# S1 BOOKS TO MOVIES AND GAMES

The results for mapping books to movies and games one-to-one are presented in Fig. S1, where we show the book-movie-game triplets as their cover images. In this example, we jointly explore these spaces via the 16 most popular books, ordered left-to-right, top-to-bottom. We see that the first match is "Twilight", "Batman v Superman" and "PUBG" ($\boxed{1}$), and it corresponds to items that are very popular but have an average, even controversial user rating. Note that by average, we mean average among the 500 most popular items. A similar case, albeit slightly less popular, can be seen in the "Lord of the Flies", "Star Wars: Episode I" and "Day Z" triplet ($\boxed{3}$). "HP and the Half-Blood Prince" is both very popular and very favorably rated, which results in it being appropriately matched to the IMDb's #1 rated "Shawshank Redemption" and the "Factorio" game ($\boxed{2}$). In general, we find that this metaphor is easier to interpret than the distance-based mapping and produces meaningful associations between the different media. Now one could explain to a person familiar with movies that "The Da Vinci Code" is the "Iron Man 2" of books.

# S2 TOPOLOGICAL MAPPING

Here, we present a prototype of topological mapping. We use it to map a taxonomy of sciences (taken from Wikidata) to a taxonomy of industries (taken from Eurostat), i.e., between two trees.

**Method.** We define a hierarchical dataset as a directed tree with edges oriented from parents to children, where each vertex of the tree is associated with a vector of attributes. For example, this could be a file system tree, with each vertex having a size and a creation date. Our goal is to map vertices of the data tree to vertices of the concept tree, such that the difference between their attribute vectors is minimal according to some cost function (e.g., MSE). This is the attribute cost from Eq. 2. Additionally, the map $M$ must satisfy the hierarchy constraint: if vertices $x_p, x_d$ in the data tree are connected by a path $(x_p, \dots, x_k, \dots, x_d)$, then their assigned concepts must also be connected by some path $(M(x_p), ..., c_k, ..., M(x_d))$. In other words, the parent-descendant relationship must be preserved.

We solve this problem with the simulated annealing algorithm described in Sec. 5. However, generating random neighboring assignments is no longer trivial due to the hierarchy constraint, so we must make some modifications. The most important modification is that we do not require all of the data vertices to be assigned, so that we can compute mappings for data trees that do not fit topologically within the concept tree. Instead, we add a loss term that penalizes unassigned vertices, making it a soft constraint. This simplifies the search of the solution space since we can move through "partial" solutions to more easily find low-cost regions. And also makes the algorithm much more practical to use with real datasets that don't align well with each other. We initialize the search with a random assignment computed in the following way. The data root is assigned to the concept root. Then, we traverse the data tree breadth-first, and assign each vertex $x_c$ with parent $x_p$, to a random descendant of $M(x_p)$, satisfying the hierarchy constraint. If concept $M(x_p)$ has no unassigned descendants, we leave $x_c$ unassigned. After initialization, we proceed as described in Sec. 5, stepping over random neighboring assignments. The sampling of the neighboring assignments also needs to be adapted. First, we sample a random data vertex $x_c$ that has an assigned parent $x_p$. Then, we assign $x_c$ to a random descendant $c_d$ of $c_p = M(x_p)$, where $c_p$ must be connected to $c_d$ with a path of unassigned concept vertices. If there are no such vertices $c_d$, we simply unassign data vertex $x_c$. Finally, we unassign all descendants of $x_c$ and $c_d$. These procedures guarantee that after changing the assignment of the data vertex $x_c$ we still satisfy the hierarchy constraint.

**Results.** In Fig. S2, we show the mapping of scientific fields onto industries. It matches the topology and one attribute: the publication number to the employee number (size of subfield/industry). We render the resulting mapping as a joint tree, where sciences are colored blue and industries are orange. We use the area of the (half-)circles to represent the number of publications/employees in each field/industry but set a minimal value to prevent the nodes from getting too small. Overall, we are able to satisfy both constraints: the natural sciences are assigned to the largest economic sector – manufacturing ($\boxed{1}$), preserving the size attribute across the two spaces. And the specific natural sciences are mapped to the descendants of manufacturing, e.g., biology becomes electrical manufacturing. Similarly, mathematics and computer science are mapped to subtypes of the information and communication industry. We see that some scientific fields, e.g. information science ($\boxed{2}$), appropriately skip a hierarchy level to better match the attribute. Another interesting case is physics ($\boxed{3}$), which is a very large field and cannot be matched well to any subindustry of manufacturing, because there is no type of manufacturing that is so much larger than the others. Nevertheless, finding good solutions requires us to introduce multiple loss functions and constraints, leading to an inelegant algorithm. We present it here for the sake of completeness, and in the future, we will pursue a specialized algorithm for tree mapping, e.g., searching for assignments hierarchically and providing better initialization by aligning nodes with similar local topology.

# S3 ADDITIONAL IMPLEMENTATION DETAILS

Next, we provide additional implementation details for our distance-based metaphors.

## S3.1 Authors to Words

The data space is an embedding of authors, and the concept space is an embedding of English nouns, both learned from data. We train the author embedding using a self-supervised model similar to word2vec [8] on the VisPubData dataset [5], which contains 3108 papers and 5415 unique authors The model is provided with a pair of authors and predicts whether they are co-authors. Each author is passed to the model as an integer index used to look up a corresponding 32-dimensional embedding vector. Then, a dot product is computed between the vectors, followed by a single sigmoid output unit. The model is trained with a cross-entropy objective to perform the classification, learning an embedding in the process, and achieves 91% accuracy on a held-out validation set.

