Health information design model (HIDEM): A replicable model of the design process for data-intensive applications in health informatics

Antonio Solano
Flatiron Health

Miso Kim
Northeastern University, United States of America

Follow this and additional works at: https://dl.designresearchsociety.org/drs-conference-papers

Citation

This Research Paper is brought to you for free and open access by the DRS Conference Proceedings at DRS Digital Library. It has been accepted for inclusion in DRS Biennial Conference Series by an authorized administrator of DRS Digital Library. For more information, please contact dl@designresearchsociety.org.
Health Information Design Model (HIDeM): A replicable model of the design process for data-intensive applications in health informatics

Antonio Solano-Roman\textsuperscript{a,b}, Miso Kim\textsuperscript{a}\textsuperscript{*}
\textsuperscript{a}College of Arts, Media and Design, Northeastern University, Boston, USA
\textsuperscript{b}Broad Institute of MIT and Harvard, USA
\textsuperscript{*}corresponding email: m.kim@northeastern.edu

https://doi.org/10.21606/drs.2022.208

Abstract: Genetic and healthcare data have experienced an explosion in the last decade and with it, a deluge of new and cutting-edge research as well as digital tools and software. However, visualizing and working with such large amounts of information also poses organizational challenges in transdisciplinary collaboration between scientific domain experts and design professionals. We propose the Health Information Design Model (HIDeM), a collaborative process model with a novel series of principles and activities that tackle the specific nature of digital, data-intensive products for scientific research while considering different organizational contexts. The model was developed in collaboration with several domain experts in the life sciences and has been tested in diverse scenarios. We also present a use case that demonstrates this model's potential and outcomes.

Keywords: health informatics; design for healthcare; information design; interdisciplinary collaboration

1. Introduction

Life science research is being revolutionized by big data, bioinformatics, and the growing trend of creating user-centered digital products. For instance, the first half of 2021 saw a record $15 billion in investments for digital health companies in the United States (Mercom Capital Group, 2021). In addition, the United States Food and Drug Administration launched a pilot program in 2018 (FDA) to allow private companies to design digital healthcare products without the historical and cumbersome constraints that have made software development slow. Since then, the FDA has continued to issue guidance documents to increase the use of electronic health data in health care decisions (Office of the FDA Commissioner, 2021). At the same time, health science researchers have turned to translational groups with diverse professionals that can provide their expertise and collaborate to solve practical problems. In this new age of emerging infectious disease and
complex medical challenges, it is imperative to create a framework for effective collaboration within these groups.

A particularly recurring need within these groups is to design interactive visualizations and Graphical User Interfaces (GUIs) to study data used in life sciences. This is due in part to the release of more, higher quality data across the world, as well as the digitalization of previously analog data. Notable examples include the innovative Circos visualization (Krzywinski et al., 2009) used for the display of multidimensional genomic data. Another example is the NextStrain platform (Hadfield et al., 2018), used for tracking disease epidemics through genomic sequencing which has gained paramount importance during the COVID-19 pandemic.

These changes call for tighter collaboration between design, data visualization, and science experts. We propose that designers can play an important role as mediators in transdisciplinary groups that come together to create user-centered digital products for non-linear visual analytics. These digital products are characterized by the visualization process being "iterative, where the user interface acts as link between data and user" (Sturm et al., 2015) and by a back-and-forth exploration process among visuals, sensemaking, and GUI.

The design of digital visualizations that allow domain-experts to interact with complex datasets and look for undiscovered relationships and patterns requires close collaboration between researchers, health specialists, data scientists, software developers and designers. However, it is only until recently that academia and healthcare industries have turned to design as a discipline to support systematic approaches to user-centered problem solving (FDA 2018). This presents both a new opportunity and many new challenges for designers, who must act as mediators between all the aforementioned stakeholders in the astringent and relatively inflexible context of medicine and scientific discovery. There is a need for a clearly defined and transparent workflow process that everybody can understand and replicate across various disciplines.

