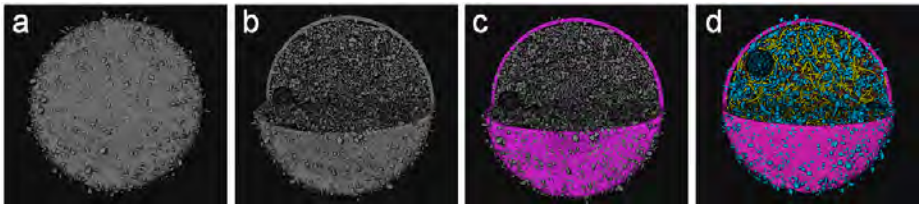


Figure 19: Examples of the use of fluorescent color palettes to show the basic structure of Micoplasma model based on information hierar- chy.

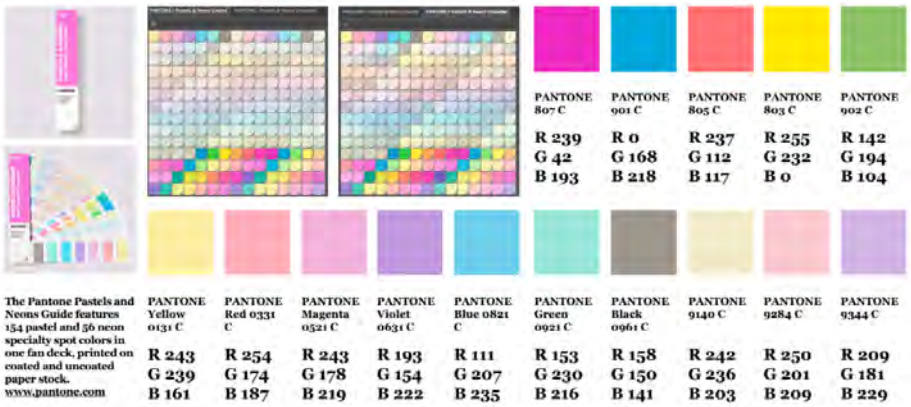


Figure 20: Samples of two-color palettes.

We set six main rules of color ontology to use the color palettes as following:

1. Create a tree-like structure to define the components and their connections.

2. If necessary, define the possible colors for each level to highlight the individual components.

3. Start with fluorescent colors and decline towards pastel colors for the subsequent levels.

4. Maintain consistency in the choice of color concerning the colors of the parent container.

5. If a color is used for one component, this color cannot be reused for another (dedicated color).

6. Colors to show the environment.

For the 6th rule, we try to express the pH level. We overlaid a colored layer to represent the cell pH level; for alkaline pH, a blue layer; for neutral pH, a yellow one; and for acid pH, a red one. Figure 21 represents a variation of Figure 19(b).

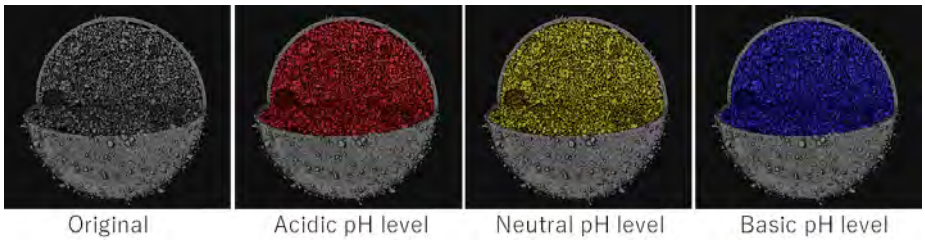


Figure 21: Examples of the pH level of Figure 19(b)

4. Automatic Coloring of Microscopy Images Using Information Entropy

We have developed a prototype program that automatically colors objects in microscopy images by leveraging information entropy. The program is written entirely in HTML and runs smoothly in modern browsers. When a user enters a textual description of the image into a form, the program calculates the information entropy as defined in Equation 1:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) \tag{1}$$

where $(p(x_i))$ represents the probability of each unique word in the description and n is the number of unique words. For example, in Figure 22, the description contains two keywords, 'liposome' and 'microtubule', each with a probability of $1/2$. Substituting into Equation 2:

$$H(X) = -\sum_{i=1}^{2} \frac{1}{2} \log_2 \frac{1}{2} = -\left(\frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot (-1)\right) = 1 \tag{2}$$

This entropy value of 1 bit results in the selection of two colors—grayscale and a single hue—which are applied to the fluorescence micrograph to distinguish the liposome and microtubule structures.

The resulting entropy value determines the number of colors to be applied. A detailed, expert-level description produces higher entropy and thus more colors, while a concise description for non-experts yields fewer colors. The program then selects the required number of hues from a predefined color palette and applies them to the image. Figure 22 demonstrates the program with a fluorescence micrograph of a liposome containing bundled microtubules. Because the description contains only two keywords 'liposome' and 'microtubule', the

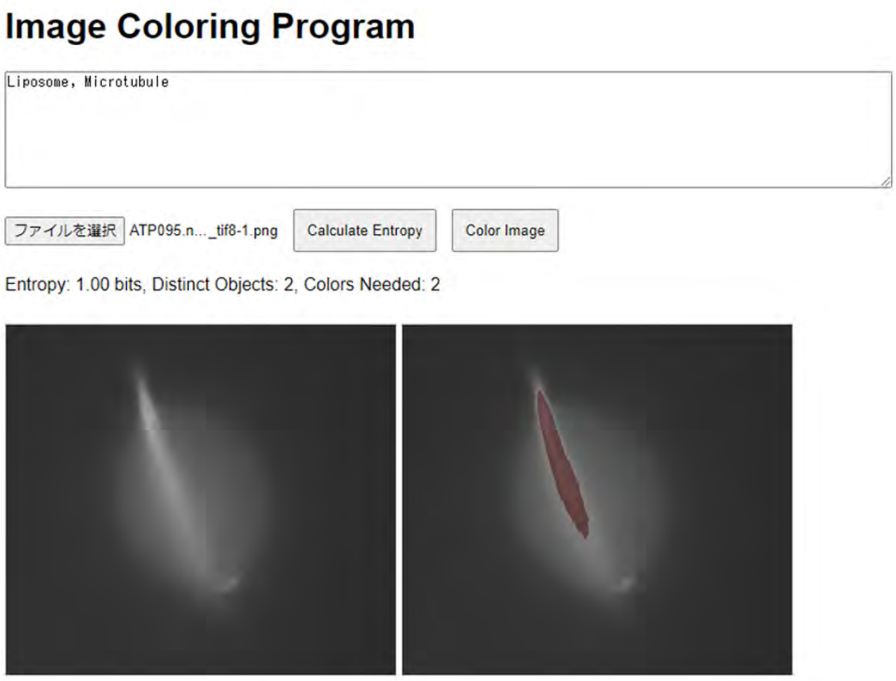


Figure 22: An example of coloring using an image coloring program.

calculated entropy selects two colors, grayscale plus a single hue, allowing each structure to be clearly distinguished.

5. Summary

The Colors Group explored the application of colors in visualizing complex cellular architectures, proposing colors as an ontological framework. By structuring cellular components hierarchically and using information entropy, they developed a systematic approach to color selection. Color palettes, varying from fluorescent to pastel hues, were designed to represent different levels of cellular organization, as demonstrated in the Mycoplasma model.

Rules were established to ensure consistent and distinct color use, including environmental factors like pH. A prototype program was created to automatically color microscopy images based on textual descriptions, using information entropy to determine the number of colors. This approach enhances clarity and communication in biological visualizations, balancing complexity and accessibility for diverse audiences.

Group discussion: Representing Uncertainty in a Multi-scale View of the Cell

Participants: Ludovic Autin, Marc Baaden, Barbora Kozlikova, Michael Krone, Daniele Savasta

In this discussion group, we focused on the problem of visually representing the uncertainty when displaying complex structures, spanning from the atomistic scale to the whole-cell environment. The key topics that we discussed were the following:

Uncertainty as a visual concept

Here we discussed our experience with existing visual representations of uncertainty, potentially coming as an inspiration from seemingly unrelated areas. Analogies discussed included the usage of fog of war in video games minimaps [49], dashed timelines used in historical timelines (e.g., Joseph Priestley)[47], as well as monsters and black holes on old medieval maps [64], indicating unknown or unknowable areas. Furthermore, we focused on the identification of different types of uncertainty, where we revealed the following basic sources: unknown structure or sequence, low-resolution density, and missing or placeholder data.

Figure 23: Artistic depiction of uncertainty as seen on old maps: sea monsters can represent unknown areas. Giacomo Gastaldi, "La Descriptione dela Puglia", 1567. Source: gallica.bnf.fr / Bibliothèque nationale de France

Techniques to visualize uncertainty

Here we were analysing the existing visual channels and features that are generally used for visualization of uncertainty. We identified the following:

- Visual metaphors
  - Typical representatives are blur, transparency, sketchiness, dark zones, or question mark bubbles.
  - Thick or dashed outlines or sharp geometric shapes are often used to draw attention.
  - Lighting effects (e.g., only the uncertain region emits light) can be used for highlighting the uncertain regions or objects.
  - Textures or hatching for architectural or surface features are also typically used in molecular visualization.
- Color
  - "Bright pink" (e.g., as used in the Unity game engine) was mentioned as a universal sign of denoting "something wrong" for users of that software.

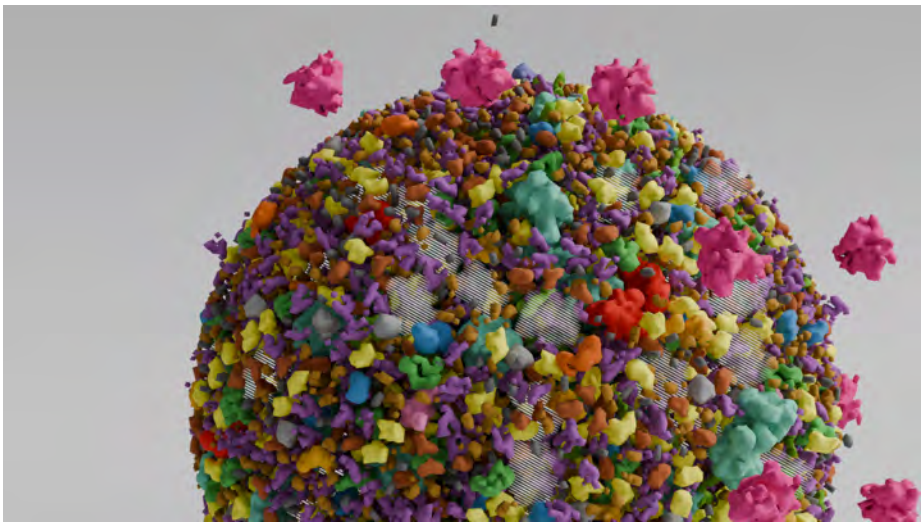


Figure 24: An early sketch of a representation adopting a combination of hatching and transparency to represent uncertainty. The example has only visual value and does not represent any real uncertain data.

  - Other commonly used and powerful options include grayscale vs. colored to highlight known vs. unknown.

- Animation
  - Motion blur, blinking, or dynamic fuzziness can be successfully used to draw attention, but it has to be used carefully to avoid possible distraction.

Context and abstraction

Here we discussed mainly the variables and context that influence the final selected representation of uncertainty. We concluded that the representation mainly depends on:

- Audience – who is the target group and how it influences the design decisions (experts vs. students).
- Scale – the range of scales in data and its resolution that influences the choice of proper representation (atomic vs. cellular vs. tissue).
- Abstraction level – determines the level of abstract or realistic appearance of the visual representation (cartoon vs. photorealistic).

It is important to note here that uncertainty at different levels might require different styles even for the same object.

Figure 25: Sketch of our proposed categorization of types of visual representations typically used for conveying uncertainty, based on several criteria. These are the scale in the data, and the type and the level of uncertainty.

Exemplary cases

In the following, we were trying to identify specific examples where we can pinpoint clear cases of uncertainty that needs to be visually communicated. The examples discussed were:

- Mycoplasma "foot"– model where only the low-resolution density is available, and it is questionable how to represent this absence of high-resolution data in such a crowded model.
- 3D chromatin models – there are multiple reconstructions from Hi-C data, which were constructed using various existing algorithms. The question is how to display the uncertainty of the correctness of the individual algorithms and the resulting models, and how to show uncertainty when the models differ significantly. A challenging aspect here is also how to communicate the whole model is uncertain, but still we want to show it in as much detail as possible to still enable visual inspection and exploration.

Guidelines

As the discussion led to several very interesting points and conclusions, we finally discussed how to communicate this information to a broader audience. We concluded that the most feasible way would be to compile a set of exemplary cases that could help the designers of visualizations to decide about the most feasible options for displaying uncertainty in molecular and cellular scenes. We decided to categorize the possible approaches according to several criteria (e.g., source of uncertainty, scale of the data) and then to divide the existing approaches to visualizing uncertainty according to these criteria (see Figure 25). The goal is not to create ultimate guidelines, as this is a highly complex or

maybe even impossible goal to achieve. Instead, our aim is to provide designers with a catalog of visualization options. As creating such a catalog cannot be reached within the meeting (due to its broadness and complexity), we agreed to continue on this topic and prototype different visual strategies on real datasets on different scales. The resulting representations will then form the desired catalogue, which can be organized, for example, as a decision tree.

Group discussion: Visual strategies for representing scale

Participants: Monica Zoppè, Paola Perin, Vera Williams, Beata Mierzwa, Ryo Mizuta

The molecular world is a vast, unknown territory for most people, and the intricate systems that make up the foundation of life don't elicit a coherent intuitive model in our mind. Although most people are probably able to visualize the DNA double helix, the organization of DNA strands in the cell nucleus does not come to mind intuitively. One issue with visualizing the molecular world, whether in an illustration or in our mind's eye, is depicting accurate relative scale.

Even for experts who are exposed to molecular-level visualizations on a daily basis, comparing the size of a strand of DNA to the cell nucleus is not as immediate as comparing the size of a car to a mountain. Our real-world experience with cars translates correctly whether we see a car depicted in close-up or from afar as it drives on the Fukuyama mountain. How, then, do we facilitate this intuitive size comparison for microscopic objects? A possible solution to this issue was proposed with the Ten Million Times (TMT) scale [29], wherein the cell and its components, all the way down to the atoms that make them up, are expanded ten million times, in an effort to enhance audience relatability by bringing most objects within the mm to km size range. In TMT, the DNA double helix would be a rope with a diameter of 2 cm, whereas the cell nucleus would be an ellipsoid with a diameter of 60 m.