Our concept space consists of 500 common English nouns, which we passed to a pre-trained word-embedding model. We used the "en_core_web_md" model from the *spaCy* toolkit [4] for the embedding, producing 300-dimensional embedding vectors. Both of the

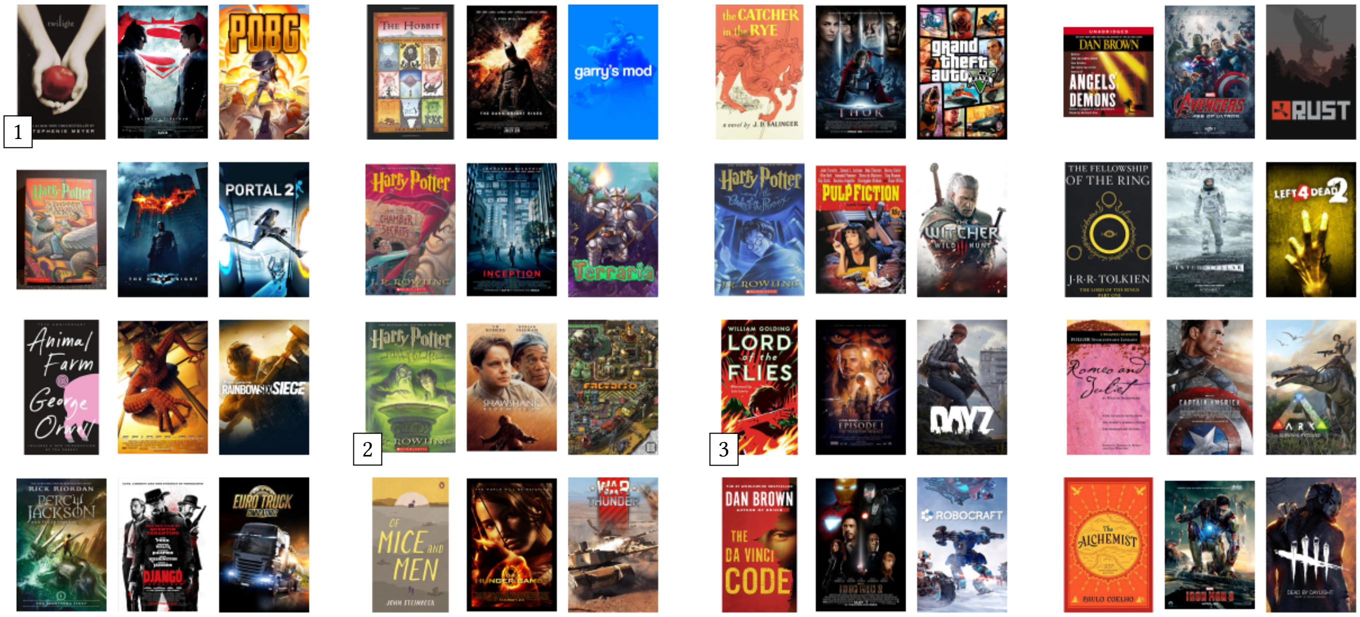


**Figure S1: Attribute-based mapping of books to films and video games. Here we explicitly map the books' rating and vote count attributes to similar film and game attributes. As a result, one can interpret the metaphor more directly since all three items in a book-film-game triplet will have similar rating and popularity. The user can rely on their sparse knowledge across all three domains to learn more about the unknown items. For example, top-left we see the "Twilight", "Batman v Superman", "PUBG" triplet (1), where all three are very popular and have a mediocre rating. While "HP and the Half-Blood Prince", "Shawshank Redemption" and "Factorio" (2) are connected because all three are popular, but are also rated very favorably.**

embedding models use the dot product, and accordingly, we also use the normalized dot product (cosine similarity) as our distance function $d$ for both spaces.

## S3.2 Authors to Cats

For the data space, we obtained publication data from Microsoft Academic, loading 14k authors who have published at CHI and their 19k keywords. This author-keyword matrix underwent a sparse Singular Value Decomposition (SVD) to compute 30-dimensional author embedding vectors for the 100 most frequent authors according to our data. To construct a cat embedding, we took the cat images from the "Dogs-vs-Cats" dataset [6] and trained a model using SimCLR [1], with ResNet18 [3] as the encoder architecture. In SimCLR, the model is trained in a self-supervised fashion to find identical images under random cropping, color distortion and blur transformations. The images are passed through the encoder that constructs a transformation-invariant representation, producing a 256-dimensional embedding vector. Then, cosine similarity is computed between the encoded images to predict which of the images were identical prior to the transformation. Deviating from the original SimCLR approach, we take the feature vector after the projection head because, in our application, we are interested in a space with meaningful distances rather than an information-rich representation for fine-tuning. After training the model, we used a sample of 1000 images and their feature vectors as our concept space.

## S3.3 Authors to Visual styles

We use images generated by StyleGAN2 [7] to anonymize the authors images. The author embedding vectors are learned from a dataset of 1090 SIGGRAPH papers and 2008 authors, using the method from Sec. 5.1. Similarly to our cat metaphor, we perform a mapping between the author embedding vectors and an image embedding of style donor images. However, to construct an image distance metric that emphasizes the style of the image (rather than its content), we make several modifications to our model from Sec. 5.2. We are again using SimCLR, but replace the encoder with a pretrained VGG16 model [9], which has its weights frozen during training. Instead of using the output of the encoder directly, we extract the style information as the activations after the convolutional layers ('conv11', 'conv21', ..., 'conv51') and compute the Gram matrix for each layer's activation (we follow [2] in how the style information is extracted). Concatenated Gram matrices are used as the input to the projection head. The idea is to constrain the encoder to only extract the stylistic features, and train the projection head to map them to a 256-dimensional vector that describes the style. We follow the SimCLR procedure as usual, but use an aggressive cropping setting (10-20% of the image size) to further encourage the encoding of the style and not of the content. The model is trained using a dataset of 11,000 digital art images [10].