This article introduces a model for a collaborative process to develop data-intensive applications in these translational life science research in healthcare and bioinformatics: the Health Information Design Model (HIDeM). This model proposes a series of principles and activities aimed specifically at designing data intensive applications for health informatics and that considers diverse organizational contexts. The HIDeM model aims to:

- Follow a collaborative and translational design process appropriate for the design of digital products aimed at scientific research;
- Identify the specific visualization needs through the use of proxy tasks and proxy values;
- Map scientific sense-making into concrete tasks that researchers need to accomplish through various service design activities.
This model is the result of joint work with the experts from Northeastern University and the Sabeti Laboratory at the Broad Institute of MIT and Harvard, focusing on Infectious Disease (ID) and evolutionary biology. It was through this collaboration in a multitude of projects that the original model was proposed. Some of the projects that benefited from this collaboration include: the creation of a historic and near real-time geographical visualization of confirmed cases of viral infections in the greater Boston region, as well as a data visualization tool to explore relationships between disease outcomes for infected patients and the genetic variations of pathogens. The HIDE draws from these experiences and borrows from existing methodologies that will be discussed in the background section.

2. Theoretical background

2.1 Issues with visualization and interface design in life sciences

Historically, life sciences research has relied on its own domain experts to create GUIs and visualizations for data analysis (Carroll et al., 2014; Cashman et al., 2018). However, the increase in the volume of available data poses great challenges in the fields of medical research, epidemiology, genomics and cancer research, and drug development, among others. The nature of these challenges ranges from how to effectively visualize massive datasets to how to interact with complex, multidimensional data in a clear and structured fashion. For example, recent innovative approaches utilize web-based technologies that can be as functional as the traditional locally installed tools (Argimón et al. 2016; Neher et al. 2017). In the field of infectious disease epidemiology, new trends in analysis tools are geared toward social network visualizations, molecular epidemiology, and GIS (Carroll et al. 2014).

The massive datasets also present new opportunities for cross-domain collaborations between designers and scientific researchers. However, there are not many studies that address the need to provide domain experts in life sciences with the right processes for collaboration to design user-centered, accessible, and interoperable applications. Relevant work includes a Health Information Technology Acceptance Model aimed at consumers of health data (Kim & Park, 2012), the description of a process for the development of a user-centered development of a mobile medical application (Peischl et al., 2015), and several sense-making mapping activities specifically designed for scientists (Mirel et al., 2011; Mirel & Görg, 2014). However, these approaches are either not aimed at domain experts or they focus on a subset of activities. There is a need to enrich the collaborative process with a transdisciplinary and holistic approach, such as service design perspective (Stickdorn, 2012), to enhance the structure of collaboration and align relevant stakeholders with a shared goal.

A holistic design approach typically prioritizes users’ needs, often depicted as customer journey. Little is known about the end to end process of scientists’ reasoning in solving complex analytical problems (Mirel & Görg, 2014). As Mirel et al. (2011) argue, “recognizing and identifying the actions and steps that drive discovery through visual analytics is an unexplored problem in the biomedical sciences.” However, a core activity such as Cognitive Task Analysis (Pirolli & Card, 2005), as adapted by Mirel & Görg (2014), can help map specific
sense-making into tasks. This activity achieves two important goals. The first is to design for the right questions by properly understanding the analytical problem. The second is to provide a holistic view of the entire research process, so that tools that are created for a specific task can be successfully integrated with the rest of the activities. Figure 1 shows an example of the adapted model proposed by Mirel & Görg (2014). A successful task analysis is carried out by observing the user in their habitual environment and noting the different tasks in which they engage to solve a particular analytical problem.

While understanding the sense-making of the users is important, creating these digital applications requires going a step further. It is also necessary to ensure that the design proposals address solutions for the right analytical questions. This is because the design of the tool should not only respect existing mental models, but also correctly support concrete use cases. Meyer & Fisher (2018) introduced the concept of “operationalization,” which is defined as “the process of identifying tasks to be performed over the dataset that are a reasonable approximation of high-level question of interest.” This activity translates analytical questions into tangible user tasks. For example, if the high-level question is, “which strains of a virus are more dangerous?”, then the first task to operationalize would be to “identify viral danger.” This is naturally a rather fuzzy term. With this in mind a good proxy for this task could be based on “identifying regions of the viral genome known to be responsible for infecting the cells of a host.” This deductive process allows designers to hone into concrete tasks that can be represented with visualizations or through user interaction.

