Data and UX Heuristics

What is this?

These set of heuristics will help you to better apply good design practices when working with data driven projects.

They may not always apply and you might not always have the answers but they are intended to guide you to a better outcome.

Who is this for?

Primary Audience

UX practitioners who contribute to data driven projects.

Secondary Audience

Anyone working on a data driven project who need to understand why UX is relevant such as data stewards, data scientists, product owners, data curators etc..

When to use?

Earlier is better!

Appropriate times such as:

- When planning a data project
- At the kick-off meeting to ensure all stakeholders are present and all the right questions are being asked
- At UX strategy meetings for data project planning
- When stakeholders need to understand why UX is relevant for a data project

How to use?

We provide a set of 9 independent heuristics for you to consider.

You don't have to provide an answers to all these heuristics but it is valuable to think about them early on in your data project.

We have provided a handy checklist on the next slide to help you cover all the heuristics.

Each heuristic comes with a description, example and activity.

Data and UX Heuristics

Pistoia Alliance

There are 9 guiding heuristics which will support you.

Understand the problem
we are trying to solve

Identify relevant data
and metadata

Design for greater longevity
and reuse of data

Use known data standards
and industry guidelines

Connect people to bridge between
stakeholders for data projects

Design to capture the context
of the data as metadata

Understand the entirety
of the data journey

Design as close to real data
and metadata as possible

Establish a consistent
flow for user feedback

Understand the problem we are trying to solve

Description

When tasked to build a solution, we are often given requirements and the prescribed solution. However, we often find that stakeholders do not or cannot fully articulate the problem, and the final solution, while perfect to specs, does not actually solve their issues. UX'ers can mitigate this by doing user research to get at the heart of the actual problem before design and development.

Examples

- Scientist requests a new UI element to be added without understanding the underlying problem. Although this was added, it did not solve the scientific problem for the user.
- An inflexible design that limits the exploration of datasets.

Activity

- Map how data is used to identify all relevant stakeholders [See Heuristic: Understand the entirety of the data journey]
- Conduct user research and explore the problem space using suggested methods such as Design Thinking or the methods at **uxls.org**.
- Learn about your users requirements so you can identify and create your **personas** quickly
- Validate the problem understanding and potential solutions with the stakeholders using mock-ups and/or prototypes.
- Iterate to improve your understanding.
- Be open to serendipity for new insights.

Pitfalls of ignoring this heuristic

- Rushing into a solution too soon might mean you design the wrong solution.

Use known standards and industry guidelines

Description

Whenever possible, known standards and industry guidelines e.g. FAIR principles should be employed in your design. Rather than creating your own controlled vocabulary or ontology, use community, industry or company standards to encourage interoperability so the data and metadata is more likely to be used by others. Not following industry standards might affect data quality and that data will require more downstream work as more time needs to be spent curating and unifying it. See reference by **Schultz et al 2019** for further information.

Examples

Drug Names are notoriously difficult to identify due to their numerous brand names. Enabling auto-suggest driven by an industry standard controlled vocabulary will support reuse and identification of what the scientist intended.

Activity

- Collaborate with data and subject matter experts to familiarize on data and metadata conventions and standards in your organization and the wider community. A useful resource to consult is **FAIRsharing.org**
- Always review your designs from a data standard perspective:
  - Are your designs using known standards? Have you had your design reviewed by a data expert?
  - Are you supporting users in data choices e.g. auto-suggest for synonyms against a data standard?
  - Push for powerful, flexible querying (semantic tagging) to accommodate the diversity of your user groups.

Understand the entirety of the data journey

Description

Data can be seen to move across various phases from the time when it is generated, stored and transformed or extracted for analysis. Additional steps include curation for quality control or standardisation, integration for business intelligence, and visualisation for exploration, reporting and/or gaining insights.

In each step of this journey, UX designers will support the various stakeholders, each with their own needs, requirements, activities and understanding of the data, by helping them to achieve their goals.

Examples

Different stakeholders deal with the data in different points in time and will focus on specific activities. A software engineer might deal with how consistently the data is stored and moved across platforms and/or APIs, a scientific curator might be interested in how accurately the data is annotated, and a researcher or BI analyst might focus on how effectively the data visualisation will support their data analysis.

Activity

- Be prepared to be presented with multiple opinions on what data are the most relevant.
- Map what data helps the stakeholders achieve their goals and/or meet their needs.
- Measure how data is being used.
- Analyse the research of (2), (3) and (4) to communicate with an appropriate tool such as:
  - The data journey (**ref 1**, **ref 2**)
  - The entirety of the evolving data ecosystem (e.g. Service blueprint **ref 3**, **ref 4**, experience maps).
- Gather user feedback iteratively.
- Review your goals and/or design based on new findings.