Then, we construct a distance-based mapping between the author and style vectors, mapping 100 most frequent authors to a small sample of 16 style images. We deliberately use a small number of style images and allow duplicate assignments to make it easier to distinguish style similarity. Once the style images are assigned, we

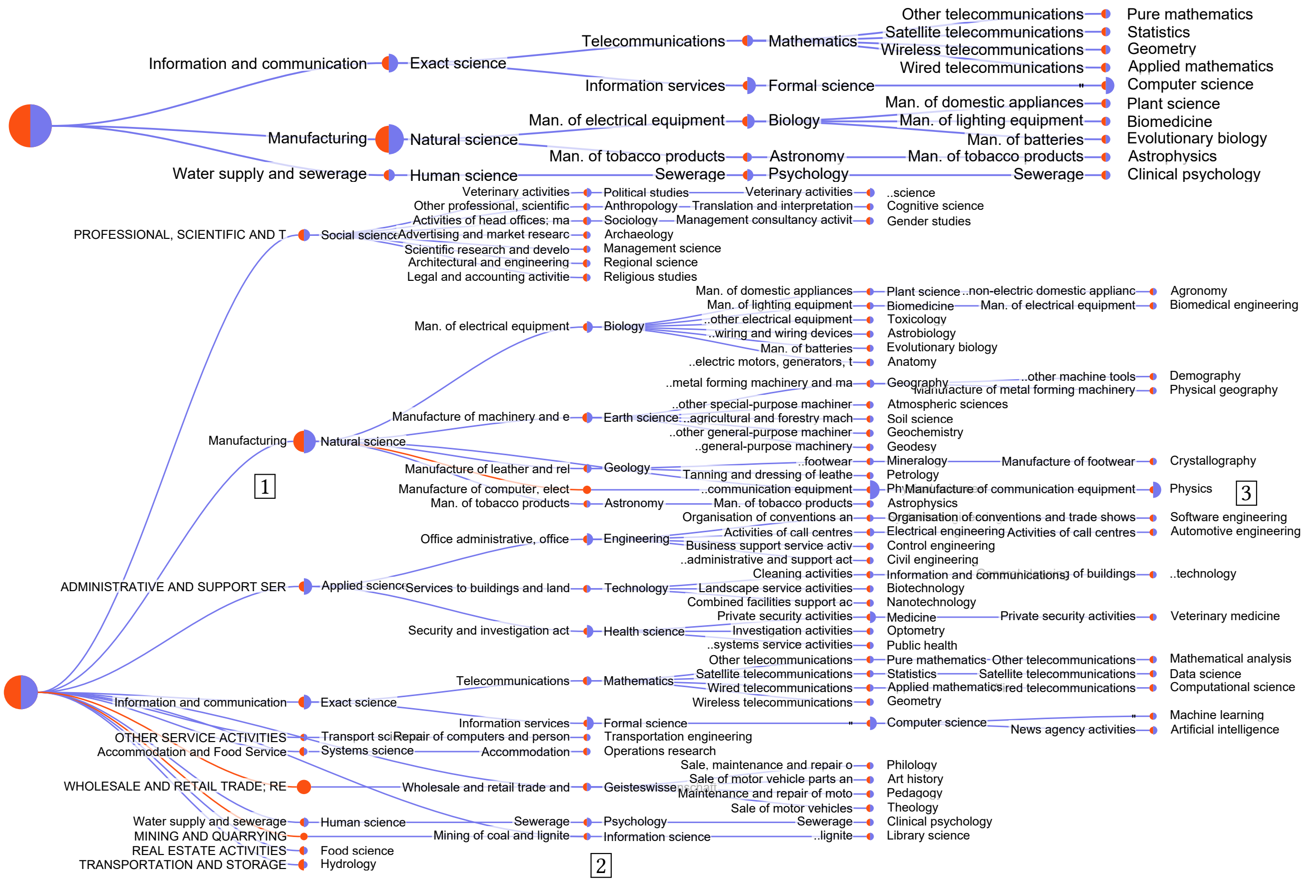


**Figure S2: Mapping the science taxonomy (blue) to industries (orange). The area of the (half-)circles encodes the size of each field or industry. The mapping preserves both the parent-descendant relationships and the size of the nodes. For example, natural science (1) is assigned to the largest industry – manufacturing, while biology and astronomy are mapped to different types of manufacturing. Fields are allowed to "skip" levels of hierarchy to better match their attributes, like in the case of information science (2). And for physics (3) there is no type of manufacturing that is sufficiently large to represent it.**

perform the style transfer for each author image with the method of Gatys et al. [2].

## S4 QUALITATIVE STUDY

To learn more about how people perceive and interact with metaphors, we conducted a user study. We opted for a qualitative study because it is better suited for our rather unconventional idea of metaphorical visualization and allows us to study aspects that we might not have anticipated.

### S4.1 Study design and analysis

We recruited 10 participants who are doctoral students working on visualization and HCI at a local department (3 female and 7 male, aged 27-36). Aiming to study metaphors in a personalized context, we purposefully collected data and built metaphors about the people at the department, which included the participants themselves. We obtained the data from Microsoft Academic, retrieving 13k authors that published in top visualization venues and 18k topics (keywords) attributed to them. Similarly to Sec. 5.2, we extracted embedding vectors for 50 authors at the department. The authors were then mapped to words (Sec. 5.1), cat images (Sec. 5.2) and visual styles (Sec. 5.3).

All three metaphors could be explored in a web-based tool that we created for the study (see Sec. S5 for details). The tool displays all 50 authors as draggable notes on a digital corkboard so that the users can perform affinity diagramming. This setup provides the participants with a simple task that encourages them to explore the metaphor. It also allows them to better illustrate the perceived similarities and clusters. The users can also create new metaphors by dragging concepts from the left panel onto the authors, thereby providing initial assignments for some authors. The tool would then compute the concepts for the remaining authors, continuing the metaphor.