![Figure 1. Diagram (tasks, data, software-supported methods) proposed by Mirel & Görg (2014).](image-url)
In addition to a good task analysis and operationalization, there are special considerations that must be considered when designing GUIs and visualizations in health informatics. We have classified them in three categories: working memory and cognition, visual and statistical accuracy, and transparency in showing data transformations and methods. To ensure the designs address these considerations, Jackson et al. (2012) suggests the use of sketches, prototypes, and discussion as a way to evaluate scientific visualizations. The goal of these prototypes is to jointly evaluate the strengths and shortcomings of the proposed interactions and visualizations, which can be considered as coincident boundary objects – see section 3.2.4 for a definition– (Star & Griesemer, 1989) that can help align the goals of designers, domain experts, and/or users. Additional support for guiding the creation of the visualizations has historically been found in the Computational Information Design process coined by Fry (2004), who examined the different design, information, and computation pedagogies, and combined them into a “method for the exploration, analysis, and representation of complex data.” While HIDE uses these sources as a conceptual basis, it also focuses on adding concepts from sensemaking theory and involves non-designers in the process.

2.2 Existing methods and theories
Since several methods and theories coalesce throughout the HIDE, we present a diagram that summarizes how they relate and align with each other in Figure 2. The visualization tasks will be explicitly described by asking the researchers to verbalize their analytical interests and to use proxies in a deductive process, as well as carefully looking at the data necessary for these analyses. Additionally, a holistic view of the research process based in observations and a service design perspective will be drawn out to provide seamless interoperability of the tool with existing tools and workflows. Typical service design activities include contextual inquiries and interviews for the purpose of incorporating diverse stakeholders’ perspectives, journey ideation and task mapping, and several forms of prototyping that aim to shape the holistic flow of activities (Stickdorn et al., 2018). Next, critique sessions between visualization specialists, scientists, and other domain experts will be held using sketches and prototypes of the visualizations, with the goal of constantly evaluating if the design efforts are aligned in the right direction in terms of usefulness, practicality, and realistic implementation. Finally, the design of the tool will follow a process that is based on the guidelines from the Computational Information Design theory.
3. The HIDeM model

3.1 Model overview

In this section, we describe the HIDeM model in detail and demonstrate how it has been applied in a real-world setting. This model attempts to answer the question, “what are the activities a design process for health informatics should consider given the specific differences of design contexts?” It is designed for developing tools for health informatics in either quick production settings, commonly found in the private sector, or for slower, more research-oriented settings such as academic environments. The model posits that it is more important that certain activities and considerations always be present if the visualization is intended for the analysis of complex Genomic Data or Real-World Data. The model is composed of four principles and nine activities separated into two classes (Figure 3). Many of these activities are tied together by feedback loops and iterative cycles.

Figure 2. Methodologies that are used in the HIDeM.

Figure 3. Diagram of the activities in the HIDeM.
While the terms “design” and “development” might seem vague, we wanted to use concepts that are relatable for the design practitioner as in many real-world scenarios there is a design stage that includes research and analysis activities, and a development phase that includes synthesis and technical implementation activities (Salah et al., 2014). As mentioned, the design activities conform an “analysis cycle,” in which the end goal is to create a comprehensive understanding of the specific sense-making models of prospective users of the tool or visualization, understand the context of use, and come up with a comprehensive list of data objects, tasks, and general information hierarchy. On the other hand, the “development activities” refer to steps where the tasks are a) to streamline and validate the design solutions via discussions and usability testing, and b) to implement the design solutions into a working tool using software development resources.

3.2 The principles of the model
The model is comprised of four guiding principles that help create the right context for the development of all the design activities. These principles can be understood as precursors to successfully embed the model in the organization, and they should be discussed with the entire team that may comprise designers, domain experts, developers, and other stakeholders.

3.2.1 Designers as mediators for discovery
This first principle helps designers frame their role in the organization. This is because while it is not uncommon to think of design as a “problem solving” discipline, in this particular case, design’s role is to establish an environment for the domain experts in which novel, user-centered ideas can offer solutions that can be embedded within existing workflows with as little pain as possible. The goal is not necessarily to make scientific breakthroughs, but to discover and optimize the processes and information flows that enable them. In other words, designers can “create conditions for systems to emerge” (Dubberly, 2019) in which domain experts can be active participants in the design process.