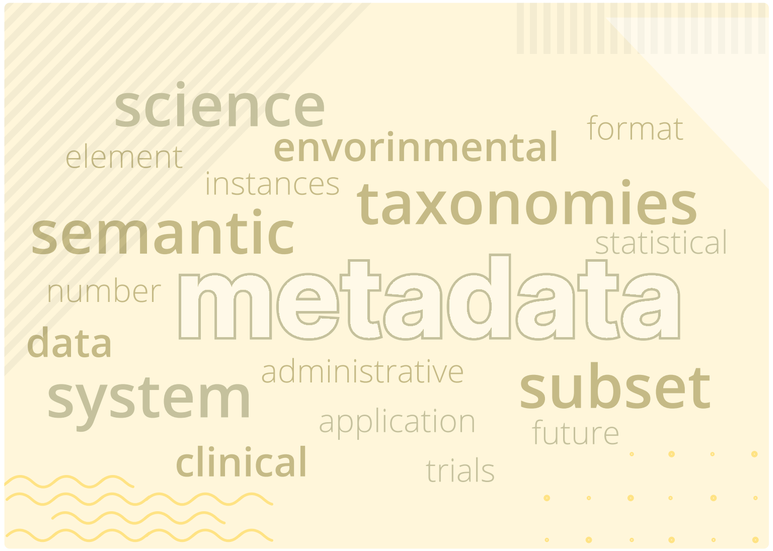
Identify relevant data and metadata

Description

Data serve multiple audiences; different types of users consume the data in the context of specific research or business needs and they might be interested in a subset of data or metadata. It's hard to know upfront what elements of the data will be relevant for which segments and when.

Understand what objective(s) are most important to the scientist. What sources of data and historical knowledge are relevant and available? How can the data and metadata bring value for an immediate purpose and be sufficiently FAIR for future secondary use? How can UX contribute to this balance?

Examples



- Contribute to minimal requirements specification and community/industry standards for particular data type.
- Recommendations for relevant FAIR data and metadata resources.
- UX support for a FAIR evaluation service

Activity

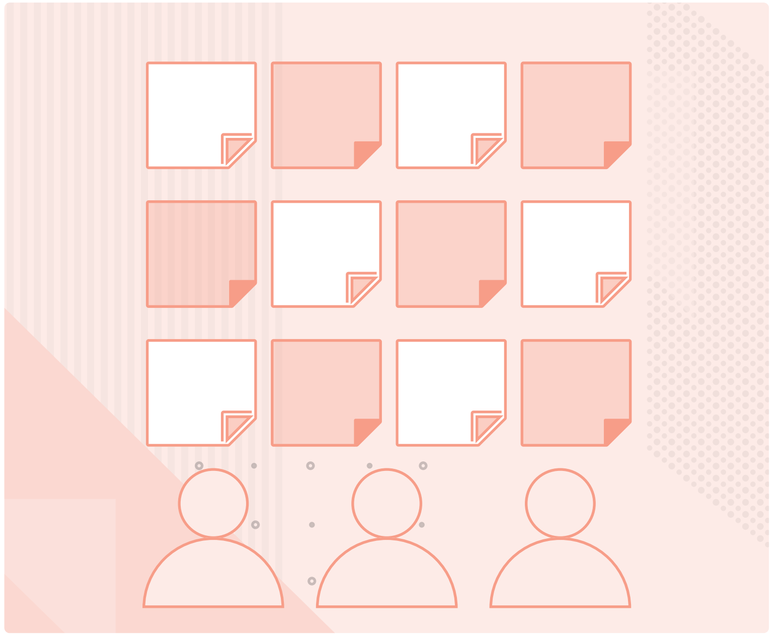
- Conduct user research to understand what data are expected to be most relevant to the objectives.
  - Who are the different types of persona or role in the science team?
- Help to determine what is the immediate, primary purpose for the data?
- Ask what secondary reuse of the data is expected and how FAIR is this data?
- Define what success looks like? Does the relevant data have sufficient quality and completeness to be successful?
  - Identify gaps and improvers
- Provide flexibility to view all data or filter, to display as table or as a visualization, offer color palettes that support users with color-deficient vision.

Connect people to bridge between stakeholders for data projects

Description

The natural cycle of UX projects means that UX designers form a natural “people” bridge between various stakeholders and users. The methodologies we use to kick-off projects such as kick-off workshop, user research to gather insights and usability testing allow us to regularly report back and keep all key groups connected.

Examples



- Running an ideation workshop to drive consensus between data providers and end users.
- Validate the teams assumptions and ideas with users and customers.