A session with each participant lasted around 45 minutes in an one-on-one video call, the screen and the audio were recorded. It began with a short introduction to metaphorical visualization. Then, the participants were shown their personal Microsoft Academic page and the dataset origin was explained. Next followed the main part of the study, where the participants used the web tool to interact with the metaphors. This usually began by asking the participant to find themselves and a few people with similar concepts. Afterward, they were instructed to search for any other similarities and arrange the notes while thinking aloud. Participants were occasionally prompted to comment on why they considered grouped concepts to be similar, or whether the groups aligned with what they know about their colleagues. After about 5-10 minutes, the participants were explained how they can change the metaphor and prompted to try it. They would then comment on their existing groupings and continue to explore and edit the metaphor for about 5 minutes. This process was repeated for all three concept spaces in a different order, and the session was concluded with a 5-10 minute semi-structured interview. During the interview, the participants were asked to comment on which spaces they found easier to interpret, and which they liked better. They were asked about any expected or unexpected groupings that stood out and any particular assignments that they remembered. Also, we asked if they had any ideas for other concept spaces or applications of metaphorical visualization. Finally, the demographic data was collected.

During the analysis phase, the recordings were transcribed and annotated to include the groupings formed by the users in the tool. Then, the transcripts underwent an iterative coding process using NVivo, eventually generating 65 codes that were grouped into 10 categories. The codes and their metrics can be found in the supplemental materials.

## S4.2 Findings

**Perception of similarity in different spaces.** During the study we observed the participants explain how they reasoned about similarity in different spaces, and during the interview, they were asked which they found to be easier. When working with words, participants most often grouped them based on the topic, for example, technical terms (application, system, data), art-related (guitar, singer, poem), business (investment, contract, client), and so on. **P6:** *"*Groups 'data', 'system', 'control', 'database', 'application'* This is kind of software-ish."* But some more subtle and multi-faceted connections were also made, showcasing the flexibility that words can offer. **P7:** *"'Explanation' [and 'poem'], poems always need explanation or interpretation."* **P1:** *"'Honey' maybe comes with 'girlfriend', it depends if honey is honey [food] or honey [endearment], like *chuckles*."* It seems that the word space requires thinking and can be harder to interpret, but can also be more flexible and interesting, a point brought up by 4 participants. **P6:** *"Words need a lot of parsing, and thinking about 'does this work?'"* **P2:** *"*Matches 'celebration' and 'football' to 'music'.* I really like this, because it seems sooo infinitely dimensional. *laughs* It works, because there are so many directions that terms can be similar."* **P1:** *"If the domain is for exploration, maybe I use the words."* **P3:** *"For words you have to really look harder, I think. But it's also fun."* We believe that the flexibility of word similarity can be especially useful in fostering playful and creative exploration of personal data.

When looking for cat similarity, people most commonly relied on the color of the cat: black, white, orange, striped, etc. **P2:** *"Oh! We can go for black cats right away. *Groups 5 people.* Yeah, this works out nicely."* But sometimes they also used other traits, such as kittens or cats in cages. **P10:** *"I'm just looking for some baby cats."* **P3:** *"Apparently, [Person A] and [Person B] and [Person C] are similar because they all have cats with this type of head."* **P8:** *"I put these two together because they are both in cages."* In general, the opinions on how easy it was to perceive similarities in the cat space seemed to diverge, with some people saying that it was easier and others that it was harder than words. **P10:** *"You don't have to interpret it so much. You just see black cats."* **P6:** *"Cats were the hardest because I didn't really know what were the primary features."* **P7:** *"And cats were funny. I liked the cats as well, but I couldn't find much."* This suggests that the cat images, being visual, are compared more quickly, but some of the finer features may require time to interpret.

The style metaphor was regarded as the easiest for finding similarities by the majority of our participants. **P9:** *"And for the style transfer it was super intuitive, because I look at it, and it looks similar or not, this is like a no-brainer."* **P8:** *"The style visualization made me fully understand the connection between the researchers."* This was somewhat surprising to us, the advantage appears to be that the content of the image can be ignored, with only the color and the texture encoding the similarity. **P2:** *"[Person], well, he fits color-wise, and also texture I think is similar, that's cool."* Another contributing factor is that we used 16 styles, making clusters of similarly-styled people more easily detectable. Nevertheless, participants were able to not only find the clusters of identical style, but also in-between cases of similar texture or color. **P10:** *"And these two are something in-between. Here you have really smooth area, and here something's just blurred."* **P9:** *"[Person A] is like a weird case. It feels like he's between [Person B] and probably this group above here."* **P5:** *"Oh and look here, a small [Professor] group cluster, [Professor] is the same as us, between us and here."* This is quite encouraging and we think that such "data-enriched avatars" might make for an interesting future study.

**Expected and unexpected findings by participants.** We observed our participants construct many similarity groupings, both around themselves and involving others. In total, we coded around 180 similarity clusters and pairs being mentioned. In 106 cases, the participants indicated whether the discovered similarity aligned with their knowledge of their colleagues. Among those instances, we counted 90 that were expected and 16 that were not. The latter cases, where the perceived metaphor does not correspond to the user's expectation are the most interesting to us, so we have manually reviewed them by computing the underlying similarity between the researchers, i.e. the data points. Interestingly, in 12 out of 16 cases, the data similarity was above the 75th percentile or 0.27 cosine, i.e. the cases were not false positives, but rather represented similarities in the research topics that were unknown to the participants but reported by Microsoft Academic. **P2:** *"My cluster doesn't make sense. ['meat', 'chocolate', 'potato']. Maybe, I don't know. [Person A] does a lot of HCI, [Person B] as well. [cosine 0.55]"* **P5:** *"I think this one is similar to these ones, but I don't see the connection between these two and [Person]. [cosine 0.32]"*