3.2.2 A culture of transdisciplinary collaboration
The second principle seeks a tighter communication between design professionals and domain experts. Collaboration between scientists is not a new idea, but collaboration between scientific domain experts and designers—aside from the established practice of scientific illustration—is not commonly seen. Since it is key for the design practitioner to understand the mental models and sense-making patterns of the user, there must be an open and transparent collaboration between the two roles. This implies that the scientific environment where the designer becomes embedded must be respectful of the practice of design, and at the very least be open to learning more about design processes and methodologies. Conversely, designers must be open to learning more about the science and conventions within the specific work setting.
3.2.3 Ignorance as a precursor for learning

Fischer (2000) stated that “complex design problems require more knowledge than any one single person can possess, and the knowledge relevant to a problem is often distributed and controversial.” Design problems in the space of health informatics are no exception, and the opinions of different domain experts can be indeed controversial at times due to the exploratory and novel nature of this practice. Fischer goes on to suggest that “knowledge workers and designers have to engage in creative activities to cope with unforeseen complexities of everyday, real-world tasks.” For instance, a discussion referencing some sketches can be a great opportunity for the domain expert to learn more about the design process, jargon, and the way that sketches help model workflows; at the same time, the designer can learn about specific terminology, scientific concepts, and nuances in the domain expert's research.

This can be an intimidating situation for a design practitioner who is not accustomed to the data, environment, context, or specific scientific domain. However, the HIDEM model proposes that they must embrace their expert knowledge in design and their ignorance in the scientific field equally. In other words, designers must embrace the often-awkward situations that stem from stepping into uncharted territory and should think of this as a positive indicator that they are moving in the right direction. Similarly, domain experts should be open to learning more about design methodologies and common practices and should embrace the process by actively engaging in design activities.

3.2.4 Boundary objects at the core of the team

The last principle addresses a practical solution to a complex problem: how can all parties share knowledge and discuss possible solutions to identified problems? We propose that boundary objects, which are defined as “both plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites (Star & Griesemer, 1989)” play a central role in this process. Coincident boundary objects in the form of sketches, prototypes, and diagrams are essential for discussions about the project goals and the pertinence of the solutions. Another example of a boundary object of the class “ideal type” is the “tasks-objects-software supported methods” diagram by Mirel & Görg (2014) and discussed in the background section. Based on experience, we have concluded that boundary objects are extremely effective at creating shared understanding among interdisciplinary team members.

To summarize, these four principles aim to lay the groundwork and to discuss some of the preexisting conditions that we believe are necessary for designers and domain experts to effectively and positively collaborate in this challenging environment.

3.3 The activities of the model

The model is composed of nine activities which follow a loosely defined sequence which are tied together by feedback loops. Naturally, the model allows for some tasks to happen
Simultaneously. Figure 4 provides descriptions, deliverables, and ways to evaluate each activity.