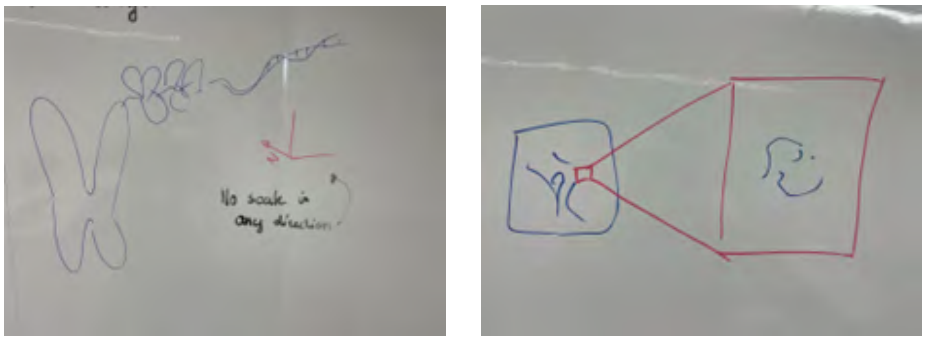
This working group started discussing around the notion of a "complete" 3D representation of a cell on the TMT scale. The aim of this up-scaled 3D model was to allow users to "explore" the cellular environment and appreciate the various components with correct relative scales and distances. The discussion points summarised below were held primarily within this specific context. Other possible representations (e.g. 2D images such as renderings, microscopy images, etc) were also included in the discussion of potential solutions later on, as well as the issue of accuracy in scale representations.

Conventional approaches for conveying scale

As a starting point, the group brainstormed existing visualisation strategies for conveying a sense of relative scaling across different cellular components, together with their respective shortcomings and challenges. These were:

- "Hierarchical" schematics, as often used for representing the structure of chromatin (see Fig. 26(a)).
- Nested zoom (magnification box) illustrations (see Fig. 26(b))
- View through a lens (combined with strong depth of focus/local occlusion effect to convey depth)
- Use of scale bars

Several challenges and shortcomings were identified in existing representation methods:



(a) (b)

Figure 26: Examples of conventional representations of different scales using (a) hierarchical representation, and (b) magnification boxes.

- Lack of biological accuracy (hierarchical schematics): Displaying multiple levels of hierarchical structures at the same time often detracts from accuracy in depicting a process or structure. For example, in the case of the chromatin structure image used in most textbooks, there are no physiological conditions in which one would find condensed, metaphase-like chromosomes unwinding as depicted.
- Inaccurate scale representation (hierarchical schematics): Depicting structure details across multiple scales within a single illustration does not always provide information on relative dimensions, especially when the structure is depicted in a 3D style (where perceived closeness to the viewer of each detail would bias its size estimation).
- Scaling continuity (hierarchical schematics, magnification box): For discontinuous representations such as the magnification box approach, it can be hard for non-technical viewers to maintain their appreciation of scale across the different images. especially when scale jumps between overview and details are very large. For example, in the typical display of viral invasion of a cell, the whole cell overview is often depicted as big as the viral particles in subsequent image panels.
- Intuitiveness vs accuracy (scale bars): While not a concern for non-technical viewers, size exact measures given by scale bars were raised as essential in scientific communication, but potentially detracting from non-technical audiences gaining an "intuition" for relative sizes.

Problem statements

Based on the shortcomings outlined above for existing representation methods, the working group narrowed their discussions in an attempt to tackle the following questions:

- How can distance information be conveyed in any view, on any scale?
- When representations transition across orders of magnitudes in size (both in static images and dynamic 3D environments), how can one represent

distances in a clear way that survives the transition? I.e. in a way where viewers can retain their appreciation of distance after changes in scale?

Possible solutions

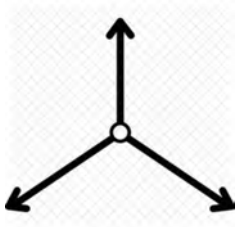


Figure 27: 3D axis

- Including 3D-scale, as in Fig. 27 (reduce closeness bias, feasible in schematizations, not in 3D)
- Camera zooming (not quantitative if zooming speed is not measurable, e.g. when too fast)
- Depth cueing, i.e. *photographic lens effects; human eye vs. camera* (only the object of interest would be in focus, whereas the background and foreground would be progressively blurred)
  - The "view" should be accurate to whatever a camera or human eye can record. All effects such as depth of field, should be realistically achievable. We therefore need an acceptable range of viewing parameters such as focal length (in mm), f-stop, etc.
  - For an interactive 3D environment (e.g. VR), besides out-of focus blurring, one could think about additional viewing modes such as introducing a magnifying lens that would allow the user to see cellular objects to greater resolution.
- Light or saturation fading with distance *i.e. atmospheric perspective*
  - What influences fading distance? Water? If so, how do we define fading rate, and should the attenuation length be scaled TMT or remain fixed in viewer scale?
  - If we are considering water, how can we represent it? In TMT viewer scaling, water molecules would be around 1 mm. This would make it too crowded: if all water molecules were represented, this would make it impossible to see anything else. Could create a gradient view of visibility and/or selectively show water molecules where needed (Marc's input - e.g. water around protein surface or within active sites of molecules?)
  - Could we show lighting as we often see underwater to nuance water whenever we are not showing it to molecular detail?

– Toggling: Starting from the toggle idea for water molecules, we suggested we can toggle other properties as well, such as level of detail (atomic, 3A, coarse surface, all from cryoEM). Discussion ensued about what is real in electron density data.
– Addition of a mini-map (with fading to represent objects that are too far to see, the overall picture may be lost, and a minimap could allow the viewer to better navigate the environment)

Level of understanding, level of detail, and learning

A test with a non-specialist "exploring" a 3D cell environment (Teresa) was conducted using the mycoplasma rendering of MolStar [15]. The results of this test clearly showed that an unguided interaction with a "complete" cellular model (i.e. displaying as many components as possible) was too overwhelming and did not produce the "exploratory" learning journey anticipated but instead yielded confusion in the user. Two conclusions were made:

- A 3D environment that assumes nothing about the audience and intended purpose is not possible. Decisions on what to show and what not to show must be made in the interest of clarity and constructive learning.
- We must acknowledge the possibility that different users may approach visualisation in different ways, e.g. starting with an exploring phase (as highlighted by the MolStar test), or preferring a more guided approach.

These points highlight that for any meaningful further work, user testing is necessary, to define the preferred visualisation style (including representation of scale).

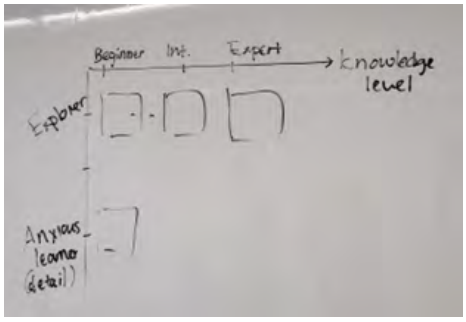


Figure 28: An illustration of possible categories/styles of approach to an interactive cell environment.

Possible experiments

Designing a scale object: As a feasible starting point, the working group focused on exploring possible scale objects, with the aim of testing their usability in a future user test. This scale object was not intended as a strict quantitative size reference, but rather as a visual anchor when viewing images on different scales. A cat (as originally used in the TMT world) was selected as the prototype scale object for two reasons.

1. It is important for the scale object to be visually familiar to any viewer. An intuitive sense of the cat size, and the relative scaling of specific parts such as eyes, ears, whiskers, are necessary for visually grounding the viewer across multiple scales.

2. The scale object should have an appropriate scale in range with the TMT representation.

Initial testing: An implementation of a prototype scale cat is shown in Fig. 29, modified from Ref. [27]. Testing was conducted using existing repository of visual representations, previously set up by the group (survey), comprising microscopy images, rendered 3D visuals, etc.

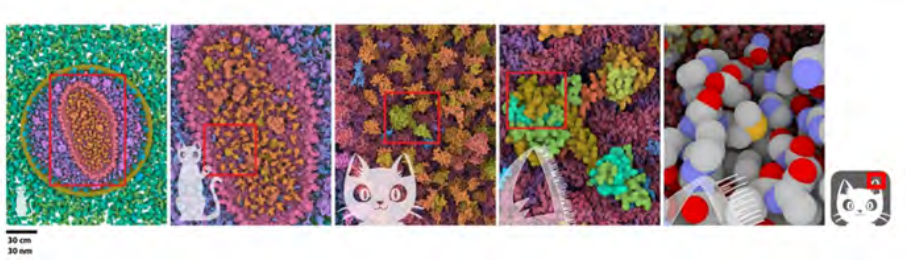


Figure 29: Prototype implementation of a scale-cat across a multi-scale 2D representation of a cell and its components.

Questions from initial testing:

- What visual style should the scale cat be drawn with? A somewhat stylised approach may be suitable, but the illustration should maintain realistic proportions
- Cat works for images with field of view that is up to 1 μm. What to do for larger structures? Perhaps e.g. 3 representative animals could be used for different scales, where at least one of the scale animals is seen in its entirety at any given magnification.
- for 3D images that include the third dimension (or for 3D environments, such as navigable or interactive representations), the scale cat might need to be 3D as well.
- Although possible, using the scale cat in 2D images that include a perspective view is more challenging, as it would be necessary to match the image's perspective (as it requires quite a bit of artistic skill to draw a cat in the same perspective used in creating the molecular scene).
- If we want to use different animals to illustrate different scales, we may need culturally "neutral" animal icons

Next steps and future work

Given the diversity in the ways people interact with visual information (particularly those that represent complexity at different magnification), this discussion group concluded that immediate user testing with the scale cat was needed. The

prototype illustrated in Fig. 29 represents one instance of its use, in which it is overlayed with an existing 2D image. Its application to other forms of images (raw microscopy images, 3D rendered illustrations) should be tested to yield corresponding prototypes. The resulting repertoire of scale representations will be used to survey and study the visual interactions/preferences of the wider audience.

Summary of new findings

The workshop was participated by 17 scholars of different disciplines and backgrounds. This made it possible to develop engaging and fruitful discussions, which spanned from the psychological aspects of color perception to the details of producing interactive visualizations.

During the first day, after the initial presentation of each member of the group and an introduction by the organizer MZ, we witnessed some of the problems relative to using colors and other visual elements in preparing images that report cellular objects, at various levels of details.

The omission of some details, to simplify the image and not overwhelm viewers, the use of metaphors relating to our everyday life, the introduction of artistic elements that may engage with large audiences, were the subjects of the afternoon session.

During the second day of presentations, in the morning we delved deeper in the theme of Color, from definitions according to different standards, issues of encoding and reproduction on different substrates (from cloth to video screen), to the psychological implications of using colors, which, being often culturally acquired, may differ among different peoples. Contrast, attractiveness, readability and other aspects of the use of colors in various situations, from the arts to the toilet indicators, were all themes that color experts were capable of illustrating to the crowd of scientists.

The afternoon was dedicated to talks and an engaging session with many virtual, mixed or augmented reality (VR, MR, AR) experiences. We witnessed what I (MZ) felt like a major leap forward in the use of this technology. We were able to navigate in the cytoplasm of some cells, to manipulate DNA by separating and sometimes re-annealing two strands of the filament, to build viral capsids starting from monomers, and to dance with molecules in the room.

Considering the wide span of disciplines involved, and the broad implications of discussing the theme of the workshop, it may not surprise that the discussions were open and sometimes disordered. Despite the vastness and complexity of the themes, some issues emerged as transversal to all discussions, including:

- SCALE
  The world of cells spans several orders of dimensions: from the subnanometer size of single water molecules or of ions (10^{-10} m), to the dimension of an entire cell, that can reach tens of micrometers (10^{-5} m), corresponding, in the TMT scale (see group discussion), to a range going from mm (10^{-3} m) to tans of km (10^{4} m).

- UNCERTAINTY
  Among the vast amount of knowledge obtained thanks to the many disciplines informing cellular and structural biology, there are still unknowns: both known, such as the details of the packing of DNA in nuclei or the structure of the proteoglycans that typically surround the surface of many cells, and unknown, which we cannot define, and that could be related to specific cellular conditions, associated with viral infection, or related to particular kind of cells found in 'strange' organisms.

- CROWDEDNESS
  Cells are typically composed of water for 70-80% of their total weight. If

all water molecules were to be depicted, no image would be readable at all. This consideration led to a general agreement that water may only be explicitly included in images only when necessary, for example when it is part of an enzymatic reaction. Even with this solution, however, the sheer amount of other objects makes the place a very crowded one, such that any view could not cover a depth of more than few tens of nanometers. One possible solution could be to agree to represent only one tenth of the objects in the cytoplasm, keeping the proportions scientifically accepted.

- ACTIVITY
  All living systems are alive and constantly active, in a dynamic equilibrium that allows life to proceed and keep homeostasis. Such activity is typically depicted in printed images making use of signs and symbols (arrows, equations ...). In a purely visual descriptive context, however, activity can only be reported as an animated video. The group convened to leave this topic, in itself extremely important, for future discussions.

- INTENDED AUDIENCE
  A major topic in all discussions, and in general about producing visual representations, is how to adapt the message to its intended audience so that the information reaches them without distortion of the underlying science. In line with the proposal that we proceed by trying to build a mental description of life in the cell (like we have of the real world surrounding us), we agreed that: 1. any representation is necessarily biased by who actually makes it, just like any person who takes a photo decides the lens, the framing, the exposure, the background and so on; and 2. the adaptation of the 'ground truth' (despite imaginary) to the audience can be prepared only after (and if) we reach a common vision of the 'ground truth'. Therefore we considered that also this issue should be discussed in future meetings.

- STORYTELLING in COMMUNICATION
  Another aspect frequently proposed by several participants is how to engage the audience, or how to let viewers relate the scientific message to direct experience. Several examples were brought to our attention, and most participants agreed that the 'trick' can be used in ways that are both artistically pleasing and effective in transmitting the message, but is also prone to misuse leading to 'bad solutions'. This fine line between engaging the viewer and distorting the message seems to be particularly difficult to set by employing IA tools, which are increasingly available and user-friendly but carry the risk of biasing the storytelling, leading to shallower learning. This topic was considered less urgent than others and, considering the limited time, left for future research.

- COLOR. CONSISTENCY, PERCEPTION, ATTRACTIVENESS
  Color was, as expected, the most discussed item during the entire 5 days of the workshop. Not surprisingly, the group dedicated to Colors was the most participated: please refer to the dedicated chapter.

Identified issues and future directions

The topic of the workshop was very broad, with the aim of progressing towards he delineation of a roadmap that could lead to solve the issue of how to harmonize the use of colors and signs in the representations of structural and cellular biology.

Accordingly, after the general dissection of the different facets explored during the first workshop in 2024, we concentrated on a subset of problems.

In particular, as reported in the group discussions, we concentrated on the use of colors with some examples, on how to represent uncertainties, and on how to convey the vast range of scales that we can encounter when visiting cells, their interior and the elements inside (see reports of groups).

To some extent, the proposed solutions can already be implemented, and the best way to proceed seems to be a test of implementation and reception by viewers. These features could be tested and assessed independently of the final outcome, and indeed some preliminary tests could inform the progression of the general work. To this aim, which implies an organized project, we might be able to submit a project proposal to funding bodies. It is common experience among the participants that getting funds for this kind of research, that spans different disciplines and implies participation of scholars from different Countries, is hard to obtain. Indeed, very few calls for applications that are interdisciplinary and international in nature exist. Nevertheless, we will consider whichever possibility arises and pursue any open avenues for furthering our research.