We also performed a similar investigation for the 18 instances where participants said that they expected people to be similar, but did not find the expected similarity. We found 12 cases where the perceived lack of similarity was explained by the data, i.e. similarity was below the 75th percentile. **P5:** *"But I expected that [Person] is *closer to our topics. [cosine 0.17]"* **P1:** *"Interviewer (I): Who should be similar? – [Person], for example? But she's not similar. [cosine -0.04]"* **P4:** *"[Person A with 'organization'] I would put here [to 'potato', 'honey'], but the word doesn't fit. [cosine -0.02]"* Overall, we observed that most of the cases where the user's perception of the metaphor did not fulfill their expectations were actually in the data. Our prototype did not provide access to the underlying data, and it should have, in retrospect, because these cases indicate opportunities for users to verify and extend their knowledge of their colleagues. **P2:** *"I would've like to have [...] the papers, the keywords for the authors, to see how this could make sense. Because for some pairings it was surprising to see."* **P4:** *"[...] it would be interesting to know what types of topics are behind these images."*

**Metaphor creation feature underused.** One thing that our study setup did not encourage enough is changing the metaphor by manually defining assignments. We structured our tool around affinity diagramming, which worked well for understanding how similarity is perceived and data connections are made, but an unintended consequence was that people focused too much on finding similarity groups, which is logical in hindsight. A few participants were even confused that the whole mapping would be recomputed once they added an initial assignment and needed additional clarification. **P7:** *"Oh, 'temperature' is different. *Tries to fit it somewhere* Ooh, now everything's different."* **P8:** *"I try a white cat for [Person]. And now, what? Did mine also change?"* In future work, we'd like to have a tool that is more suited for building fun metaphors, e.g. by presenting only a few people at a time, prompting the user to define assignments, and automatically presenting some of the interesting outcomes.

**Fun and personalization.** One of our main goals for this qualitative study was to see if metaphorical visualization can provide a fun and casual way of exploring data. And throughout the study we observed participants find amusing assignments and associations. **P10:** *"*Laughs* Good, [Person A], 'error'. I have to make a screenshot. [And later after the study:] *Chuckles* The [Person A] error. I still have to send it to him."* **P2:** *"Let's see if there are more baby cats around. Oh [Senior researcher] does not, *chuckles* this is like an old cat."* **P9:** *"But then it's pretty funny that [Person A] has 'maintenance' and [Person B] has 'system' it's like a maintenance system they do together *chuckles*."* **P7:** *"['cousin' is put next to 'marriage'] *Laughs* and I don't marry my cousin. But yeah, *chuckles* maybe we need the police here."*

And even more interesting were the many cases where people connected the concepts to their associations about themselves and their colleagues. **P7:** *"Uh.. *chuckles* so I like .. *laughs* I'm one of the small cats because I'm one of the youngest here."* **P1:** *"I like this [Person] 'appointment'. – I: Why? – Like, when I remember [Person], the first thing that pops into my mind is the [Seminar] timing thing. *laughs*"* **P3:** *"He looked tired like this cat *chuckles* this morning when I saw him."* **P5:** *"Look! All [Project] people are easily "confused" or arguing with each other and explaining to each other things. *Groups 'depth', 'confusion', 'argument'.*"* We believe that creating such associations between the data and their personal experience, and especially making assignments that reflect them, can allow people to introduce their personal knowledge and connect with the data metaphor.

Of course, the participants come from our own department, but still, we were pleasantly surprised that many people spontaneously expressed that they enjoyed the experience, without being prompted. **P9:** *"I'm surprised that it worked so well, really. – I: Really? – It's crazy, yeah. I mean, if I would've see a graph layout of those, just for reference, I would argue that you would have to do some major trickery to get the amounts of freedom you need to describe something like this."* **P1:** *"This was fun."* **P5:** *"*Prompted to finish* Sorry, I'm obsessed now, I feel like I finally acclimatized to cats. *Continues to group cats.*"* **P2:** *"This was really cool."* **P4:** *"It's like playing "Memory" – I: Where you find similar pairs? – Yeah, exactly."* **P3:** *"I really like it. And I kind of would use it for memorization. I think if you connect vocabularies to funny images and so on, would be easier for learning."* **P7:** *"This is like that game where you have to insert words for a sentence and then something funny or politically wrong comes out. – I: Cards Against Humanity? – Yes *chuckles*."* This quality of data metaphors to be fun and enjoyable in themselves could be an important advantage when bringing data to casual users and applications.

## S5 STUDY TOOL

In Fig. S3, we show screenshots of the tool that we built for our qualitative study (Sec. S4). It is a single-page web application that uses a Python/C++ back-end to compute the metaphors. The tool allows participants to explore their colleagues' research topics through three metaphors: words, cat images and visual styles. The main area of the user interface ($\boxed{1}$) contains a sticky note for each person in the dataset, which also displays the concept assigned to them within the current metaphor. The cards can be freely dragged around, enabling the participants to build pairs and clusters of related concepts. The metaphor can also be changed using the list of concepts on the left ($\boxed{2}$). A word/cat/style can be dragged onto any note to make an assignment. As a result, the mapping is recomputed, finding a suitable concept for each remaining unassigned author, essentially continuing the user's metaphor.