**Health Information Design Model (HIDeM): activities**

<table>
<thead>
<tr>
<th>activity</th>
<th>description</th>
<th>deliverable</th>
<th>evaluation</th>
<th>[owner] [collaborators]</th>
</tr>
</thead>
<tbody>
<tr>
<td>operationalize</td>
<td>take hypotheses or goals and break them into concrete tasks using operationalization techniques (proxy values and proxy tasks)</td>
<td>a diagram of tasks – objects – software supported methods</td>
<td>Interview current or potential users and probe for tasks and activities</td>
<td>[designer] (current or prospective users, domain experts)</td>
</tr>
<tr>
<td>cognitive tasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>review literature</td>
<td>A scholarly research of previous tools and techniques, and user profiling</td>
<td>a report on current techniques, technologies, patterns and similar tools/software</td>
<td></td>
<td>[designer]</td>
</tr>
<tr>
<td>analyze interoperability</td>
<td>model how the data and the user behave before and after using the tool/software</td>
<td>user and data journeys that explicitly describe the formats and activities that happen outside of the tool</td>
<td></td>
<td>[designer]</td>
</tr>
<tr>
<td>architect information</td>
<td>create content and data inventories and hierarchies, and propose a navigation</td>
<td>an information architecture, a site map or other hierarchical representation of data and objects</td>
<td>cognitive walkthrough, card sorting, usability testing</td>
<td>[designer]</td>
</tr>
<tr>
<td>ideate</td>
<td>brainstorm and sketch possible visualizations, taking into account user patterns and data structures</td>
<td>sketches, paper prototypes, digital representations using no or synthetic data</td>
<td>evaluate navigation, clarity of data mappings, non-traditional data representations</td>
<td>[designer and domain experts co-own this activity]</td>
</tr>
<tr>
<td>explore paradigms</td>
<td>evaluate visual paradigms with dummy data or a representative subset</td>
<td>summary of test results with scientists, developers and other stakeholders</td>
<td>cognitive performance, statistical support of visual relationships and limitations</td>
<td>[designer] (domain experts)</td>
</tr>
<tr>
<td>test prototypes</td>
<td>show a tool with limited functionality to end users and evaluate behavior and task completion</td>
<td>results of usability testing, task completion and possible new features</td>
<td>evaluate whether the tool supports the tasks, user interaction and possible additional features</td>
<td>[designer and some domain experts may co-own the design of the tests]</td>
</tr>
<tr>
<td>finalize tool</td>
<td>eliminate major bugs, ensure that core data visualisation and GUIs features are fully operational</td>
<td>production-ready tool/software with core features and capabilities</td>
<td></td>
<td>[software engineer] (designer)</td>
</tr>
<tr>
<td>document algorithms</td>
<td>provide clear explanations on the statistical and object transformations of the data</td>
<td>documentation files, markdown or website</td>
<td></td>
<td>[software engineer; data scientists] (designer)</td>
</tr>
</tbody>
</table>

**Figure 4. Activities with deliverables, evaluations, and participants.**

We provide examples for all these activities in the following section, so we will not discuss them in detail here. However, we would like to highlight a couple of key activities that we think are not common for design practitioners:

### 3.3.1 Operationalize cognitive tasks

An activity that many designers may not be familiarized with conducting—particularly in a structured fashion—is the operationalization of cognitive tasks. However, this activity is key in driving the creation of concrete visual artifacts that support the goals of the domain experts and should be used as the core insight that guides the rest of the project. The theory that drives this activity was discussed in the background section and is based on the work of Meyer & Fisher (2018) and Mirel et al. (2011). Operationalization considers four main components: data, tasks, stakeholders, and visualizations.
3.3.2 Review literature

Literature reviews (LR) are a common activity in academic and research environments that help build knowledge about a particular field and provide an understanding of the current state of the art regarding a particular topic. While design practitioners generally don’t conduct them, there are many benefits that stem from a thorough LR. First, it provides designers with domain-specific understanding that would otherwise be very slow to attain on their own. It also gives insight into the existing mental models of the users and the existing usability patterns that are present when interfacing with similar visualizations and data exploration tools. Meta-analyses provide a quick and effective way to understand the current landscape of digital tools, existing issues, and needs.

For instance, in one of our projects, our LR yielded a systematic review of several visualization tools for infectious disease epidemiology (Carroll et al., 2014) that stated, “The architecture of the tools was inconsistently described, and few tools in the review discussed the incorporation of usability studies or plans for dissemination. [...] Existing tools offer a range of features and functions that allow users to explore, analyze, and visualize their data, but the tools are often for siloed applications.”

Finally, in some cases, designers will be lucky enough to find studies that typify users, roles, and tasks, which is a rare and invaluable finding that fast tracks the research phase. For instance, in a non-academic project one of the authors was tasked with developing a digital tool for a user in the domain of clinical trials known as a Clinical Research Coordinator. The LR allowed him to find a study with 456 participants (Owens Pickle et al., 2017) that listed common roles and tasks, motivations, reported salaries, and a wealth of additional information about this target user. All of these benefits allow designers to gather a comprehensive understanding of the systems that domain experts are already using and, possibly, of the potential users themselves.

3.3.3 Analyze interoperability

The previously cited review on infectious disease tools (Carroll et al., 2014) stated that, “interoperability of tools and the importance of data sharing and integration were important goals that should factor into the design of visualization tools” and that, “commonly cited barriers to widespread adoption (of these tools) included lack of organizational support, access issues, and misconceptions about tool use.” We believe, in general, that is the case for many of the tools that are designed both in academia and the private industry. While the reasons may be diverse, we posit that for adequate adoption and use of a newly designed tool the designer should conduct a thorough analysis of the software ecosystem in which the new tool is meant to be embedded. They should create an inventory of tools, the different data transmission formats that the tools support, as well as diagrams that outline typical workflow and transitions between one tool or activity and the next.