While starting with some initial testing, it is clear that there is still a long way to go, and we have discussed the possibility of convening another meeting to refine the proposed solutions, and/or to prepare a project to be submitted to grant agencies.

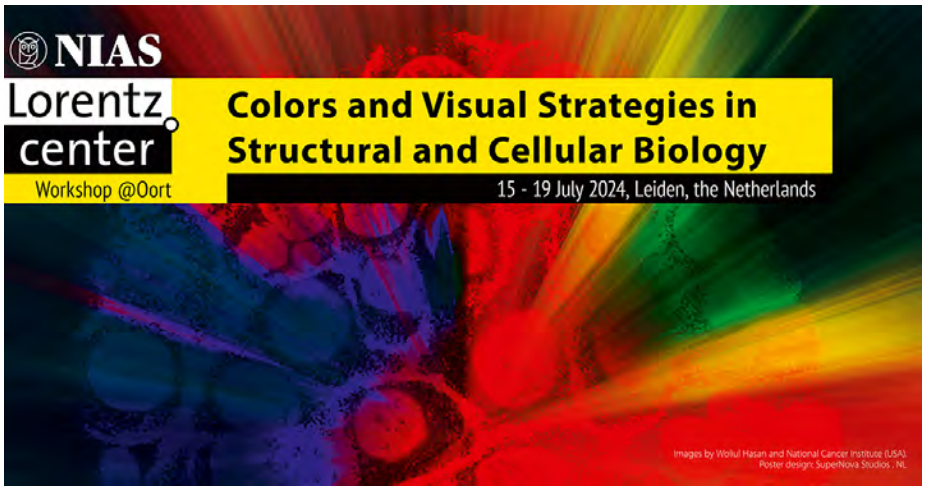
Regarding a new meeting, several participants mentioned possible venues and institutions that could fund and/or host the next workshop:

- Fondation des Treilles
- Centre Européen de Calcul Atomique et Moléculaire
- Lake Como School of Advanced Studies
- Ettore Majorana Foundation
- The Nesin Village
- European Cooperation in Science and Technology

We will examine the various possibilities and submit applications to those that seem most suited for our purpose.

The thorough discussions, informed by the direct experience of many participants, also produced some ideas that can already be shared with the wider community, and we are planning the preparation of one or more ‘white papers’, or opinion articles, in which we present some of the ideas discussed, with the intent of raising awareness and engaging many professionals in the development of a common basis for depicting structural and cellular views.

A Report of the previous Lorentz workshop



A.1 List of Participants

- Franco Achilli, Istituto Europeo di Design - Milano, Italy
- Maria Teresa Artese, Institute for Applied Mathematics and Information Technologies - CNR, Italy
- ★ Marc Baaden, IBPC - CNRS, France
- Sander Ewen, Leo Kanner College, Leiden, The Neteherlands
- Matteo Farinella, Columbia University, Zuckerman Institute, USA
- Erna Fiorentini, Institute History of Art and Architecture, KIT - Karlsruhe Institute of Technology, Germany
- Brady Johnston, The University of Western Australia, Perth, Australia
- Kèvin Knoops, Maastricht University, The Netherlands
- ★ Valentina Manchia, Politecnico di Milano; University of Bologna, Italy
- Paola Perin, Department of Brain and Behaviour Science, University of Pavia, Italy
- James Procter, University of Dundee, UK
- Micol Riva, Politecnico di Milano, Italy
- Bjorn Sommer, Royal College of Art, London, UK
- ★ Vera Williams, LIACS Computer Science Inst. Leiden University, The Netherlands
- Yoshie Kiritani, Design Research Institute, Chiba University, Japan
- ★ Monica Zoppè, Institute of BioPhysics - CNR, Milan, Italy

---

★ Organizers

A.2 Introduction

Monica Zoppè
monica.zoppe@cnr.it

Similarly to the Shonan workshop, the Lorentz workshop, held in Leiden, NL, in July 2024, organized by Monica Zoppè, Marc Baaden, Valentina Manchia, Friedrich Förster and Vera Williams, lasted 5 days and saw the participation of less than 20 scholars. The participants were coming from a slightly more varied range of disciplines, which gave to the meeting a very strong interdisciplinary character. The topic of discussion was exactly the same, but, being the first time that we convened to face all aspects of visualization for cellular and structural data, we had to dedicate some time in the definition of the problem, and in the development of a common vocabulary that could permit a fruitful exchange among all of us. The major part of the first days was set to present each other our expertises and 'tools of the trade', which are very different for structural biologists *vs.* semiotic experts, *vs.* designers and communication designers *vs.* teachers etc. The details of this effort are reported in Chapter A.3 dedicated to methodology and mutual presentations and understanding.

Another foundational issue (reported in Chapter A.5) was to reach a clear definition of the problem: **What** is it that we want/need to represent? We discussed the nature of objects, which are typically proteins, nucleic acids, lipids and carbohydrates, all combined in mixed forms and at various levels of size; we discussed issues related to dimensions, which span several orders of magnitude, yet are all too small to be perceived intuitively; also, considering the active nature of biological objects, we discussed about activity and forces, which are often totally extraneous to our experience, such as pH, osmotic pressure, RedOx potentials, electric charges and more.

We rediscovered the need of a full 'catalogue', which would ideally include everything, and we even started to prepare it, before realizing that several attempts of organizing structural knowledge are already active for the classification of proteins, of genes, of cells and more, although they not always are connected to one another, and we may not be able to exploit them for our purpose. It might also be humanly impossible to actually build a catalogue that includes everything, especially because of the variability of biological entities.

A second step in the discussion (reported in Chapter A.6) aimed at defining **how** we could represent the cellular and structural information: which tools (from printed paper to VR), which color could be used (too many items with a limited range of discernible colors), which other channel of communication could be used (beyond colors: texture, light, sound...).

The presence of teachers and communicators allowed the group to discuss the question of **for whom** the representations would be prepared: clearly, preparing illustrations to be printed in elementary school books, in scientific publications or in comics books aimed at young people, implies, besides the difference in content, also a different visual language.

Other aspects proposed were related to the representation of **unknowns**, and of the different levels of uncertainty, a theme re-proposed in the second meeting at Shonan. We also briefly debated the various possibilities of showing **activities**: minute vibrations, active transport, or major cellular changes, such as during cell division.

As initially foreseen, the subject of the meeting was way too vast to be fully examined during a 5-day workshop, and we did not reach any meaningful conclusion. However, we did set some firm points that served as a basis for the following meeting in 2025. We agreed to delve deeper in the discussion, as we did in Shonan and present in this report, and we imagined the possibility for obtaining some basic information onto which to build further insights. To this aim we proposed to set up a survey that could enable us to discriminate what works best, at least regarding colors, in some exemplary representations. A first draft for such survey was set up, and its content served as examples during the Shonan meeting. However, the Survey has not been distributed widely, and it will need more items that users can evaluate in order to be meaningful in a statistical study. We expect that the results will provide indications on how to manage colors in cellular representations, once it will be completed.

On a personal level, I can safely state that the ample discussions during the days at Lorentz were very open, friendly and enjoyable, but above all made it clear to me that the way to proceed towards reaching an agreed set of guidelines on how to use colors and visual signs in the representation of structural and cellular biology, need to proceed step by step. This awareness led me to sketch the proposal presented on page 29.

A.3 A successful interdisciplinary method of investigation

Erna Fiorentini, Valentina Manchia, Paola Perin*
* Corresponding author: paola.perin@unipv.it

The discussion on visualization in structural and cellular biology would not have been possible without a better understanding of how scientific images work and how color can convey information and meaning. For this reason, the workshop was designed to cover a wide range of scientific fields and expertise, from molecular and structural biology to molecular visualization, from image theory to semiotics and visual culture studies, from color theory to color science, from science education, communication and dissemination to design, illustration, and comics. From their perspectives and with a few visual examples and materials to share and comment on, scientists and experts presented and discussed a wide range of topics shown in this report.

In particular, the scientific talks were mainly aimed at giving nontechnical explanations of the structure of scientific images and the solutions found for communicating complex information. In more detail, three issues were discussed:

1. how to communicate the scale at which a particular image detail or pattern is visible: possible solutions include inferring scale from an object of reference, as in Zoppè, 2017 [29], or showing progressive magnification steps, such as in the classical video "Powers of ten" [53];

2. how to differentiate images depicting raw data from those showing derivative models: here solutions appeared to be more heterogeneous, but one effective strategy employed differentiating color saturation and/or transparency between raw data and models, in particular for Electron microscopy or Electron tomography, where data are intrinsically colorless; data can be shown in black and white, and models in colors [10];

3. which common visual conventions and metaphors are used in visualizing abstract data: commonly used ones include field lines as arrows, protein display styles as ribbons or ball-and-stick, pH and redox potential as color gradients [5].

Lectures on color theory then explored in detail the mechanisms of color perception, including the issue of color blindness, the visual standards and color space and palettes currently used in color coding, and the technical application of the color spectrum.

Theoretical reflections on the typology and strategies of visualization have expanded the scope of reflection in two ways. On the one hand, from the perspective of image theory, an attempt was made to define the differences between types of visualization and identify the parameters that would serve as instruments for this differentiation (as proposed by Erna Fiorentini [37]).

On the other hand, an attempt has been made to define the informational and expressive role of color by evaluating its use as a specific communication tool in the context of the single scientific image, starting from the fact that every scientific image is the result of a *multi-level translation process*, from the recording of laboratory data to its visualization (as proposed in Manchia, 2020 [40],

and Manchia, 2022 [13]). Specifically, there are several ways in which color, as a *visual variable* (Bertin, 1967 [57]), can be used for communicative purposes. In atomic models, for example, color has an informative value, labeling the different elements. In an electron tomography, on the other hand, which has no intrinsic color, an applied color may act by signaling a specific area of the image. It may function as a focusing element for the reader, a tool for "channeling the meaning" (Bastide, 1990 [2]).

In both cases, we can speak of a local color meaning, depending on the context of and the relationships between the represented elements. The choice of colors to be used can, therefore, depend on visualization strategies based on highlighting contrasts and relationships between elements. The effectiveness of science communication was also the theme of other presentations, which focused on the visual and expressive aspects of scientific images, as well as their communication and dissemination, especially in contexts such as 3D animation for the general public and education. In particular, the importance of extracting the fundamental data to be communicated in visual form in image processing and molecular modeling, in order to obtain clearer and more direct information for the viewer, was frequently emphasized. *Mediation*, *abstraction*, and *simplification* have emerged as keywords. All these considerations can be valuable to start thinking about how to work more consistently on images, as well as how to achieve greater consistency between sets of images (within a publication, but also in an animation). Given the large number of disciplinary orientations and languages, it proved challenging to develop a common transdisciplinary language.

Short lectures on single topics helped us to understand each other's concepts and language. After the first round of talks introducing the specific fields of expertise of the speakers, several talks (some in the form of slide presentation, some more informal such as blackboard or coffee machine discussions) took place, explaining e.g. the concepts of pH and redox potential for non-scientists, the principle of printing design for non-designers, or showing topical images or videos as a starting point for discussion.

Indeed, each disciplinary language exhibited its own sectoral lexicon, and cross-disciplinary comparison was also useful in reflecting on the different meanings that terms attested in one research tradition can have in other scenarios.

For example, a term like "construction", which in the perspective of image theory and semiotics alludes to the mediated and situated nature of scientific images, in the microscopy field technical terminology alludes to building an image that derives from data in an indirect way (e.g. by averaging several incomplete instances, as in cryoEM), as opposed to "reconstruction," which indicates extracting an image directly from data (e.g. obtaining the 3D volume of a cell by delineating its 2D profiles throughout its depth). For this, see Fiorentini [37], pages 35-36, Table 3.

Other concepts needed further specification for communication between scientists and visual people, such as the concepts of "artifact" and "mesoscale". The term "artifact", in the scientific world, indicates an unwanted product of the experimental procedure that affects the meaning of the experiment. For example, tissue shrinking due to dehydration is an artifact. In the visual culture world, on the contrary, an artifact is a valuable product of human skills and artistic inventiveness. The concept of mesoscale is instead more ambiguous within the sciences [19], since it indicates an "intermediate" scale between microscopic and

Figure 30: We all take different aspects from the same reality. The cat in the middle (which we may think is the same for all) is seen in a different way by a veterinarian, a child, a robot, a mouse, and an artist (image credit Marzia Ravasio).

macroscopic, but different scientific disciplines use it on different size ranges (nanometers to microns for molecular visualization, microns to centimeters for neurosciences, and even more variable ranges in material science).

Regardless of size, in the mesoscale approach understanding small structures is essential to explain the functioning of the macroscopic system (for example, a small splinter in the foot can affect a runner's ability to run a marathon, so the spatial range necessary to understand what happened during the run needs to include the millimeter-size splinter and also the 42 km marathon path). Mesoscale visualization poses the challenge of balancing the need for giving the essential microscopic detail, the danger of information overload due to unnecessary details, and the necessity of orienting within the different scale ranges at play. Since modern biology is increasingly focused on mesoscale systems, visualizing the mesoscale is a central problem in modern biology.

For the sake of effective communication, real translation between different languages took place on the fly, creating a kind of common "service language" that clarified the concepts involved. For example, rather than speaking directly of reference *objects* [36], as those items necessary to mark the informational space of the scientific image, the space occupied by the object, it was preferred to speak more simply of *reference context* or *informational framework*. In this respect as well, the workshop has proven how interdisciplinarity can be effectively used as a method rather than merely as an intention or state of mind. As theorized by Gilles Deleuze, disciplines can in fact be seen as the sprouts of a rhizome, a common root that must be characterized in order to achieve mutual understanding. In our work in Leiden, a successful cross-pollination has taken place and new seeds have been planted for an advanced interrelation between disciplines. Further developing this kind of workshop seems to be the way to go if we want to achieve an effective interdisciplinary communicative language.

A.4 Understanding Color: Coding and Perceptual & Cognitive Aspects

Maria Teresa Artese*, Yoshie Kiritani*, Bjorn Sommer
* Corresponding authors: mariateresa.artese@cnr.it; kiritani@faculty.chiba-u.jp

Colours convey meaning in very different ways. For this reason, we can make a preliminary distinction. We could define code-based colour meaning as meaning that is based on a shared, fixed and conventional code, such as that used for traffic lights or basic colour conventions in chemistry. Alternatively, we could define colour-local meaning as the specific value that colours can have in a given image, according to their function (e.g. marking different areas) or their relationship to other colours within the same image. In this section, we will focus particularly on code-based colour meaning.

A.4.1 What is color?