Activity

- Kick-off workshops at the start of data projects allow everyone to identify stakeholders and user groups.
- Communication plan to regularly report user research insights to key stakeholders and take new designs to users.
- Usability testing is an invaluable method to keep stakeholders connected.
- Create a contact list of data experts to answer quick questions

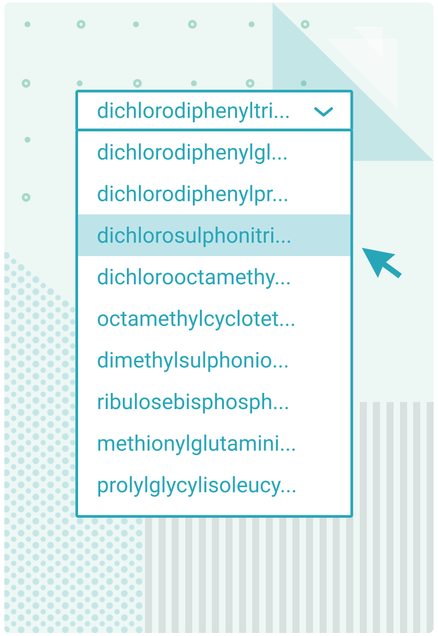
Design as close to real data and metadata as possible

Description

Data representation that is unfamiliar (e.g. Lorem Ipsum) can derail thought processes.
Pay attention that the design follows data standards and specifications, which are understood by the target audience.

Designing with as close to real data and metadata builds trust and confidence with your users.
The design for data presented should be aligned with the mental model of your users.

Examples



- Data displayed is not aligned with existing scientific nomenclature or common data standards.
  - Usage of ambiguous or non-standard terms.
  - Simulated (synthetic) data is used as if it were real data, rather than using “Lorem Ipsum”.
- The user’s mental model is not aligned with the data presented in design deliverables and causes additional explanation effort and an increased cognitive load.
- The design of a control element (e.g. drop-down selection) in a test environment does not scale to the productive environment in terms of size, length or complexity. The proposed design cannot be applied.

Activity

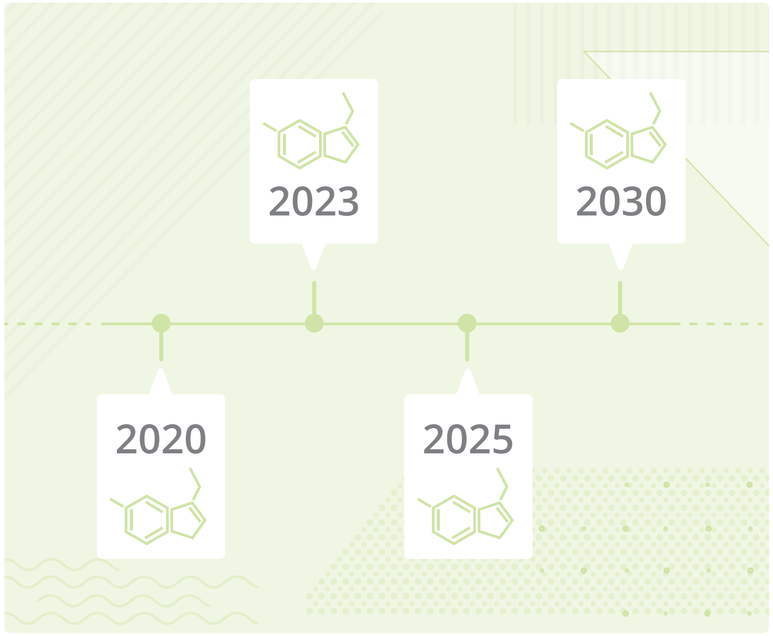
- Whenever possible, ask end users for real-life data sets to use in your design before you start wireframing. Mock data can be fine as long as it follows data conventions & standards.
- Validate design and provide walkthrough design sessions with key core users and data stewards to check for data consistency and data flow.
- Understand data dependencies and include them in the design specifications.
- Ensure the data used within testing is realistic for the real world.
- Consider the license for data usage or data confidentiality e.g. for IP implications of data display in your designs.

Design for greater longevity and reuse of data

Description

Design to improve the experience of data stewards and scientists to increase the longevity and reuse of data. Strong design is required to support the dynamics of data growth and complexity of data types. Reusability of FAIR data and the associated metadata includes links to data usage license, provenance and community or industry standards.

Examples



A design system is applied to infrastructure support for ensuring data is sufficiently FAIR to support immediate primary usage and possible future secondary usage.