This is a powerful activity that aims to reduce unnecessary data transformations, encourages the use of conventional formats, reduce the barriers for adoption, probes into existing
mental models that are shared across tools, and allows for faster and seamless deliverability of results and outputs. For example, during contextual inquiry activities with hospital staff we learned that it was not uncommon for them to have to quickly hop between the EHR system, paper notes, a piece of software known as an Electronic Data Capture (EDC) system, and a Clinical Trials Management System (CTMS). Identifying the “software fatigue” and multitude of data formats that were represented in these daily transactions represented an invaluable insight and a great opportunity for optimization.

3.3.4 Document algorithms
Carroll et al. (2014) also found that a “lack of trust and misconceptions about the use of the tool, may also prevent adoption,” and that many users felt that, “many existing tools were too complex and had a substantial learning curve.” With this in mind, one way to increase trust in the system is to create clear and transparent documentation so that users understand the data transformations that happen “behind the curtain,” and to, whenever possible, show snippets of information directly in the application that help explain the results on the screen. In addition to these benefits, documentation may be necessary to make sure that the tool is compliant with certain regulations or that it meets certain quality controls. Therefore, designers should assist in the creation of documentation, as they were directly involved with creating the GUIs and the visualizations that are at the core of the solution.

4. Use case: NX4 – a visualization tool for viral genomics

4.1 Background and overview
In this section, we discuss the real-life use case of the HIDE M on a scientific research project. The designed tool was released as an open-source software for genomic analysis called NX4 (https://www.nx4.io/) and was also validated through a peer-reviewed publication (Solano-Roman, 2019). The tool provides a novel visualization approach for the joint analysis of thousands of genomic sequences, also known as Multiple Sequence Analysis (MSA).

Historically MSA tools have used the same visualization paradigm for decades: for example the notably performant tool “AliView” (Larsson, 2014) uses the same visualization paradigm as “SEA VIEW” (Galtier et al., 1996), a tool that was released 18 years before. However, the amount and complexity of data has exponentially increased ever since.

With this in mind, while we were carrying research projects (both as part of daily work and also some under IRB approval), we started to conduct some contextual inquiry exercises with seven domain experts at the Sabeti Laboratory at Broad Institute of MIT and Harvard. During these exercises we discovered a series of potential opportunities for improvement of the domain expert’s workflow as it relates to MSA activities. Particularly, some researchers were interested in better understanding the relationship between different viral genomes and their impact on the outcomes for patients. An archetypical question would be: “Can we predict which patients will face worse outcomes solely based on the genetic makeup of the
virus that infected them?” Based on this, we designed a tool to support a specific portion of the research process in which MSA is conducted.

4.2 Activities
Examples of the activities in the model are described in detail below.

4.2.1 Operationalize cognitive tasks
Starting from the original needs we continued to probe the users following the operationalization methods to create diagrams such as these that describe the task, data, and methods that support them (Figure 5).

![Figure 5. Tasks, data, software-supported methods diagram for NX4](image)

4.2.2 Literature review
The literature review that we concurrently carried out aligned what the domain experts were telling us: for instance, we learned that there has been an increasing trend to design web-based tools instead of desktop-based applications (Yachdav et al., 2016) and that novel visualizations with radically different paradigms were increasingly being used in genomic research (Asnicar et al., 2015; Ribeiro-Gonçalves et al., 2016). With this knowledge in mind, we decided to build a web-based visualization because we could not only harness the power of modern, JavaScript visualization libraries, but we could also reduce the barrier to adoption by not asking users to install another program on their computers.
4.2.3 Analyze interoperability

Next, we conducted an interoperability analysis with several potential users of the tool and created a “user and data journey” diagram as an output (Figure 6) that exposed the multitude of data, software, and activities that revolved around this type of genomic analysis. The analysis was conducted during interviews and contextual inquiries that occurred regularly at the laboratory, where we could observe and interact with the users in their natural environment. We eventually learned that users “hop” between many different tools that required installation directly on each computer they use and that this causes frustration and fatigue.