Color is said to be a psychophysical phenomenon: the stimulus reaching the eye is the product of spectral reflectance and the spectrum of light falling on the object. A lot of experiments and color space definitions reflecting human eye behavior have been approached, starting from the first studies of 1929, by which researchers tried to model a color space in which the 3 coordinates can match human perception enabling also to compute distances between colors. The characteristics of a color space model to be useful for color coding and palette selections are:

1. to be perceptually uniform, so that an euclidean distance can be used to compute color distances and distance can be a good measure for judging whether two colors are similar or not
2. to be device independent, so that color coordinates can be reproduced with high fidelity
3. to have the three axes that match those of the human visual system, which are hue, chroma and lightness:
    - Hue is the main property of a color, used to distinguish one color from another, it involves primary colors (red, yellow, blue), secondary colors (green, orange, purple) and all degrees used to describe the others (reddish yellow, bluish green and so on).
    - Chroma is the colorfulness of the perceived color, related to chromatic intensity.
    - Luminance, or lightness, is the measure of the luminous intensity per unit area of the light reflected from a particular area and refers to the particular shade of a color and how dark or how light that particular color is.

The color space that match all these requirements is the HCL(uv) which is a polar transformation of the CIELUV color space, to be used for electronic devices, together with the HCL(ab) obtained from CIELAB color system, useful for printing devices.

A.4.2 Color coding

Color coding is the process of mapping information onto color space.

For the purpose of color coding follow the six principles: the chosen colors should be discriminable, detectable, perceptually equal, meaningful, consistent and aesthetically pleasing. Also define the information to be coded as:

1. **quantitative**, where information is distinguished by its quantity or amount and can be ordered in numerical series, leading to the choice of a color scale. Use small hue differences, or a saturation or luminance code. These scales emphasize the relationship between different levels of the information;

2. **qualitative**, where information is distinguished by its quality or nature and groups are categorically separated, there is not perceptual ordering in the representation. Use large differences of hue as a general rule to emphasize its categorical nature;

3. **ordered categories** or names require large hue differences unless it is important to maintain the rank order of the data. In this case scales are the best choice, as for quantitative information guidelines.

Except for quantitative information, restrict the number of colors to around a half-a-dozen, when more is necessary use multidimensional coding, by combining color with some other attribute, as shape or texture. For pleasing displays use complementary colors, avoid saturated blue for fine details and saturated red and blue in conjunction. Where important color judgement are to be made, objects should be viewed against an achromatic background.

Available tools for selection of colors and color palettes are:

- the ColorSpace package to create different types of palettes based on the HCL model;

- ColorBrewer 2.0 which is able to generate and export palettes colorblind safe, print friendly and photocopy safe;

- AdobeColor to select colors according to color harmony principles.

A.4.3 The constraints on color use from the user's perception

According to the communication model, there are three components: the message sender, the receiver, and the message itself. When indicating information using colors, constraints often arise from the receiver's perspective. These constraints shape how the message is conveyed. Considering human perception and cognition, we should focus on at least three key points: the functions of colors such as visibility, color illusions, and color images—including consideration of color vision deficiencies.

1. Functions of colors

A fundamental question is: why do we use colors? Why do we apply color to objects? One answer is that colors serve specific functions or effects. Therefore, we should pay attention to three aspects: attractiveness, visibility, and distinguishability of colors.

Figure 31: Today's Tokyo subway map (Tokyo Metro, 2025)

- **Attractiveness** refers to how much colors naturally draw our attention, even when we are not actively looking for the colored objects. It is also known as noticeability. For example, when driving, we usually do not search for stop signs or road closure signs, but these must be easily noticeable because they are vital for safety. The same applies to fire extinguishers. Colors that are typically highly attractive include chromatic, highly saturated, warm, and lighter colors. Thus, red, orange, and yellow are usually very noticeable.

- **Visibility** is also a type of noticeability but applies when we are actively looking for objects like information signs. When the object is text, this is referred to as "legibility." Monochrome imaging is often used to verify visibility. Dark figures on a white background and light figures on a black background are clearly visible. To achieve high visibility, the contrast in lightness between the figure and ground should be maximized. Lightness contrast is more important than differences in chroma or hue. Relying solely on hue differences is risky because it can cause effects like the Leibmann effect, which will be explained later.

- **Distinguishability**, is the ability to differentiate between multiple elements. This is often used in graphs and colorful maps, such as subway maps. Figure 31 shows the Tokyo subway map, where distinguishability between colors may not work well. When the number of elements is small, this system works fine. In the past, Tokyo's subway colors were easy to remember and helped users navigate. However, combinations of colors and objects are usually arbitrary with no strict rules. Therefore, once a color is assigned to represent a particular object, it should be used consistently. In specialized fields, such as industrial standards, these combinations are standardized.

			L cone	M cone	S cone
Normal	Trichromatism	Normal Trichromatism	○	○	○
Anomalous	Protanopia	Trichromatism	△	○	○
		Dichromatism	None	○	○
	Deuteranopia	Trichromatism	○	△	○
		Dichromatism	○	None	○
	Tritanopia	Trichromatism	○	○	△
		Dichromatism	○	○	None
	Monochromatism	Cone monochromatism	Any of LMSs		
		Rod monochromatism	None	None	None

Figure 32: Classification of color vision deficiencies by cone type (Japan Color Research Institute, 2020)

2. Color vision deficiency

Color vision deficiency occurs due to the absence or malfunction of certain cones in the eye (see Table of figure 32). Figure 33, from simulates normal vision alongside three common deficiencies. Compared to normal vision, protanopia and deuteranopia lack sensitivity to red and green, while tritanopia, which affects blue perception, results in red and green appearing overly dominant. Since red-green deficiencies are the most common, they are of particular concern. Some vision simulators allow quick testing of these deficiencies [60].

Figure 33: Color vision deficiency simulation: Normal (upper left), protanopia (upper right), deuteranopia (lower left), and tritanopia (lower right).

Barrier-free design solutions include outlining: in Figure 34 (a) shows a fictional flood disaster TV report with red crosses indicating washed-out areas. In the deuteranope simulation (b), the crosses disappear. However, when outlined (c), they become noticeable again (d). This method improves visibility

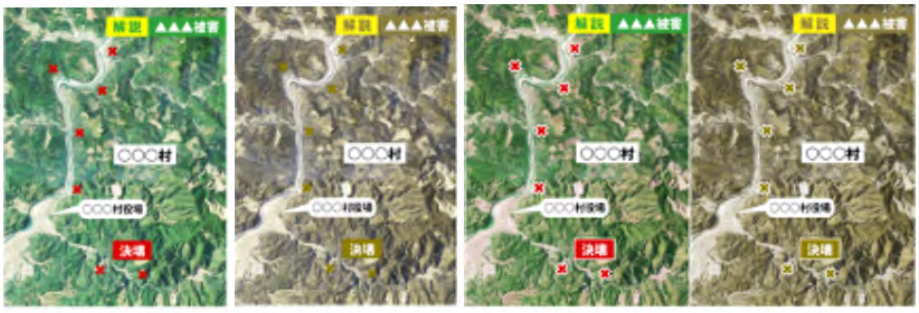


Figure 34: Examples of inadequate color representation (a), deuteranope simulation (b), improved version with outlining (c), and its simulation (d) (Japan Color Research Institute, 2020).

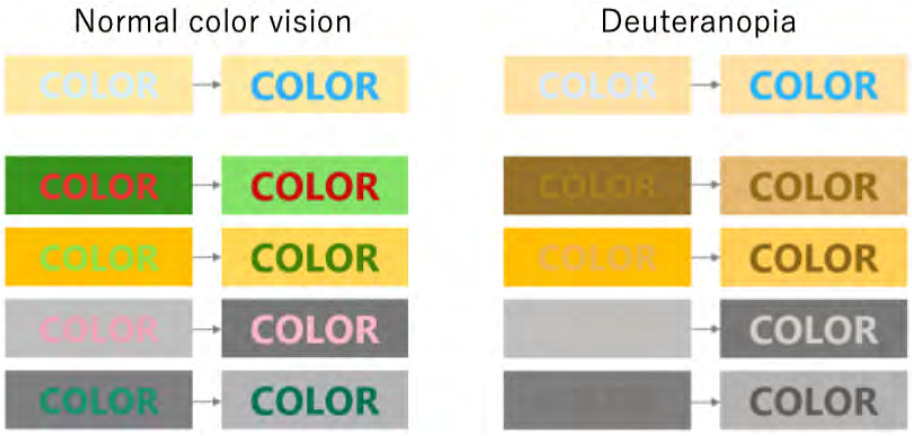


Figure 35: Increasing lightness contrast improves visibility for normal vision and deuteranopia (Japan Color Research Institute, 2020).

for both color-deficient and normal vision users. Another solution is increasing the lightness contrast between figure and ground (Figure 35), which benefits both normal and color-deficient viewers. This also improves text legibility. A final, although less universal, solution is to label colors by name on the object. For example, saying "Give me the red one, not the green one" is common, and adding color names helps color-deficient individuals identify the correct item.

3. Color illusion

Colors do not always appear the same; their appearance changes depending on the background. Color contrast occurs in hue, brightness, and chroma. For example, a yellow figure on a greenish background appears more orange, while the same yellow on an orangish background looks greener. A neutral gray figure looks darker on a white background and lighter on a black background. Similarly, a low-chroma figure on a highly chromatic background appears less saturated, while on a less chromatic background, it looks more saturated. Perceptual contrast also occurs at edges: distinct color boundaries create a sequence

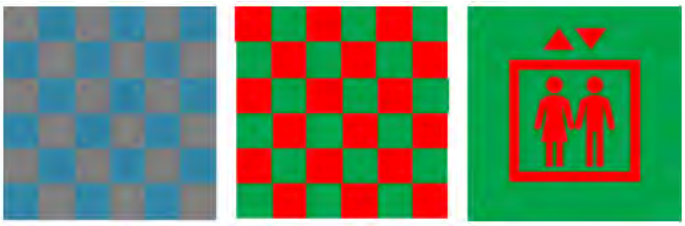


Figure 36: Examples of the Leibmann effect.

Figure 37: It is challenging to say the colors of the words (Stroop effect).

of color patches instead of a smooth gradient—this is called marginal contrast. The opposite phenomenon is **assimilation**, where thin colored lines blend with the background color, occurring in hue, brightness, and chroma.

Two important color illusions to keep in mind are the **Leibmann effect** [38] and the **Stroop effect** [26]. The Leibmann effect highlights the importance of lightness contrast for figure-ground segregation (Figure 36). For example, red and green placed together can cause flickering and eye irritation. To avoid this, the lightness difference between areas should be increased. The Stroop effect involves interference in naming colors due to conflicting verbal information. Figure 37 illustrates how reading the words is easy, but naming their color is difficult. For instance, when the word "green" is written in red, we can easily read "green" but find it harder to say "red." The meaning of words influences also perception.

4. Color images

Another cognitive aspect of color relates to color images or emotions, which operate on three levels: fundamental color emotion, color symbolism, and personal color preference. Fundamental color emotions are universal, while personal preferences are individual and uncontrollable (Figure 38 [31]).

When using colors to convey meaning, special attention should be paid to color symbolism, the second level, as meanings vary by culture and age. For example, Western cultures typically associate the sun with yellow, whereas in Japan, the sun is often represented as red. Traditional colors also carry symbolic meanings.

Concluding remarks

There are two essential points for color use:

- First, increase the lightness difference to improve visibility and accommodate color vision deficiencies.
- Second, consider the cognitive aspects of color, such as the Stroop effect and cultural color symbolism.

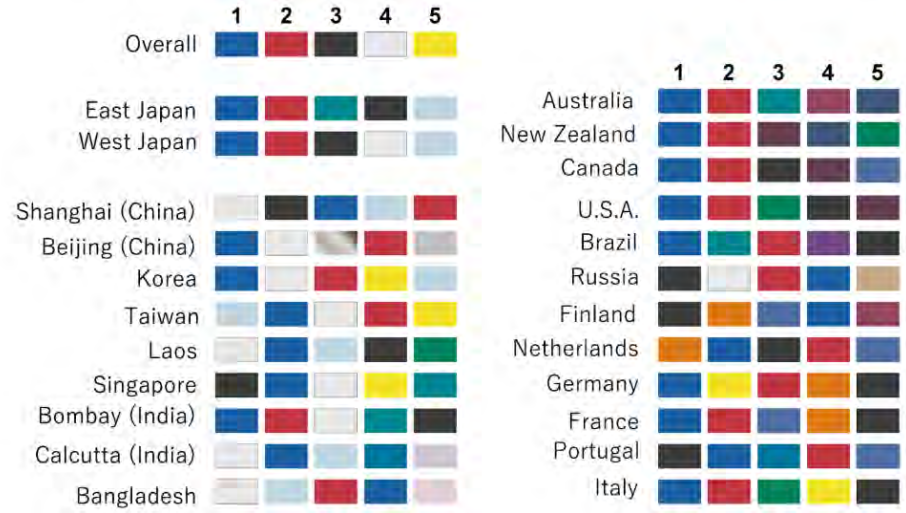


Figure 38: Color preference across different countries

A.4.4 Two application cases from Molecular Biology

A relevant problem discussed during the workshop was the use of color scales to encode different properties in molecular biology. Two application cases were discussed:

1. How to distinctively color-code the 118 chemical elements?
This is a standard example for ACS which uses color coding to show the different categories as described in the corresponding index. But, obviously, the colors are not encoding individual elements (Figure 39).

PRO: the number of elements is limited.

CONTRA: the number of elements is relatively large - it lies beyond the number of easily differentiable unique colors.

POTENTIAL SOLUTION: find a color scale which can be used to encode 118 distinct colors in a meaningful way based on the rules to be developed in the context of perceptual visual attributes that better address color spaces, as the HCL color space and the constraints on color use from the user's perception.

2. How to distinctively color-code cell components?
Color coding based on unique and highly differentiated colors, as used in the CELLmicrocosmos CellExplorer based on [8], as in Fig 40.

PRO: the number of cell components is/can be relatively small - in the range of differentiable unique colors.

CONTRA: the definition of cell components is not clearly limited and depends on various properties, scales taken into account. Clear limitations will be required, which can be e.g. based on the inventory to be developed in the context of this work. Sources like Gene Ontology or Reactome might be relevant.

Discussion
In general, many components of the cell have no color and are just transparent.

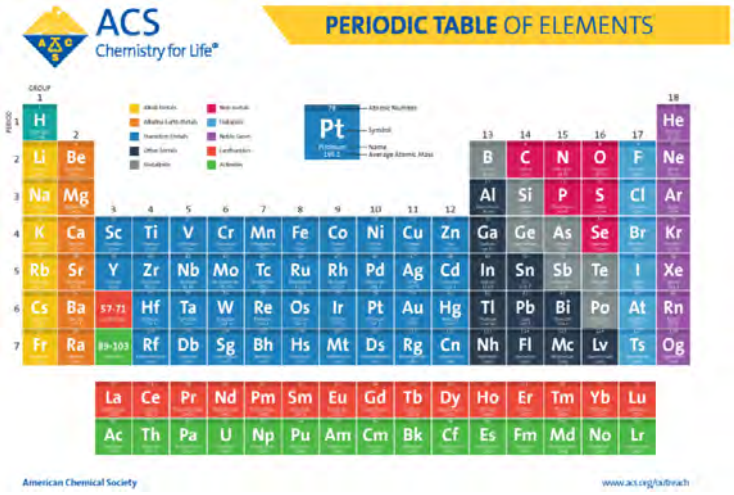


Figure 39: ACS Periodic Table of Elements

Ribosomes e.g. are usually transparent, but looking at them in an Electron Microscope shows clearly black structures depending on the fixation and imaging technique. On the other hand, Chloroplasts are essential cell components of plant cells and they contain chlorophyll - therefore, a green coloring of this cell component can be easily justified. In Light microscopy, using particular staining, whole areas of the cell might be shown in a specific color. E.g., using Haematoxylin and Eosin staining, leads to a blue central part of the cell/cell core/Nucleus area. The surrounding area is colored red.