Activity

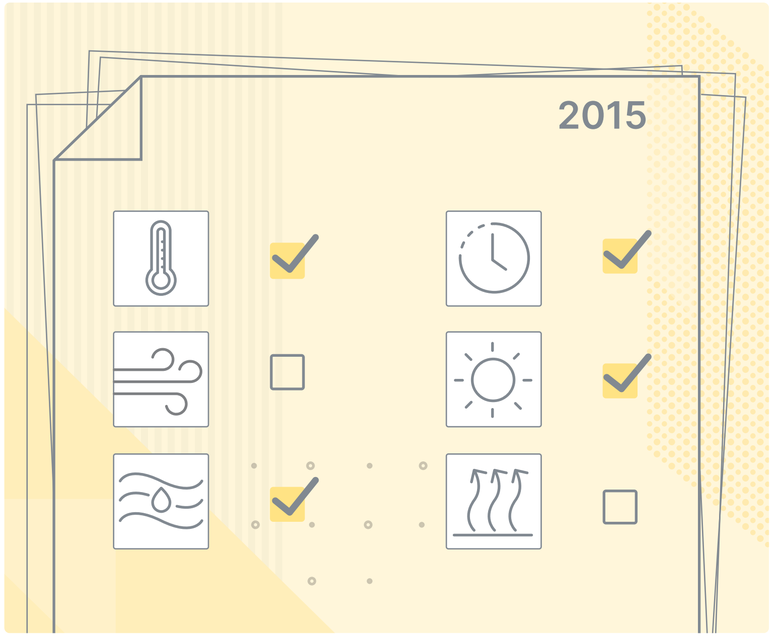
- Determine and apply the local UX design processes to support data management.
- Mature design systems can provide a much more unified platform experience.
- Contribute to making data more FAIR, especially for the three reusability maturity indicators (data usage license, provenance and community or industry standards).
- Help the scientist users to consider the expected longevity of the data.
- Support the associated metadata being available beyond the lifetime of the data.
- Understand how the data has changed over time to future proof for potential changes.
- Always review your designs from a user accessibility perspective for the user.

Design to capture the context of the data as metadata

Description

It is important that the design captures the context of the data as metadata to make scientific sense even when the data is no longer available. Data and metadata should each have identifiers which are global, unique, persistent and resolvable by machines and humans. This will reduce tedious manual data entry and is a driver for making the data Findable, Accessible, Interoperable and Reusable (FAIR) for users and machines.

Examples



- Recreate an experiment that a scientist ran a year ago but is unable to do so because not all the metadata from the instruments was captured at the time.
- Bioinformatics programme to be run on experimental data the versioning and parameters need to be captured as metadata to ensure reproducibility.

Activity

- Use the output of Heuristic (Identify data and relevant metadata) to clarify the metadata.
- Evaluate the data and metadata using the FAIR Maturity Indicator methods described in the **FAIR Toolkit**.
- Design improvements, as necessary for making the data and metadata more FAIR, especially the identifiers.
- Implement the improvements, especially to ensure the use of global, unique, persistent and resolvable identifiers (GUPRI) for data and metadata.
- Check that the metadata are sufficient for future reuse of the data and is also able to support reproduction of the data, **even if no longer available**.
- Seek opportunities to reduce tedious manual data entry.

Establish a consistent flow for user feedback

Description

It is important to establish a consistent user feedback flow that runs throughout the project. The prime method used is usability testing and we encourage you to complement this with other methods such as surveys and web analytics to triangulate your user feedback.

Usability testing a data science project which enables scientific discovery can be considerably challenging and requires exceptional UX thought. Scientists and data specialists are always busy, so it's important to establish a consistent flow and participation for usability testing when new data-driven interfaces are designed.

Examples

Lighting conditions and distance are often overlooked but important considerations when testing. If scientists have to work in a darkened environment, will the data interfaces be easily be readable and reduce eye strain? Likewise, if the scientist has to stand some distance away from the interface, is the data still accurately readable at larger distances?

Activity

- **Test with real data:** Avoid usability tests that have placeholder text (e.g. lorem ipsum). If the specific data is not ready or may change, work with data stewards to find a suitable substitute or mock data set.
- Are you testing with multiple user personas and scenarios? A data interface that works for one persona may not work for another.
- Are you getting input from users to identify unknown potential scenarios you may not be aware of yet?
- To the extent possible, are you testing in the environment the users are working in (or a reasonable facsimile) and and on the device(s) that will be used to view data?
- Are you taking the lab context into account? For example, will a data interface be usable if the person is working at a distance from a display, or if their hands are using lab equipment?

Data and UX Heuristics Versioning

Version 1.0	23.03.2021	First UXLS addressing UX aspects and Heuristics
Version 1.1	28.05.2021	Updates from the FAIR project team.