<table>
<thead>
<tr>
<th>touchpoints</th>
<th>tools</th>
<th>data</th>
</tr>
</thead>
<tbody>
<tr>
<td>get samples from outbreak extract viral samples from infected patients</td>
<td>on-site sample collection kits</td>
<td>biological samples (e.g. saliva, blood), might be from different locations in the body</td>
</tr>
<tr>
<td>sequence the data a series of chemical and computational processes to obtain genomic sequences</td>
<td>biochemical preparations, genomic sequencer</td>
<td>genomic sequences in plain text files</td>
</tr>
<tr>
<td>combine with clinical data on tables parse, organize and match the clinical and the genomic data of every patient</td>
<td>spreadsheet software to combine epidemiological tables with clinical data</td>
<td>epidemiological tables lookup tables and ids that link to genetic sequences stored elsewhere</td>
</tr>
<tr>
<td>generate phylogenetic trees create a tree to look at genetic divergence, viral subtypes and outbreak onset</td>
<td>phylogenetic tree generator software</td>
<td>nexus, newick or other phylogenetic tree formats</td>
</tr>
<tr>
<td>generate sequence alignment files run all the sequences through an alignment software</td>
<td>multiple sequence alignment generator software</td>
<td>alignment fasta, clustal g, g, msf or other alignment file formats</td>
</tr>
<tr>
<td>look at divergence and variations search for regions of genetic mutations and explore functional structures</td>
<td>multiple sequence alignment visualization tool</td>
<td>alignment file, data on functional regions of the genome, derived statistics</td>
</tr>
<tr>
<td>establish relational hypotheses compare alignment data with clinical data and hypothesize about causal relationships</td>
<td>phylogenetic tree viewers with tip coloring for categorical data</td>
<td>clinical, genomic, statistical</td>
</tr>
<tr>
<td>conduct experiments in the wet lab verify or discard hypotheses</td>
<td>laboratory tools and techniques</td>
<td>clinical, genomic, statistical, biological</td>
</tr>
</tbody>
</table>

Figure 6. User and data journey for interoperability analysis.

Interoperability analyses recognize the service ecosystem around the users and help us design more seamless solutions or think about integrations between systems.

4.2.4 Architect the information, ideate, explore prototypes, and evaluate paradigms

These four activities are tightly connected by several feedback loops and rarely follow a linear pattern (see Figure 3). Instead, the activities inform each other, and this is why we’re discussing them together in this section. Several information architecture (IA) sessions and
exercises were conducted using well-established methods and practices, including card sorting exercises and validation via wireframes, such as the one in Figure 7. This wireframe portrayed “dummy” graphics and was used to evaluate the IA during user-testing sessions.

![Figure 7. Example of one of the wireframes used for information architecture and tasks evaluation.](image)

Several other sketches used for discussion and evaluation of visualization paradigms and interaction techniques are shown in Figure 8a (paper format) and Figure 8b (digital, data-driven prototypes).

### 4.2.5 Finalize tool and document algorithms

This activity is carried out in close collaboration with the software developers to ensure that the final result still supports the tasks that were identified during the first activity in this model. It’s not uncommon to reach the software development stage and find technical constraints that force the design and interaction of the application or of the visualization to change. The designer should be able to provide support in these situations and find pragmatic, yet effective solutions to these problems. Since, as we discussed before, a notable barrier for adoption can be the lack of documentation, we decided in this case to create a dedicated documentation site using the service GitBook (https://nx4.gitbook.io/documentation/) The documentation explains the aim of the tool, the formats it supports, and provides a use guide (Figure 8c).
Additionally, the algorithms that drive the functionality of the tool were discussed in the peer-reviewed article that was written shortly after completing the tool. For non-academic settings we still recommend documenting all relevant data transformations in help articles (e.g., Zendesk knowledge hubs).

A final version of the tool can be seen in Figure 9. The open-source tool provides a color-blind friendly interface and allows the visualization of thousands of genomic samples without the need for vertical or horizontal scrolling. The user can also export sample IDs for further exploration.