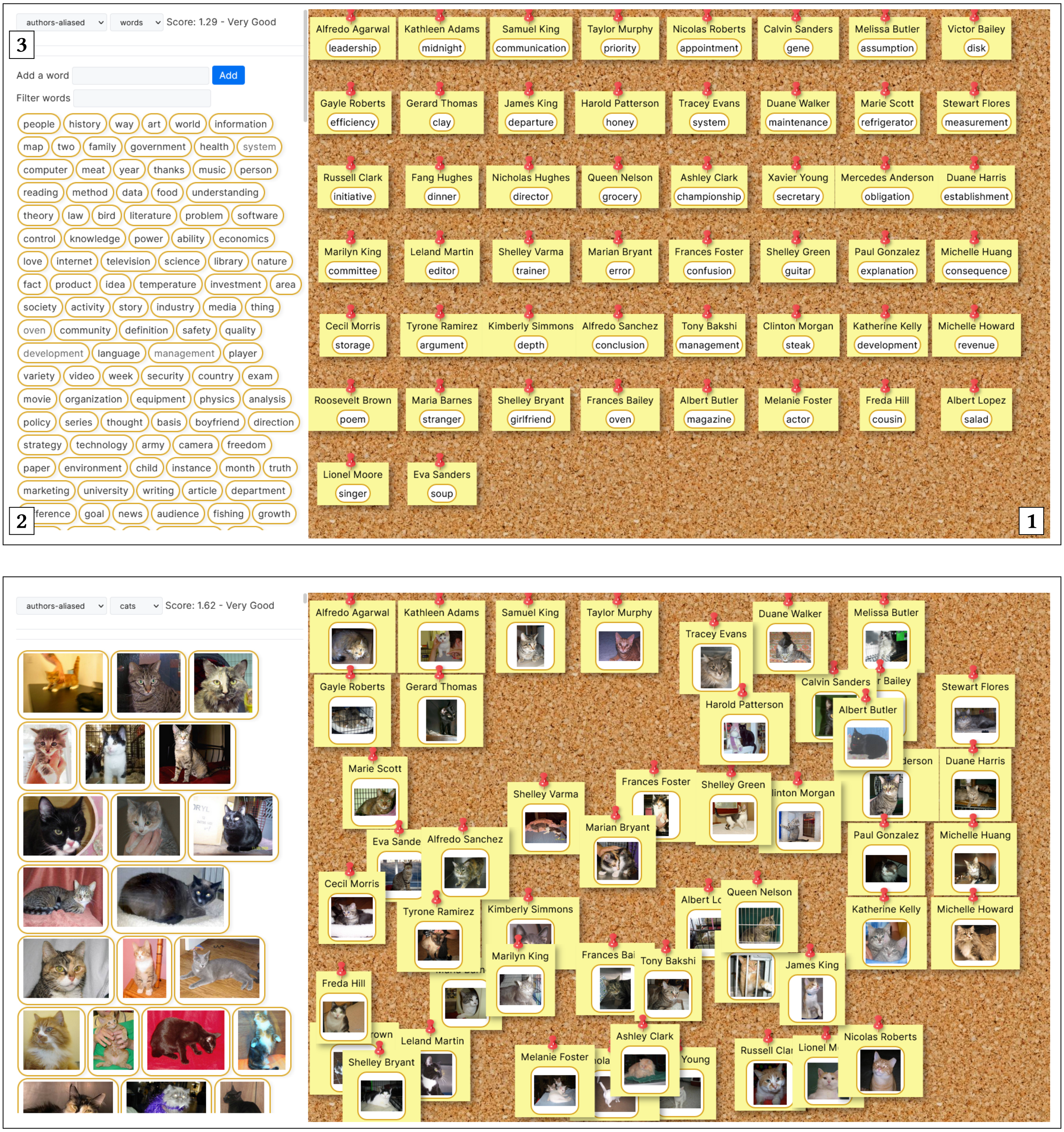


**Figure S3: The metaphor tool used in our study. All the names are anonymized. Top: The initial state of the tool displaying the word metaphor. 1: The main area used for affinity diagramming. Each draggable note displays a person's name and the word that they are currently mapped to. 2: A list of concepts that can be assigned to any person by dragging the concept onto a note. 3: The controls for switching between metaphorical spaces. Bottom: The tool during the exploration of the authors-to-cats metaphor.**