Additional challenges related to color coding discussed are strongly related to the output medium:

- Using a 2D visualization, colors can be easily differentiated and boundaries can be clearly drawn, but usually 2D visualizations are strong abstractions. If microscopic images are used and then colored afterwards - e.g. electron microscopic images - then the structure is already very complex with a large amount of gray scales/contrast/etc.

- Using 3D visualization, the representation of colour becomes drastically more complicated. It depends on the perspective, on the shading model (contrast, the way, shadows are drawn or omitted) and positioning/type of light sources, on transparency and related overlays of different structures.

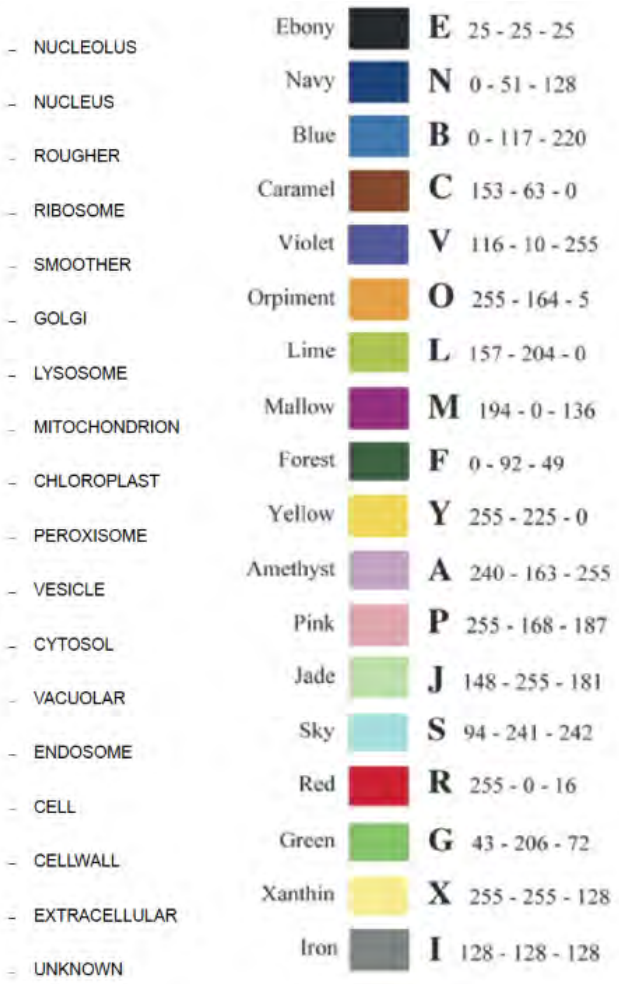


Figure 40: CellExplorer Example

A.5 Laying the foundations for Coherent Representation in Structural Biology

Maria Teresa Artese, Marc Baaden, Paola Perin, Micol Riva*
* Corresponding author: micol.riva@polimi.it

In the spirit of Fiorentini [37] and Elkins and Fiorentini [35], we attempted to isolate possible issues related to the objects intended to be represented (experimental results / raw data) and discussed an appropriate format to usefully store them (table, container, reservoir, Wunderkammer, catalogue raisonné, inventory .. ; see Fig. 41).

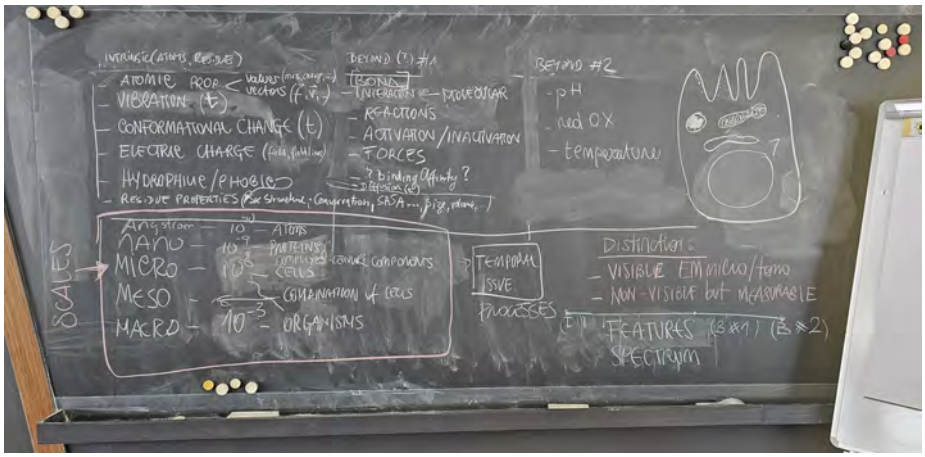


Figure 41

A.5.1 Introduction: overview of initial discussions

Initial discussions and problem framing of the early days identified several critical points in cataloging and consequently representing scientific objects in structural biology, particularly regarding the cellular scale. Key considerations include:

- Scientific representations cover a wide range of cell components and properties (i.e. properties of atoms, molecules, nucleus, surfaces, etc.; properties in context/environment; motion, forces, interactions; etc.). Depending on scales, there are a variety of relevant properties that need to be represented.
- There are several representation styles and methods: from hand-drawn styles to advanced Computer Grpahics; from static 2D to 3D images, to 4D dynamic visualization.
- Scientists use many databases, often relating to specific methods (e.g. PDB, IUPAC, SBGN, etc.), as well as many softwares, such as dedicated domain-specific scientific tools (e.g. PyMOL, ChimeraX, segmentation programs, etc.) and general graphic tools for science (e.g. Molecular Maya, Blender Molecular Nodes, etc.), each with its own presets for layout,

shapes, hues, etc. This variety of tools leads to a lack of standards and/or shared conventions among the scientific community.

- Nowadays, old and modern high-throughput technologies (e.g. Crystallography, Electron Microscopy, Cryo-Electron Microscopy, Cryo-Electron Tomography, etc.) and techniques generate huge amounts of intrinsically multidimensional data.

- Given the social nature of scientific visualization, it is essential to establish a common theoretical framework for effective representations. According to the terms introduced by Erna Fiorentini [37], this process implies:

  1. starting from the "What"–degree of visibility and degree of materiality,
  2. moving to the "How"—considering visualization as both a process and an image from selection, to choice and critical approach to a synthesis,
  3. considering the "Who" is looking and
  4. "For Whom" the visualization is intended, i.e. variants of looking (Elkins and Fiorentini [35], Parts 3 and 6).

  This process requires considering the "translator agency" topic and paying attention to the entire process of constructing knowledge (and visual information) from data, in all steps leading from data to visualization ([2], [40], [13]).

To briefly summarise the previous discussions on the representational topic, what emerged is the need to:

1. define a Theoretical Framework based on the "What", the "How", the "Who" and "for Whom" the representation process takes place;

2. build a common strategy by framing operations and selecting paradigmatic features, narrowing the field, and focusing on details ([2]) through four fundamental representational criteria:

   - reduction ([14]), abstraction and simplification as a way of excluding irrelevant information and highlighting the most important patterns, structures, processes in what we are observing;
   - clarity and structure;
   - use of visual metaphors ([1], [39]) and storytelling to enhance emotion and engagement, especially when dealing with a non-expert public;
   - use of local color meanings instead of color codes: hue can be considered as a second layer of meaning after defining layout, spatial variables and shapes.

Furthermore, a wide range of representational tools are available that allow for representing objects depending on the intended audience, including maps and atlas (i.e. Allen Institute Human Brain), 2D or 3D hand-drawn or digital illustrations, 2D or 3D animations, to mention a few.

Beyond the tools, to develop shared guidelines for representation, we must consider two major trends within the scientific context that may influence potential solutions. First, the predominance of printed publications. Second, the great variation in the types of data, ranging from raw data to interpreted data to data models. Both trends lead to a significant challenge: managing a high degree of communicative complexity.

A.5.2 Investigation Process: defining the "What"

To address these issues, we organised small working groups to tackle the different topics before opening the discussion to the whole team and questioning initial assumptions.

To narrow the field and better define the "What" of representation, we embraced the following 4-steps procedure:

1. Framing the problem
2. Analysing disciplinary tools and conditions
3. Discussing and conceptualizing appropriate data cataloging formats for representation
4. Laying the foundations for an inventory

1. Framing the problem
To circumscribe the problem of defining "What" we need to represent, the group first defined a list of objectives to achieve, namely:

- defining the different scales of "objects" (see above Chapter A.3) that scientists handle in various application fields of structural and molecular biology;
- listing the most relevant features scientists need or want to represent at different scales in publications and scientific studies;
- clustering the relevant features by giving shared definitions to build a more abstract framework to simplify and reduce the complexity of representation;
- finding common definitions for the different clustering categories and kinds of data that emerge from heterogeneous types of scientific measurements and observation instruments;
- finally, proposing an appropriate format to collect information by using well-known cataloguing tools (e.g. inventory; catalogue raisonné; table; etc.).

2. Structural and cellular biology tools and discipline conditions
Looking at the discipline of structural and cellular biology, we observed that measurement and observation tools and techniques vary greatly depending on the scales of the observed objects. This implies different ways of looking at objects at different scales.

Indeed, there are two main categories of tools. On the one hand, tools that primarily produce visual raw data (e.g. Crystallography, Electron Microscopy, Cryo-Electron Microscopy, Cryo-Electron Tomography, etc.), which serve as the basis, together with other (numerical) data, for deriving interpreted images. On the other hand, instruments that produce numeric raw data and require at least a double interpretive step to be visualised (e.g. molecular dynamics simulations, bioinformatics studies, etc.).

3. Discussing and conceptualizing appropriate data cataloging formats for representation

Before imagining proposals for representation, the need for shared criteria of collecting and organising data arises.

Considering the huge variety of features and properties of biological structures and their relevance at different scales, we explored the idea of **building an inventory**. Indeed, an inventory has the intrinsic property of reducing complexity by limiting and defining a concrete number of objects and values starting from a shared set of rules.

The inventory, understood as a comprehensive list of items related to the studied objects grouped into macro-categories, may in fact serve as a first step in the construction of a shared archive. It can also become an interoperable database once the scientific community as a whole agrees on the proposed criteria and rules.

To start building a coherent inventory, the set of rules proposed by the group were:

1. listing items and definitions;
2. range of scales: from Ångstrom (10^{-10}) to millimeters (10^{-3});
3. rank features/properties according to grades of relevance, depending on scales;
4. considerations on how to address time dependency in case of dynamic processes, interactions, reactions;
5. assess grades of measurability for each item, which may imply different grades of "visibility" (see the above paragraph "Introduction: overview of initial discussions" and A.4.3 "The constraints on color use from the user's perception").

4. Laying the foundations for an inventory

Step 1: listing properties and first clustering. As a first step, the small working group started collecting and listing the most important properties and characteristics of the cellular components to represent. At this stage, we disregarded which characteristic corresponded to which scale or component. Then, we shared this first attempt with the other participants.

After gathering some of the items, we made an effort to group the information into a more organised scheme. What emerged was a table (see Figure 42) divided into three levels of categorization identified as:

- Intrinsic properties

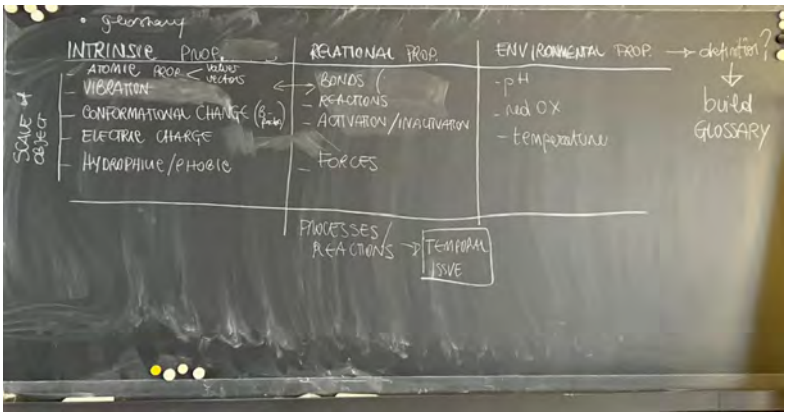


Figure 42: An initial scheme for an inventory based on three levels for categorizing properties to be displayed.

Table 1: Items and scales

Component scales	**Items (examples)**
Ångstrom (10^{-10})	atoms, molecule ...
Nanometers (10^{-9})	molecules, macro-molecules, molecular assembly, ...
Micrometers (10^{-6})	bacteria, organelles (nucleus, lysosomes, mitochondria, ...)
Meso (10^{-6} - 10^{-3})	cells, tissues, ...
Millimeters (10^{-3})	organisms

- Relational properties
- Environmental properties

In discussing this scheme, we faced a first challenge. Unable to agree on how to name the main categories, we felt the need to establish a glossary with detailed definitions, a task yet to be implemented, that could be discussed and shared by all members.

Step 2: second iteration on listing properties and clustering. Not to interrupt the listing work, we decided to leave the issue of nomenclature aside for the moment, while maintaining the original tripartition of the scheme. We then renamed the two controversial definitions "relational" and "environmental", assigning the neutral terms "beyond #1" and "beyond #2". Here, "beyond" refers to a property that we consider to exist at some level beyond the intrinsic property. In this second iteration some new topics emerged for discussion. It became fundamental to take the issue of scale into account (see Figure 43). To address this topic, we thus proceeded by listing the main scales handled by chemists and biologists and exemplifying which cellular components they may refer to items in Table 1.

Again, further challenges emerged, both in terms of features and representation.