Figure 9. Screenshot of the resulting tool.
5. Conclusion

The current revolution in life sciences and healthcare will continue to force diverse professionals to come together to tackle the information challenges of tomorrow. Complex data analyses will continue to be commonplace in these fields, where experts must attack problems from a multitude of angles, and carefully interpret data to answer intricate questions. We posit that, given the critical needs in life sciences, it is imperative that more designers collaborate and become involved in scientific research. As industries and academia continue to pay more attention to their users and attempt to improve their interaction with digital products, designers can play a major role in facilitating better ways to visually explore data, by creating tools that can support the research activities with the needs of the user at the core of the design process.

The HIDeM model is novel in that it is the first systematic design process model developed specifically to facilitate the collaboration of designers and domain experts in life science research, an area where a participatory approach to design has been underutilized. The model is also unique and practical in that it has been optimized to work in both academic and corporate settings, which are notorious for having very different paces and motivators, while industry-academia collaboration is common in life science research. Our proposed model supports such collaboration and contributes to other process models by providing a flexible framework that focuses on shared principles and activities to promote fundamental understanding across experts from different disciplines, as opposed to following rigid step-by-step tasks. HIDeM also recognizes the nonlinear nature of scientific research and has been validated through its use in diverse settings with varying team sizes, software development paces, and stakeholders. For instance, one of the authors now works at a fast-paced, health informatics startup company where the model has been applied to a variety of data visualization and data-intensive GUI design projects in the RWD domain, a distinctly different environment to where the model was originally conceptualized.

Another major contribution of HIDeM is that it recognizes the creative abilities and contributions of domain experts by encouraging them to transfer knowledge and participate in the design process as co-creators. For example, the scientists who participated in the development of NX4 obtained tools to communicate their ideas. It empowered them to be active drivers of visualization solutions, which lead to an open mindset for team work with designers and a positive impact on the culture of collaboration. With its collective context, HIDeM can also contribute to improving models like the Computational Information Design process by placing the entire team, not just the end-user or the designer, at the forefront of the creation of solutions.

The active role of domain experts also indicates the need for expanding the role of designers in life sciences and other areas where complex data is involved. Traditionally, designing with data has been seen as an activity of visualization for the purpose of effective communication, often utilizing a single designer’s perspective. However, this study reveals that working with complex data now requires cross-domain collaboration with multiple
agents, and the outcome is often an interactive system that invites more participation. Designers working on complex data now need to consider their designs as a dynamic interface to support research and further creation. Therefore, a solid yet flexible process, holistic understanding about teamwork context, and diverse activities to mediate stakeholders’ participation will increasingly become essential parts of an information designer’s work.

We have found that fostering a culture of respect and trust is required for the successful implementation of the model. The HIDeM model posits the need for cross-domain collaborations as part of the organizational culture to solve the complex problems inherent to the field. In other words, it is paramount that the research groups or organizations that want to implement HIDeM are open to an active collaborative culture between domain experts and designers who are oftentimes not experts in specific scientific fields, thus effectively establishing a symmetry of ignorance relationship that acknowledges that each party is ignorant of the other’s domain of expertise. (Fischer, 2000). This sense of perceived ignorance, coupled with honest vulnerability, are the precursors for learning about the needs of researchers, and are necessary ingredients for the development of successful digital products.

In future studies, we plan to draw upon information from these models to create toolkits, including templates and activity guidelines, that designers and domain experts can easily access and use. HIDeM can also be applied to other data-intensive domains like clinical research, population health, and health economics. This is because visual analytics in data-intensive domains are generally bound to the same human sensemaking behaviors and require that domain experts collaborate with their design counterparts to create effective visualizations and interfaces. In that sense, the contributions of the model could also benefit designers working in fields other than healthcare. The essentiality of promoting interdisciplinary collaboration and a culture of trust and respect across domains is generally applicable to any design practices that involve complex data generated in specialized domains. In this respect, the HIDeM model provides a solid foundation and examples for its initial implementation.

6. References


About the Authors:

**Antonio Solano-Roman** is a Senior UX Designer at Flatiron Health, a US-based oncology informatics company. He holds an MFA in Information Design from Northeastern University and has worked at the Broad Institute of MIT and Harvard, the Fred Hutch, and JAX Laboratory.

**Miso Kim, Ph.D.** is an assistant professor in the Department of Art and Design at Northeastern University, where she explores the humanist framework of service design with a focus on dignity, autonomy, and participation.