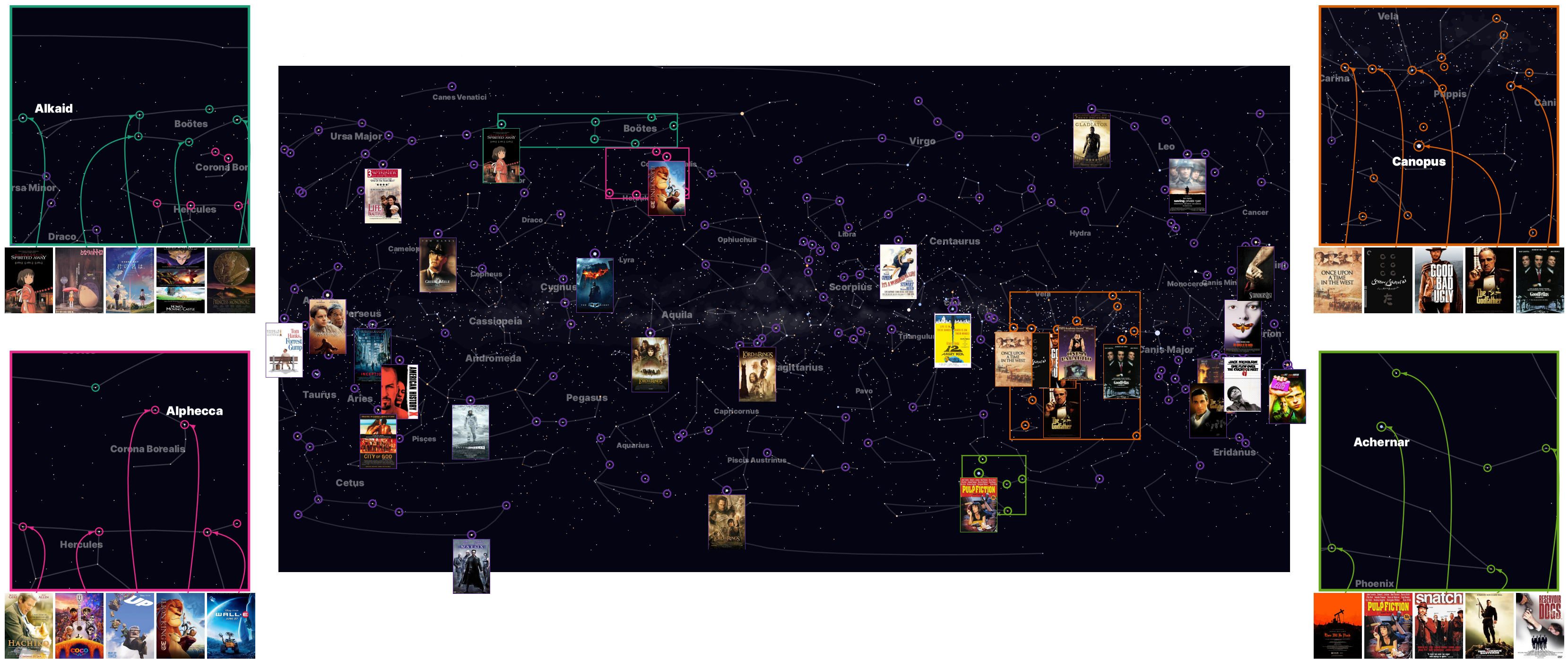


**Figure S4: A larger version of Fig. 5. The metaphorical mapping of popular movies to stars in the night sky. We assign well-rated movies to brighter stars, while also attributing related movies to neighboring stars. Users can explore this engaging infographic to build connections between stars and movies and to find out more about both in the process. For example, in the top-left corner, we see Miyazaki’s animated films (in green, “My Neighbor Totoro”, “Howl’s Moving Castle”, etc.) mapped to the Boötes constellation. Here the brightest star Alkaid was assigned to “Spirited Away”, suggesting that it is the highest-rated of the films. Just below, the animated film classic “The Lion King” (in purple) became Alphecca – the jewel of the northern crown (Corona Borealis). And the exceptionally positively rated “The Godfather” became the second-brightest star in the sky – Canopus, surrounded by the mafia, western and samurai movies. See also the supplemental video.**

## S6 CODING TABLE

Below is the coding table from our qualitative study. It consists of 65 codes grouped into 10 categories. For each code, we specify how many sessions (participants) included the code at least once, how many times the code has been applied in total and give an example extracted from the transcripts.

| Name | Partic. | Refs | Example |
|---|---|---|---|
| CLUSTER CONSISTENCY | | | |
| Cluster consistent after reassignment | 10 | 45 | “[Person A] and [Person B] stay really similar, as [Person C] and [Person D].” |
| Cluster consistent between spaces | 4 | 8 | “The [Project] people they always matched, they were very similar. For the cats, and also for the words.” |
| Cluster inconsistent after reassignment | 8 | 29 | “And ‘salad’, ‘potato’, this fits quite good, except for ‘drawer’ I would not put the drawer here.” |
| Cluster inconsistent between spaces | 1 | 2 | “I think we didn’t have the same similar words that I would put together, but now with images, I think they have a similar style.” |
| EXPECTED AND UNEXPECTED | | | |
| Expected dissimilarity | 4 | 5 | “They’re not black cats anymore, which is also something that I’ve expected, because I’ve never published with them.” |
| Expected similarity | 10 | 90 | “I found a cat similar to mine, and it’s [Person]! And this makes sense because [. . . ]” |
| Unexpected dissimilarity | 8 | 18 | “Actually, I feel a bit weird about [Person], the keyword is ‘midnight’, but [Person] and I, we published a paper together.” |
| Unexpected dissimilarity - data agrees | 4 | 6 | |
| Unexpected similarity | 7 | 16 | “Actually, [Person] looks quite similar, which is kind of surprising. I’m not sure how they relate.” |
| Unexpected similarity - data agrees | 2 | 4 | |
| FINDINGS | | | |
| Hypothesis building | 3 | 5 | “Like if this is the hypothesis, according to it, [Person] should be here, but he’s not.” |
| Interesting | 6 | 10 | “But when you change the style of a related person, you would see she got also changed. That is very interesting.” |
| Outlier | 4 | 13 | “And [Person] is still not really close to anything.” |
| Remembered cases | 6 | 10 | “I remember mine ‘obligation’, ‘appointment’ for [Person].” |
| Repeated finding | 2 | 6 | “I appreciate that even after the change [Project] people are still [Project] people.” |
| Surprising finding | 3 | 5 | “I would’ve thought that I would have more similarity with [Person] than him. That was kind of surprising.” |
| Unknown person | 8 | 13 | “Oh, [Person] ‘soup’. I don’t know her.” |
| FUN AND CREATIVITY | | | |
| Creative similarity | 4 | 8 | “‘Explanation’ and ‘depth’, I interpret it as a deep explanation.” |
| Cross-space association | 6 | 14 | “[Person] *laughs* ‘supermarket’ is still fine, I think he goes to the supermarket from time to time.” |
| Fun, cool, like it | 9 | 37 | “It’s a cool thing” |