- When it comes to the characteristics of atoms and cells pertaining to the "beyond #1" and "beyond #2" categories (which include actions/reactions,

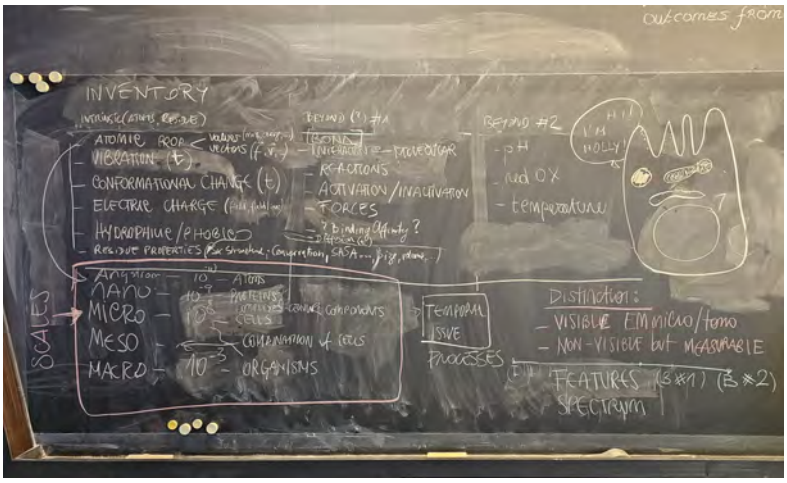


Figure 43: A collection of properties that may be visualized and how to arrange them from intrinsic object properties to emerging features beyond a given object.

forces and other dynamic processes) the temporal issue becomes extremely important. This topic must undoubtedly be taken into account when discussing methods and forms of representation so as not to undermine the communication goals.

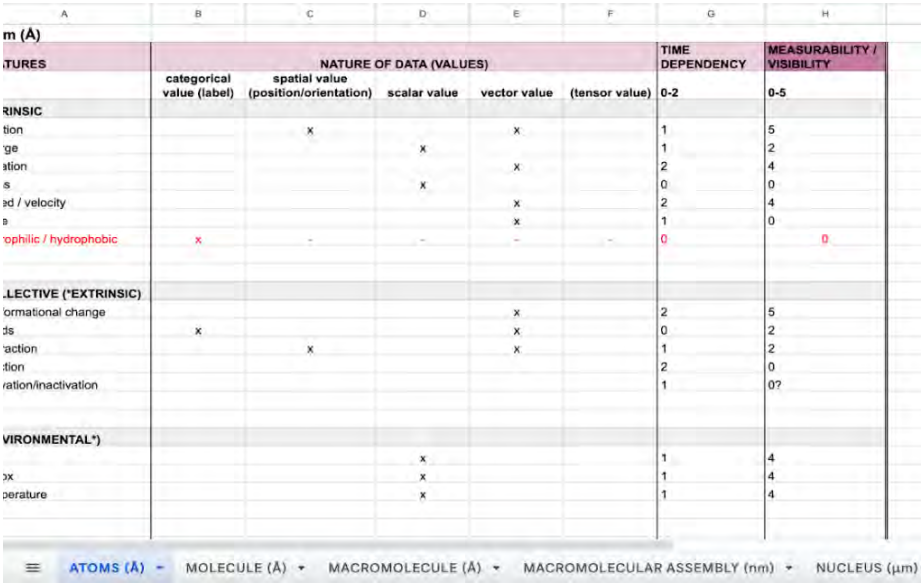
- When discussing dynamic processes, we must consider that the dynamic features of cellular components at different scales do not always constitute a discrete series, but sometimes a spectrum.
- Finally, when it comes to representation, there is the need for distinguishing between visible features derived from visual raw data, and non-visible but measurable features derived from numeric raw data (see above "Structural and cellular biology tools and discipline conditions").

Step 3: third iteration: tentative inventory in spreadsheet format. Given the complexity of the topic and the large amount of information to be organised in an orderly structure, the third step consisted of building a provisional inventory in the form of a spreadsheet.

We proceeded by organising a series of spreadsheets (Figure 44), each dedicated to different items at different scales, as specified in Table 1, and then organising the components' features and the values of data on the x and y axis.

Within the column of features, all the component properties were clustered into three main categories:

1. Intrinsic properties: inherent to objects and independent of any outside forces acting on them;
2. Beyond #1 properties [Collective/Extrinsic*]: leading to and/or emerging from interrelationships with other objects;
3. Beyond #2 properties [Environmental*]: contextual, i.e. environmental properties that influence the intrinsic and extrinsic properties of the components.



		NATURE OF DATA (VALUES)				TIME DEPENDENCY	MEASURABILITY / VISIBILITY
	categorical value (label)	spatial value (position/orientation)	scalar value	vector value	(tensor value)	0-2	0-5
RINSIC							
tion		x		x		1	5
ge			x			1	2
ation				x		2	4
s			x			0	0
ed / velocity				x		2	4
a				x		1	0
ophilic / hydrophobic	x	-	-	-	-	0	0
LECTIVE (*EXTRINSIC)							
ormational change				x		2	5
ds	x			x		0	2
action		x		x		1	2
tion						2	0
ation/inactivation						1	0?
VIRONMENTAL*)							
			x			1	4
ox			x			1	4
perature			x			1	4

ATOMS (Å) | MOLECULE (Å) | MACROMOLECULE (Å) | MACROMOLECULAR ASSEMBLY (nm) | NUCLEUS (µm)

Figure 44: A tentative inventory in spreadsheet format

We then tried to list all the possible natures of data, namely values, that can be encountered and retrieved from the "visual" (e.g. EM, C-EM, C-ET, etc.) and 'numeric' tools (e.g. models, simulations, bioinformatics, etc.) and are commonly used within the structural and cellular biology discipline.

In particular:

- categorical values: labels;
- spatial values: position and/or orientation;
- scalar values: quantity with magnitude but no direction (e.g. charge, mass, etc.);
- vector values: quantity defined by magnitude and direction (e.g. force, speed, etc.);
- tensor values: quantity defined by several magnitudes and directions (e.g. angular momentum, mechanical stress inside a rigid body, etc.).

During the construction of this provisional inventory, we pragmatically addressed two other important issues, both influencing the representation. On the one hand, the temporal issue, which we identified in the table as a measure of time dependency of the values, on a scale of 0 to 2. On the other hand, the visibility degrees that depend on the measurability of data sources, whether produced by "visual" or "numerical" measuring instruments, on a scale of 0 to 5. In summary:

1. Time dependency (0-2 scale):
    - 0 = time independent;
    - 1 = may change with time but may be pictured as static;

- 2 = time dependent

2. Measurability (0-5 scale) / grades of visibility: the scale goes down from 5-(directly) visible (e.g. a microscopic picture due to light or electron absorption ,etc.), via intermediate values signaling invisible but measurable (e.g. temperature determined by measuring tools) properties, to 0—theorized but not directly measurable (e.g. phylogenetic distance).

A.5.3 Final considerations

As work progressed, the complexity of creating a comprehensive inventory (including definitions, categorisations and clusterings) became evident. Trying to obtain a list of characteristics at the different scales of interest poses many challenges just by moving from the atomic to the nanoscale, where we started from, already.

Major open questions emerged:

- Feature-related
    1. Are the features local, or not?
    2. Are they inherently time-dependent, or not? If yes, how much?
    3. Are they actually measurable? If they are, is the measure direct, emergent, or derived from theoretical assumption?
- Communication Goals
    1. What features do we want to communicate?
    2. What is the purpose of communication?
    3. To whom are we speaking?
    4. Should we remain more abstract while defining the core features?
- Ethical and Theoretical questions
    1. Is there a consensus on a core set of characteristics?
    2. What are the limits of constructing such an inventory?
    3. Why not use existing ontologies to build a conceptual framework shared by the scientific community of the various fields of molecular and structural biology, so as to avoid semantic and categorisation misunderstandings ([22])?
    4. Are we trying to construct a "meta-ontology"?

It became obvious that there is no clear way to define an inventory without first finding answers to these questions. Above all, it is essential to extend the discussion to experts and scientists from different fields of application and dealing with different scales, starting from a shared basis. At the same time, using already existing ontologies could solve the problem of "starting from scratch" by "inventing" terms and categorisations that, in some cases, have already been shared and adopted by the scientific community.

A.6 Levels of looking: Towards a Multiscale Model in Scientific Imaging

Franco Achilli, Sander Ewen*, Matteo Farinella
* Corresponding author: sanderewen@gmail.com

After having shared and reviewed relevant examples of color coding in structural biology (A.4) and compiled an inventory of the desired properties that an ideal visualization should include (A.5), we proceeded to look for a way to structure this abundance of information.

For a good transfer of information to the public, it is important to provide information about the given context. Objects and properties of these objects, such as position, charge, vibration, speed, pH, temperature, etc., are not visible at all levels. For example, a molecule is not visible on a cellular scale. Moreover, even if multiple parameters could, in theory, be visible at the same scale, including too much information in a single image could make it unintelligible. For example, it may be counterproductive to visualize both pH and temperature gradients within the same image, since they both require color gradients. Our brain can only absorb a limited amount of information at any given time. A good scientific illustration should take that into account, visualizing relevant properties, without overwhelming our senses.

This gave rise to the primary idea of looking at an object within a cell from different levels of looking (for more on looking and levels of looking see [35]). A level of looking provides information about the context in which objects and properties of these objects can be seen. Informed by lessons in science communication [6], we turned to visual metaphors in order to provide an intuitive or perceptual understanding of this concept [29]. To use an analogy, when mapping a city (i.e. a cell) we may want to include simple representations of houses and buildings, but not individual rooms within those buildings, or objects within those rooms. At the other end of the scale, we may want to focus on an individual dog (i.e. a molecule) inside a specific room, and even zoom in on a specific flea on that dog (Figure 45).

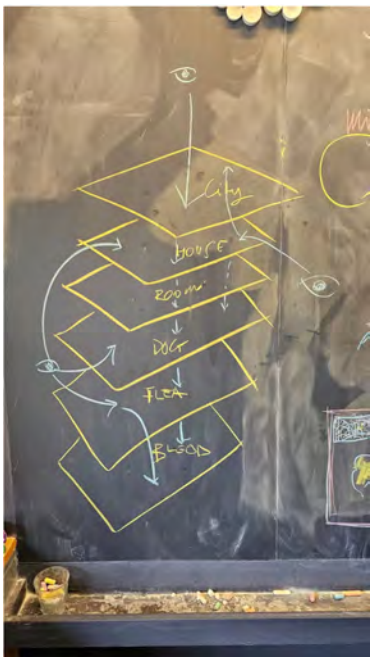


Figure 45: Levels of looking

Similarly, the total of possible properties in a cell are described in an inventory, in which only the relevant information for the viewing level is shown at one time. When creating an image, objects and any associated properties are displayed at a certain level of looking. The image is a fragment of reality, in which the information provided must match the viewer's frame of reference in order to place the information in the larger whole of knowledge of the cell.

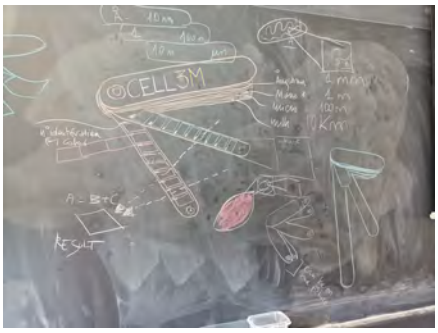


Figure 46: Pantone catalogue with properties from the inventory.

The idea developed further into an analogy of the Pantone overlay selector, where the identifying colorimetric shades (codes) are transparent and represent the different objects and properties from the inventory (see Figure 46). Objects and properties relevant to the information transfer, can be "laid on top of each other", whereby information can be combined with each other to form a more complex result.

When zooming in or out from a certain level of looking, new elements become visible and others disappear from view. Certain properties, such as pH and temperature, can move to a new level of looking, while this is more difficult for other properties or objects, since these are more scale-dependent. Reflection towards a more flexible, three-dimensional image is needed: a model of multi-level connections of the information (data) contained in images. Inspired by a multi-storey car park, structured for horizontal and vertical connections and passages, we then envisioned an initially prismatic structure, with connecting walkways at different levels, which could be chosen by the investigator or user of the system (Figure 47).

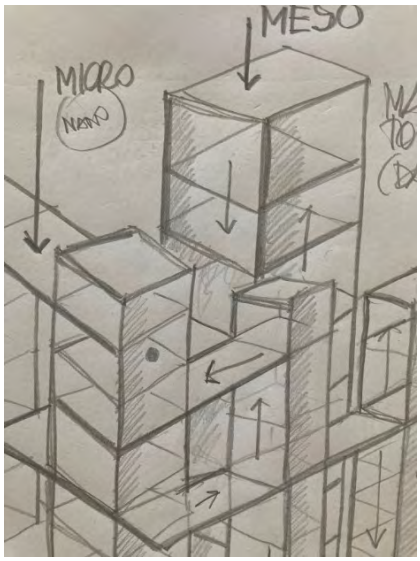


Figure 47: Multi-level connections of information (data).

Through consultation, the idea developed further into a more dynamic model, first with a cube (Figure 48) and then quickly with a cylinder (Figure 49). The cylindrical model represents a cylinder that can change in diameter, with the internal surfaces indicating the levels of looking. At the bottom of Figure 49 is the intrinsic level, the starting point in this case, where one can zoom in or out (levels of looking). By increasing the diameter of the cylinder (scale), the focal point moves from a singular object in relation to its environment. The cylinder represents a snapshot of a biological process, like a frame of an animation. When changing the cylinder in time, a dynamic arises like an animation. In time, the levels of looking and the scale are dynamic. In this case, one could conceive of a model where technology allows one to use 2D or 3D images or combine images into an animation, using even color alternatives to build ad-hoc scientific “narratives” on any image.

Further development of the reasoning proposes (Figure 50) to also consider “focal distance,” which we can interpret as an “extensible” point of observation: we then imagine the lens of an ideal camera moving along the axis of “focal distance,” just like the lens of a microscope, which either brings the relevant portion closer (isolating it) or – by moving away – allows us to consider a larger

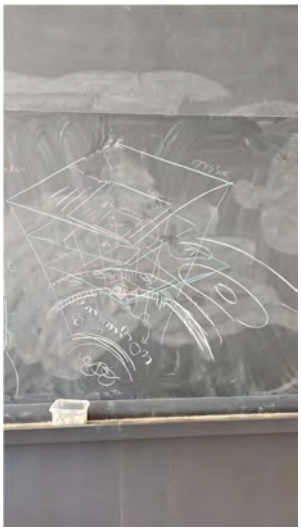


Figure 48: Dynamic model with a cube

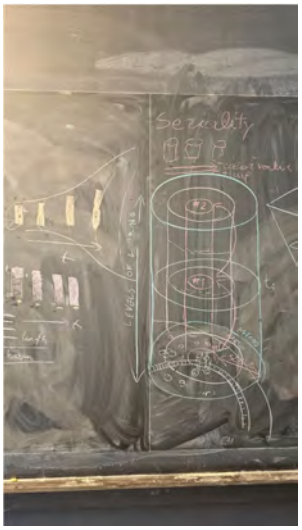


Figure 49: Cylindrical model with dynamics between the levels of looking and the focal scale

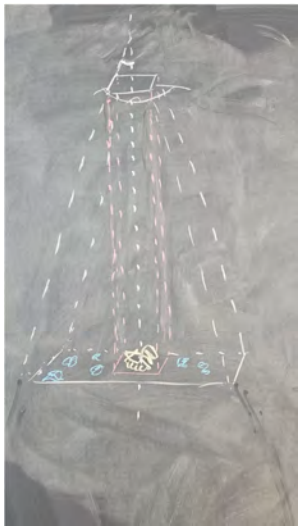


Figure 50: Dynamic model with focal distance in consideration

and more complete environment.