| Name | Partic. | Refs | Example |
|---|---|---|---|
| Funny (dis)similarity | 3 | 4 | “And ‘confusion’ [Person] with ‘poem’. *Laughs* For me, at least.” |
| Funny assignment | 9 | 41 | “I like myself [the cat]. *chuckles*” |
| Funny cluster | 5 | 10 | “I would say the cluster *chuckles* is old people from [Department].” |
| Prefers some space | 3 | 3 | “I: Which space next? – Cats!” |
| IDEAS AND SUGGESTIONS | | | |
| Feature request | 4 | 6 | “I would’ve like to have [...] the papers, the keywords for the authors.” |
| Idea metaphor application | 7 | 9 | “I would like this on my Facebook data, or Instagram. I think this is good for personal visualization.” |
| Ideas for metaphors | 7 | 10 | “Maybe flowers? Because there are so many different types of flowers and colors.” |
| PAIRS AND CLUSTERS | | | |
| Cat cluster | 10 | 46 | “This is similar here, [Person A] and [Person B], cats with black-and-white faces. [Person C] also fitting into there.” |
| Cat pair | 8 | 42 | “This one, and this one.” |
| Style cluster | 9 | 56 | “Me, and then [Person A] have similar style, [Person B].” |
| Style pair | 10 | 43 | “*Puts [Person A] to [Person B].*” |
| Word cluster | 9 | 53 | “Okay, then ‘actor’ [Person A], ‘director’ [Person B]. *Pulls [Person C] ‘singer’ close too.*” |
| Word pair | 10 | 54 | “‘leadership’ [Person A] and ‘championsip’ [Person B] maybe.” |
| STRENGTHS | | | |
| Better than traditional vis | 5 | 5 | “I think the style visualization can play at least the same role as a scatterplot for this use case.” |
| Flexible | 2 | 2 | “[metaphors] provide you more degrees of freedom [that] you usually don’t have.” |
| Reassignment for exploration | 3 | 3 | “But I think it’s good because by switching this assignment you can filter out fake metaphors from the true metaphors.” |
| Works well | 5 | 16 | “And after all, the style visualization is ... the style visualization made me fully understand the connection between the researchers.” |
| SHORTCOMINGS | | | |
| Assignment to ‘fix’ the metaphor | 3 | 4 | “Let’s give [Person] ‘reading’ so that he’s now more in the ‘literature’.” |
| Hard to choose a concept | 2 | 5 | “I: What would you [assign to] yourself? – I don’t know.” |
| Interpretation subjective | 1 | 1 | “The interpretation is really subjective.” |
| Offensive concern | 1 | 2 | “For the words it’s maybe a bit awkward because you don’t get control for the words that you get there, if it’s offending.” |
| Confusion after assignment | 4 | 9 | “Oh, everything has changed.” |
| SIMILARITY INTERPRETATION | | | |
| Learning similarity concepts | 3 | 5 | “So you have to build some kind of a base, and then it’s possible to interpret what these terms mean together.” |

| Name | Partic. | Refs | Example |
|---|---|---|---|
| Cat black cluster | 7 | 11 | "I will make a cluster of black cats, see who's there and comment on the results." |
| Cat similarity concepts | 10 | 37 | "Okay, here's a red cat, and orange cat." |
| Data similarity interpretation | 10 | 39 | "What I also recognize now, there's a cluster of professors." |
| Gradual similarity | 9 | 22 | "It feels like he's between [Person] and probably this group above here." |
| Style similarity concepts | 7 | 12 | "The image of [Person] got very strong blurring effect." |
| Photo impacts style similarity | 2 | 3 | "I have a black-and-white picture and most of the others are colored. That might be an issue." |
| Word similarity concept | 9 | 26 | "All the words I hear at work, I put then in this corner." |
| SPACE COMPARISON | | | |
| Cats are difficult | 6 | 7 | "But for images, it was for me a little bit complicated." |
| Cats are easier to interpret | 4 | 5 | "I mean, you don't have to interpret it so much. You just see black cats." |
| Cats are fun | 3 | 3 | "But I also like the cats." |
| Cats not enough variation | 2 | 2 | "Cats were also kind of visual, but it was … the features are not different enough." |
| Styles are difficult | 2 | 2 | "Yeah, [styles] work the worst, I would say." |
| Styles are easy | 8 | 20 | "Style transfer was much easier to sort." |
| Words are difficult | 6 | 11 | "For words you have to really look harder." |
| Words are easy | 3 | 5 | "Ah words, words were easy. Words fit well most of the time as well." |
| Words are flexible | 2 | 5 | "I think the text metaphor was very broad, many different categories." |
| Words are fun/interesting | 4 | 4 | "I knew it was a stretch to put food in refrigerator. But still kind of cool." |
| Words good for exploration | 2 | 4 | "But if you'd like to explore more, maybe the words is better, because it offers more." |
| Words need thinking | 5 | 5 | "I think, I don't dislike the words, but you need more time to find clusters there." |
| MISC | | | |
| Assigning concepts to others | 10 | 29 | "Soo, let's do 'paper' for [Person]." |
| Assigning concepts to themselves | 8 | 17 | "With me, I would put 'video'. Since this is my main stuff." |
| Important metaphor considerations | 4 | 6 | "The metaphor should have some characteristics that are easily distinguishable." |

## REFERENCES


[1] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. 2020. A Simple Framework for Contrastive Learning of Visual Representations. *arXiv:2002.05709* (2020). arXiv:2002.05709

[2] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. 2015. A Neural Algorithm of Artistic Style. *arXiv:1508.06576 [cs, q-bio]* (Aug. 2015). arXiv:1508.06576 [cs, q-bio]

[3] K. He, X. Zhang, S. Ren, and J. Sun. 2016. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 770–778. https://doi.org/10.1109/CVPR.2016.90

[4] Matthew Honnibal and Ines Montani. 2017. spaCy 2: Natural Language Understanding with Bloom Embeddings, Convolutional Neural Networks and Incremental Parsing. (2017).

[5] Petra Isenberg, Florian Heimerl, Steffen Koch, Tobias Isenberg, Panpan Xu, Charles D. Stolper, Michael Sedlmair, Jian Chen, Torsten Möller, and John Stasko. 2017. Vispubdata.Org: A Metadata Collection About IEEE Visualization (VIS) Publications. *IEEE Transactions on Visualization and Computer Graphics* 23, 9 (2017), 2199–2206. https://doi.org/10.1109/TVCG.2016.2615308

[6] Kaggle. 2014. Dogs vs Cats. https://www.kaggle.com/c/dogs-vs-cats.

[7] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2020. Analyzing and Improving the Image Quality of StyleGAN. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 8107–8116. https://doi.org/10.1109/CVPR42600.2020.00813

[8] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed Representations of Words and Phrases and Their Compositionality. In *Advances in Neural Information Processing Systems (NIPS)*. 3111–3119.

[9] Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *International Conference on Learning Representations (ICLR)*.

[10] Suddala Srujan. 2020. Art for Generative Modelling. https://www.kaggle.com/suddalasrujan/art-for-generative-modelling-images/version/3.