Finally, the intention to make the "progressive penetration" into the available information even more flexible led to the visualization of an additional, particularly compelling metaphor: the surface of a city that can be unveiled through the progressive lifting of (transparent) layers with obvious subjects. In this visual metaphor: the city is the level of organization; the train station is organelle; the train is the vesicle; the animal species in the train car are the protein complexes; the faces of the passengers are the single proteins; the passenger's eye is sugars on individual proteins; we finally arrive at the cell with detail at the atomic level.

A.7 Representing complexity in structural biology through visual and audiovisual storytelling

Jim Procter*, Micol Riva*, Vera Williams*
* Corresponding authors: j.procter@dundee.ac.uk, micol.riva@polimi.it, vera.williams@pasteur.fr

The first investigation established a framework and catalogs of observations of the players and behaviours comprising the biology of the cell. In the second, a conceptual device was developed through which observers can experience and engage with phenomenologies taking place in cells across temporal and spatial scales. The device is general, in that it can be employed by observers with different levels of knowledge, and diverse goals - be it to explore, measure, or gain deeper understanding of what is known and what is yet to be understood.

For our third investigation, we aimed to identify the modes and design parameters that will allow communication between this conceptual mechanism and the observer subjected to its experience. In reality, these communications are frequent and remarkably commonplace. For instance, a teacher narrating a series of animations to provide context allows for more broad recognition of critical aspects so they are more quickly understood by learners. Similarly, visual and audio design are prerequisites to effective production of experiences that are intended to affect the observer.

As a preliminary step, examples from world renown biological communicators were critically discussed. Principal amongst these were Anders Ynnerman's narration of Drew Berry's animated film "Chemistry of Life", a film designed as an immersive experience accompanied by expert narration. The viewer is taken on a journey through the scales of life; from the familiar experience of walking through a park and observing fruiting apple trees, through their leaves and down into the subcellular compartments of chloroplasts. In this way, the observer's real experiences are related to molecular processes taking place at near infinitesimal timescales through storytelling techniques enacted through visual and auditory media.

A.7.1 Designing for effective communication

A critical element to take into account when visualising such complex systems as cellular interactions on a molecular level, is cognitive load. When dealing with complex subjects that are far removed from everyday reality, it is easy to get overwhelmed or lose one's footing. Cognitive load can be managed with storytelling, controlled use of color, values, and composition, pointed use of illusions (isometric visualisation), and sound, for example.

1. Storytelling as a guide into the molecular world
When designing a visualisation on the molecular level we aim to communicate topics that are far removed from our reality. Molecular visualisations don't have an intuitive frame of reference, as its reality is not rooted in our day-to-day reality. A relatable and understandable frame of reference is essential for effective communication of scientific topics. The question "for whom is this visualisation intended?" is a necessary guide for any scientific visualisation. In practice, this means that a visualisation that is intended for in-reach purposes (scientist to

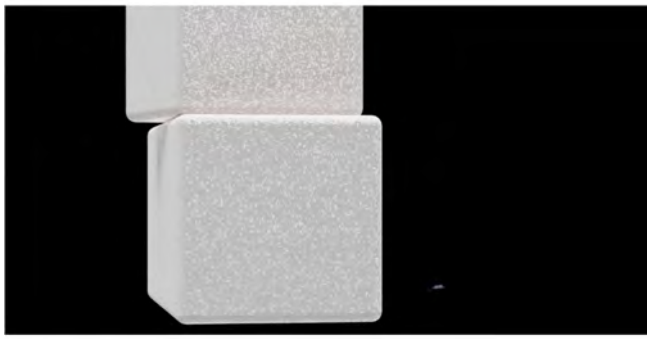


Figure 51: Using a sugar cube as a frame of reference

scientist communication), has a different frame of reference than a visualisation on the same topic that is designed for out-reach purposes (communication to a layperson).

When a suitable frame of reference has been decided on, it can be helpful to guide the viewer through several steps "into" the molecular world. This is especially helpful for out-reach communication as lay people have little to no experience with molecular visualisations. It is, for example, not too difficult for a layperson to follow the steps from taking a sample from our mouth, to putting that sample under a lens, to looking at the bacterium Treponema denticola and its internal machinery that drive its behaviour as in the short "Treponema through the ages". In three understandable steps, a layperson takes a huge leap into the depth of the molecular world and its molecular machines. This example also utilises a recognisable everyday item (a sugar cube, Figure 51) as a measuring device to help the viewer understand the relative size of a microbe in an attempt to elucidate the difference in scale of this microscopic world compared to our everyday reality.

2. Optimizing the output type

As a next step, we can optimise the output. Ideally, any visualisation generates output that stays as close as possible to the most accurate scientific data available on the topic; if a molecular model is available of a protein, use that model from the database in the visualisation with minimal abstractions (such as lowering the resolution, adjusting the shape etc). However, the visualisation is always limited by the output possibilities (analog/digital, static/dynamic, one-directional/interactive), as this will automatically result in abstractions of the original data.

For example: the viewer loses the ability to view a 3D protein model in its totality when it is printed in a magazine. Steps can however be taken to compensate for this; by creating a visualisation with good lighting to indicate its shape (Figure 52), by providing multiple views to give the audience the ability to visualise the protein from all sides in their mind's eye (Figure 53), and by providing a gimbal to demonstrate the relative orientation, for example. In essence, any visualisation should keep in mind the fundamental artistic rules of good value, composition and colour use to provide the viewer with an experience that resembles our day-to-day visual experiences [32].

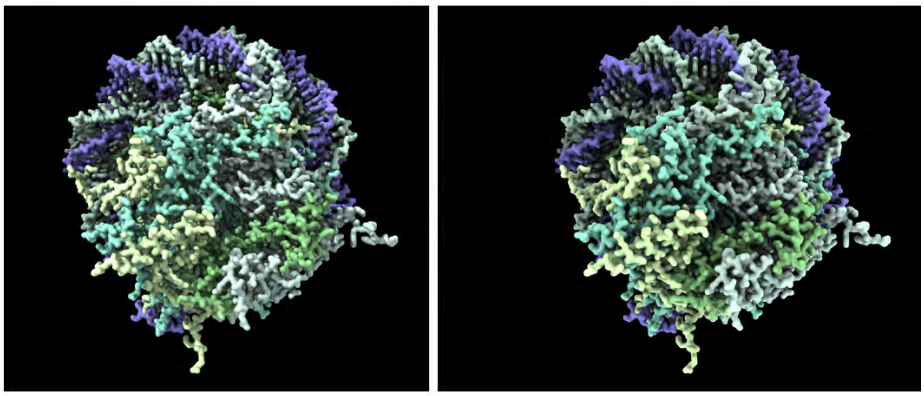


Figure 52: The effect of good lighting on conveying depth in an image: Due to the use of shadows, the 3D shape on the right is easier to read.

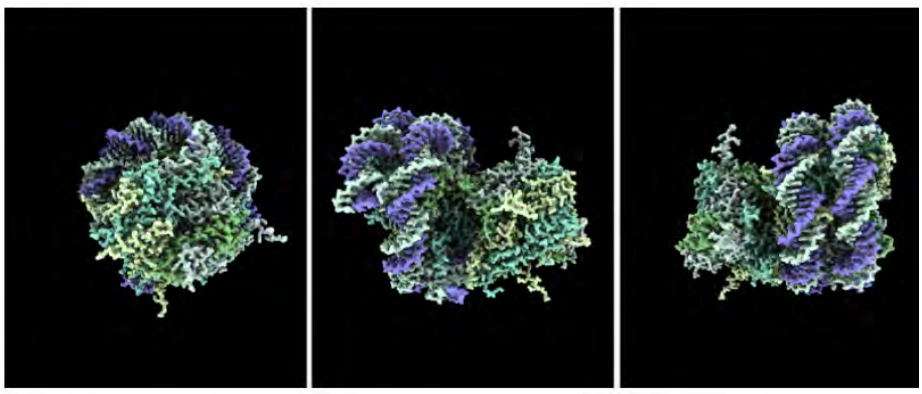


Figure 53: Showing a protein from multiple sides to provide an idea of its 3D shape (3LEL from PDB.org was used as a demonstration)

Figure 54: An example of illustrating a complex system (although it's not a molecular one) with an isometric illustration. The items in front have the same size as the objects in the back.

3. Isometric visualisations for overview
Isometric illustration offers an additional method to provide a clear overview of a complex system (Figure 54). As the angles are drawn equally on all axes, an isometric illustration offers the viewer an overview with depth, but without any foreshortening. It is important to note that this type of illustration is in itself an abstraction of reality as it lacks perspective. However, precisely because it lacks perspective, it provides an opportunity to show interconnected systems however far removed from each other in space. This allows the creation of a full overview of an entire system in one illustration. Naturally, a big limiting factor of such an illustration is the resolution, however when used in the form of an animation a bigger overview can be created and close ups of the individual systems within the greater whole can be highlighted in subsections.

A.7.2 Designing sound to enhance dynamic representations

Due to its inherent spatiotemporal qualities and its embodiment in perception and action, sound—and listening—can serve as valuable means to enhance dynamic visual narratives. Indeed, the cross-modal integration between auditory and visual stimuli can evoke unexpected sensory effects and emotional responses, shaping audience perception and cognition in meaningful ways. The "corporeal reverberation" induced by cross-sensory metaphors [16] can connect cultures, languages, and backgrounds [39] through embodied representations. Therefore, designing consistently the sound-image dramaturgy can undoubtedly boost the interpretation of complexity in structural biology representations.

1. The "audiovisual contract": a multimodal narrative approach
Many of the features defined in Chapter A.5, such as processes, vibrations, speed, conformational changes, reactions, and interactions, imply time as a pri-

mary characteristic whose effects are at least measurable, if not visible. Since time, or rather duration and rhythm, are inherent in sound, and sound is profoundly embodied in perception and action, narrative sounds and sound-driven design strategies [48, 18, 51] can undoubtedly enhance dynamics representation, especially when it comes to representing processes through 2D and 3D video animations. The interplay between narrative and experiential time allows for manipulating perceptions, enabling stories to expand and contract while offering multiple levels of interpretation.

Moreover, as thoroughly explored by Michel Chion in "L'audio-vision. Son et image au cinéma" [34], despite sound and visual perception being diverse by nature, "in the audiovisual contract these perceptions influence each other and lend each other, by contamination and projection, their respective properties" thus enhancing and modifying their respective meanings. Therefore, in the dynamic representations of structural biology, we can use sound to convey content that the abstraction and simplification required by moving images cannot represent. Hence, a multimodal approach integrating audio and visual narratives seems particularly appropriate when there is a need, as in our specific case, to translate complex datasets into accessible outputs aimed at different kinds of audiences (i.e. scientists, students, the general public).

Designing sound is a proper communicative act [45] on par with visual design: both, indeed, serve to design relationships [12] among different objects, processes, and values. In its multiple expressions—i.e. vocals/storytelling, effects, concrete sounds, music, silence, noise—and thanks to its physical and psychophysical properties—i.e. frequency/pitch, amplitude/volume, timbre—sound owns intrinsic narrative functions, either aesthetic, evocative or informational. For these reasons, it can undoubtedly serve both as a message and a medium [42] for interpretation, emotion, learning, engagement, inclusion and identity [28].

2. Sonic Translations

We can translate sound features into structural and molecular representation in two primary ways:

1. The literal approach traces a direct connection between the sonic spectrum (waveform, frequency/pitch, amplitude/volume, timber) and time-dependent molecular features. This approach, however, poses two sets of problems. On the one hand, apart from rare cases, there is a lack of shared sound-image correspondences [16], which adds complexity to the data sonification process. On the other hand, the large amount of data to be sonified favours the risk of generating cacophonies.

2. The metaphorical approach can suggest and convey time-dependent molecular characteristics sonically by appealing to or evoking everyday perceptual situations.

However, to achieve the communication goals, we should aim at a third strategy that consists of an integrated mix of the two modalities through the association of bio-data input values to more musical and rhythmic elements, such as in the examples below:

1. Mouse hippocampus sonification
2. Kyle Johnsen – 'Step the Brain along a Path'

3. Stuart Mitchell – The human genome music project
4. Protein Music Sound Atoms
5. Sound of Atoms
6. Drew Berry – bio animations
7. Marc Temple – Can COVID sound Beautiful
8. Music with plants ch.4

3. Problem of scales
This problem stems from the inherent difficulty of consistently translating process time (quantitative data values) into narrative time, which can be both linear and non-linear, into perceptual time, which is inherently subjective and qualitative (Figure 55). This issue can negatively influence the interpretation process, affecting cognitive load (see above "Designing for effective communication").

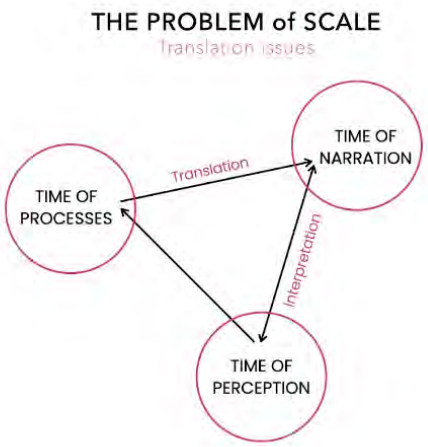


Figure 55: The problem of scale. Issues in translating temporal values going from processes to perception to narration.

A.7.3 Narrative design as a way of managing complexity

"The love of complexity without reductionism makes art; the love of complexity with reductionism makes science.". EO Wilson, Consilience, 1998.

The producers of representations of the cell's complexity must forever strike a balance if they are to achieve their goals in accessible and impactful ways. Compromise is inevitable; not only because of limitations of the representation's medium, but also due to the diverse range of phenomena that we know of, and those we continue to discover.

Our investigations in this third step reinforced the central role of narrative in communicating the biological meaning of phenomena in the cell. Narrative

allows the conjunctions of physical phenomena that comprise these complex events can be decomposed into facets, allowing them to be depicted and understood when experienced by the observer. Narrative is one of the most important techniques of science communication, but predates contemporary notions of the field. In order for our conceptual device to be wielded with effectiveness, we must also consider other techniques employed in the arts. Aristotle's Unities of Action, Time and Place lay the foundations for the design of effective narrative [33], both as rules to be upheld, and rules to be broken. Whilst narrative should ideally adhere to the unities, transgressions of any one are permitted, if only to ensure the observer is not overwhelmed, and continues to engage. In a similar manner, rigid adherence to convention or standardised representations limits the capacity for communication of complex biology, particularly those that have yet to be formalised as concise referential knowledge.

Whilst the number of actors on the stage of any one of the many fundamental biological processes is small, it is the multiplicity and scales of time and space that challenges understanding at a cellular level. Producers and designers across the creative sectors have considerable experience in channeling narrative complexity in sound, vision and other senses. Continued engagement - not just between scientists and artists, but also with other creatives will facilitate creation of new ways to approach this challenge.

References

[1] Roland Barthes. "Rhétorique de l'image". In: *Communications* 4.1 (1964), pp. 40–51. DOI: 10.3406/comm.1964.1027.

[2] Françoise Bastide. "The iconography of scientific texts: principles of analysis". In: *Representation in scientific practice* (1990), pp. 187–229.

[3] Jonathan Brecher. "Graphical representation standards for chemical structure diagrams (IUPAC Recommendations 2008)". In: *Pure and Applied Chemistry* 80.2 (2008), pp. 277–410.

[4] Muyuan Chen. "Rendering protein structures inside cells at the atomic level with Unreal Engine". In: *bioRxiv* (2024). DOI: 10.1101/2023.12.08.570879. eprint: https://www.biorxiv.org/content/early/2024/03/09/2023.12.08.570879.full.pdf. URL: https://www.biorxiv.org/content/early/2024/03/09/2023.12.08.570879.

[5] James Elkins. "Art History and Images That Are Not Art". In: *The Art Bulletin* 77.4 (1995), pp. 553–571. DOI: 10.2307/3046136.

[6] M. Farinella. "Of Microscopes and Metaphors: Visual Analogy as a Scientific Tool". In: *Comics Grid-Journal of Comics Scholarship* 8 (2018). ISSN: 2048-0792. DOI: 10.16995/Cg.130.

[7] Moshe Glickman and Tali Sharot. "AI-induced hyper-learning in humans". In: *Current Opinion in Psychology* 60 (Dec. 2024), p. 101900. ISSN: 2352250X. DOI: 10.1016/j.copsyc.2024.101900.

[8] Paul Green-Armytage. "A colour alphabet and the limits of colour coding". In: *JAIC-Journal of the International Colour Association* 5 (2010).

[9] M. Gail Jones, Amy Taylor, James Minogue, Bethany Broadwell, Eric Wiebe, and Glenda Carter. "Understanding Scale: Powers of Ten". In: *Journal of Science Education and Technology* 16.2 (2007). Publisher: Springer, pp. 191–202. ISSN: 10590145, 15731839. URL: http://www.jstor.org/stable/40188687 (visited on 09/26/2024).

[10] Kèvin Knoops, Rinse De Boer, Anita Kram, and Ida J. Van Der Klei. "Yeast *pex1* cells contain peroxisomal ghosts that import matrix proteins upon reintroduction of Pex1". In: *Journal of Cell Biology* 211.5 (Dec. 7, 2015), pp. 955–962. ISSN: 0021-9525, 1540-8140. DOI: 10.1083/jcb.201506059.

[11] Akihiko Konagaya, Gregory Gutmann, and Yuhui Zhang. "Co-creation environment with cloud virtual reality and real-time artificial intelligence toward the design of molecular robots". In: *Journal of Integrative Bioinformatics* 20.1 (June 2022). DOI: 10.1515/jib-2022-0017.

[12] Stefano Luca, Nicolas Misdariis, and Alberto Calleo. "Designing Sounds Means Designing Relationships". In: *DIID* 1 (78 Dec. 2022). DOI: 10.30682/diid7822m.

[13] Valentina Manchia. "Beyond immediacy and transparency. A semiotic approach to discursive and rhetorical strategies in media visualization and data visualization". In: *Punctum. International Journal of Semiotics* 8 (2022), pp. 85–113. DOI: 10.18680/hss.2022.0006.

[14] Lev Manovich. "What Is Visualisation?" In: *Visual Studies* 26 (Mar. 2011), pp. 36–49. DOI: 10.1080/1472586X.2011.548488.

[15] Martina Maritan, Ludovic Autin, Jonathan Karr, Markus W. Covert, Arthur J. Olson, and David S. Goodsell. "Building Structural Models of a Whole Mycoplasma Cell". In: *Journal of Molecular Biology* 434.2 (2022), p. 167351. ISSN: 0022-2836. DOI: 10.1016/j.jmb.2021.167351.

[16] Lawrence Marks. "Synesthesia, Then and Now". In: *Intellectica. Revue de l'Association pour la Recherche Cognitive* 55 (Jan. 2011), pp. 47–80. DOI: 10.3406/intel.2011.1161.

[17] Lisa Messeri and M. J. Crockett. "Artificial intelligence and illusions of understanding in scientific research". In: *Nature* 627.8002 (Mar. 7, 2024), pp. 49–58. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-024-07146-0. (Visited on 06/12/2025).

[18] Stefano Delle Monache, Nicolas Misdariis, and Elif Özcan. "Semantic models of sound-driven design: Designing with listening in mind". In: *Design Studies* 83 (2022), p. 101134. ISSN: 0142-694X. DOI: 10.1016/j.destud.2022.101134.

[19] Steven D Palkovic, Dieter B Brommer, Kunal Kupwade-Patil, Admir Masic, Markus J Buehler, and Oral Büyüköztürk. "Roadmap across the mesoscale for durable and sustainable cement paste–A bioinspired approach". In: *Construction and Building Materials* 115 (2016), pp. 13–31.

[20] A. Rychlý, J. Byška, B. Kozlíková, and K. Furmanová. "Tasks and Visual Abstractions for 3D Chromatin Representation". In: *Computer Graphics Forum* 44.3 (2025), e70142. DOI: https://doi.org/10.1111/cgf.70142.

[21] Karsten Schatz, Juan José Franco-Moreno, Marco Schäfer, Alexander S. Rose, Valerio Ferrario, Jürgen Pleiss, Pere-Pau Vázquez, Thomas Ertl, and Michael Krone. "Visual Analysis of Large-Scale Protein-Ligand Interaction Data". In: *Computer Graphics Forum* 40.6 (2021), pp. 394–408. ISSN: 1467-8659. DOI: 10.1111/cgf.14386.

[22] Nadine Schuurman and Agnieszka Leszczynski. "Ontologies for Bioinformatics". In: *Bioinformatics and Biology Insights* 2 (2008), BBI.S451. DOI: 10.4137/BBI.S451.

[23] Björn Sommer, Daisuke Inoue, and Marc Baaden. "Design X Bioinformatics: A Community-Driven Initiative to Connect Bioinformatics and Design". In: *Journal of Integrative Bioinformatics* 19.2 (June 2022). Publisher: De Gruyter. ISSN: 1613-4516. DOI: 10.1515/jib-2022-0037.

[24] Davide Spalvieri, Anne-Marine Mauviel, Matthieu Lambert, Nicolas Férey, Sophie Sacquin-Mora, Matthieu Chavent, and Marc Baaden. "Design – a new way to look at old molecules". In: *Journal of Integrative Bioinformatics* 19.2 (2022), p. 20220020. ISSN: 1613-4516. DOI: 10.1515/jib-2022-0020.

[25] Anna Sterzik, Nils Lichtenberg, Michael Krone, Daniel Baum, Douglas W. Cunningham, and Kai Lawonn. "Enhancing molecular visualization: Perceptual evaluation of line variables with application to uncertainty visualization". In: *Computers & Graphics* 114 (2023), pp. 401–413. ISSN: 0097-8493. DOI: 10.1016/j.cag.2023.06.006.

[26] John Ridley Stroop. "Studies of Interference in Serial Verbal Reaction". In: *Journal of Experimental Psychology* 18.6 (1935), pp. 643–661.

[27] Nicholas Waldin, Mathieu Le Muzic, Manuela Waldner, Eduard Gröller, David Goodsell, Autin Ludovic, and Ivan Viola. "Chameleon - Dynamic Color Mapping for Multi-Scale Structural Biology Models". In: *Eurographics Workshop on Visual Computing for Biology and Medicine* (2016). Artwork Size: 10 pages ISBN: 9783038680109 Publisher: The Eurographics Association, 10 pages. ISSN: 2070-5786. DOI: 10.2312/VCBM.20161266. URL: https://diglib.eg.org/handle/10.2312/vcbm20161266 (visited on 05/11/2022).

[28] Kathleen Wiens and Eric de Visscher. "How Do We Listen To Museums?" In: *Curator: The Museum Journal* 62.3 (2019), pp. 277–281. DOI: 10.1111/cura.12318.

[29] M. Zoppè. "Towards a perceptive understanding of size in cellular biology". In: *Nature Methods* 14.7 (June 29, 2017). ISSN: 1548-7105 (Electronic) 1548-7091 (Linking). DOI: 10.1038/nmeth.4300.

[30] Monica Zoppè. "Colors in the representation of biological structures". In: *Journal of Integrative Bioinformatics* 19.2 (July 2022), p. 20220021. ISSN: 1613-4516. DOI: 10.1515/jib-2022-0021. (Visited on 08/10/2022).

[31] Hajime Chijiiwa, ed. *Encyclopedia on Color Cognition of the World's Youth*. In Japanese. Tokyo: Kawade Shobou Shinsha, 1999.

[32] Publishing 3dtotal. *Artists' master series Color and Light*. 3dtotal Publishing, 2021. ISBN: 9781912843411.

[33] Aristotle. *Poetics*. Ed. by D.W. Lucas. Oxford University Press, 1968. ISBN: 9780198140245.

[34] Michel Chion. *Audio-Vision: Sound on Screen*. Columbia University Press, 2019. ISBN: 9780231185882. URL: http://www.jstor.org/stable/10.7312/chio18588 (visited on 06/06/2025).

[35] James Elkins and Erna Fiorentini. *Visual worlds: looking, images, visual disciplines*. Oxford / New York: Oxford University Press, 2021. 435 pp. ISBN: 978-0-19-939091-5. URL: https://global.oup.com/ushe/product/visual-worlds-9780199390915?cc=it&lang=en&.

[36] Jörg von Engelhardt. *The language of graphics: A framework for the analysis of syntax and meaning in maps, charts and diagrams*. Yuri Engelhardt, 2002.

[37] Erna Fiorentini. *On Visualization: A Multicentric Critique Beyond Infographics*. Vol. 21. LIT Verlag Berlin, 2024. Chap. "Three (and a Half) Parameters to Disentangle Visualization," 19-54. ISBN: 9783643905352. URL: https://lit-verlag.de/isbn/978-3-643-90535-2/.

[38] Kurt Koffka. *Principles of Gestalt Psychology*. London: Routledge & Kegan Paul Ltd., 1935.

[39] George Lakoff and Mark Johnson. *Metaphors We Live By*. Chicago, IL: University of Chicago Press, 1980.

[40] Valentina Manchia. *Il discorso dei dati. Note semiotiche sulla visualizzazione delle informazioni*. FrancoAngeli, 2020.

[41] Scott McCloud. *Understanding comics: the invisible art*. Northampton, MA: Tundra Pub, 1993. 215 pp. ISBN: 978-1-56862-019-0.

[42] Marshall McLuhan. *Understanding media : the extensions of man*. New American Library, 1988. ISBN: 0451624963. URL: http://www.worldcat.org/search?qt=worldcat_org_all&q=0451624963.

[43] Ettore Sottsass. *Vorrei sapere perché = I wonder why*. Italy: Mondadori Electa, 2007. ISBN: 9788837057343.

[44] D. Travis. *Effective Color Displays: Theory and Practice*. Computers and people. Academic Press, 1991. URL: https://books.google.it/books?id=ut40mwEACAAJ.

[45] Barry Truax. *Acoustic communication*. Ablex Pub. Corp, 1984, p. 244. ISBN: 0893912638.

[46] Edward Tufte. *Envisioning information*. USA: Graphics Press, 1990. ISBN: 0961392118.

[47] Florian Kräutli and Stephen Boyd Davis. "Known Unknowns: representing uncertainty in historical time". In: *Proceedings of Electronic Visualisation and the Arts 2013:* ed. by K Ng, JP Bowen, and S McDaid. London: British Computer Society, 2013, pp. 61–68. ISBN: 978-1-780172-15-6. URL: http://ewic.bcs.org/content/ConWebDoc/50996.

[48] Stefano Delle Monache, Nicolas Misdariis, and Elif Ozcan. "Conceptualising sound-driven design: an exploratory discourse analysis". In: *Proceedings of the 13th Conference on Creativity and Cognition*. C&C '21. New York, NY, USA: ACM, 2021. ISBN: 9781450383769. DOI: 10.1145/3450741.3465258.

[49] Claire Dormann, Corina Pötscher, and Guenter Wallner. "A Classification of Video Game Cartographic Maps". In: *Conference Proceedings of DiGRA 2023 Conference: Limits and Margins of Games Settings*. Tampere: DiGRA, 2023.

[50] August Wilhelm von Hofmann. "On the combining power of atoms". In: *Proceedings of the Royal Institution of Great Britain*. Vol. 4. Reprinted in Chemical News 12 (1865): 166–190. Royal Institution of Great Britain. 1865, pp. 401–430.

[51] Stefano Delle Monache, Elif Özcan, and Nicolas Misdariis. "Designing [The, With, Against] Sound [For]: Towards A Semantic-oriented Coding Scheme For Protocol Studies In Sound-driven Design". In: *DRS*. June 2024. DOI: 10.21606/drs.2024.761.

[52] C. Schulz, K. Schatz, M. Krone, M. Braun, T. Ertl, and D. Weiskopf. "Uncertainty Visualization for Secondary Structures of Proteins". In: *IEEE Pacific Visualization Symposium*. 2018, pp. 96–105. DOI: 10.1109/PacificVis.2018.00020.

[53] *Powers of Ten*. 1977.

[54] *Mycoplasma Model*. URL: https://ccsb.scripps.edu/gallery/mycoplasma_model/.

[55] *Mesoscale Explorer*. URL: https://molstar.org/me/.

[56] *Pantone colors*. URL: https://www.pantone.com/hk/en/.

[57] Jacques Bertin. *Semiology of graphics: diagrams, networks, maps*. 1967.

[58] Jeff Pettiross. *Crafting an effective data visualization color palette*. Apr. 2024. URL: https://observablehq.com/blog/crafting-data-colors.

[59] Maureen Stone. *How we designed the new color palettes in tableau 10*. Apr. 2025. URL: https://www.tableau.com/blog/colors-upgrade-tableau-10-56782.

[60] Kazunori Asada. *Chromatic Vision Simulator on Website*. 2019. URL: https://asada.website/webCVS/ (visited on 06/05/2025).

[61] Vera Williams and Meindert Lamers. *Ageing on a molecular level*. Nov. 2024. URL: https://animini.nl/aging-on-a-molecular-level.

[62] Vera Williams, Alise Muok, and Ariane Briegel. *Treponema through the ages*. Dec. 2023. URL: https://animini.nl/treponema-through-the-ages.

[63] Vera Williams and Rouchen Ouyang. *Flowerphage ϕKp24 animation 2.0*. Nov. 2022. URL: https://animini.nl/flowerphage-kp24-animation-20.

[64] Jason Ryan Berg. "How Monstrosity and Geography were used to Define the Other in Early Medieval Europe". Unpublished. Oct. 2015. URL: https://etheses.whiterose.ac.uk/id/eprint/11353/.