

PivotPaths: Strolling through Faceted Information Spaces

Marian Dörk, Nathalie Henry Riche, Gonzalo Ramos, and Susan Dumais

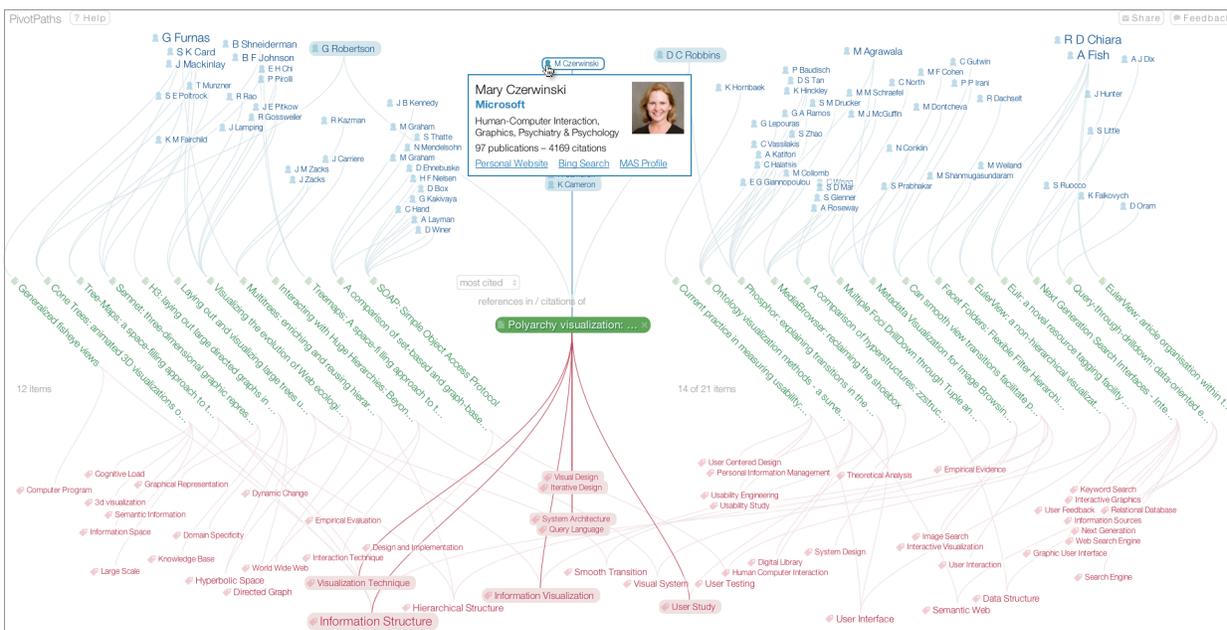


Fig. 1. A resource anchor selects a paper's references and citations. Details are shown for one of its authors, Mary Czerwinski.

Abstract—We present PivotPaths, an interactive visualization for exploring faceted information resources. During both work and leisure, we increasingly interact with information spaces that contain multiple facets and relations, such as authors, keywords, and citations of academic publications, or actors and genres of movies. To navigate these interlinked resources today, one typically selects items from facet lists resulting in abrupt changes from one subset of data to another. While filtering is useful to retrieve results matching specific criteria, it can be difficult to see how facets and items relate and to comprehend the effect of filter operations. In contrast, the PivotPaths interface exposes faceted relations as visual paths in arrangements that invite the viewer to ‘take a stroll’ through an information space. PivotPaths supports pivot operations as lightweight interaction techniques that trigger gradual transitions between views. We designed the interface to allow for casual traversal of large collections in an aesthetically pleasing manner that encourages exploration and serendipitous discoveries. This paper shares the findings from our iterative design-and-evaluation process that included semi-structured interviews and a two-week deployment of PivotPaths applied to a large database of academic publications.

Index Terms—Information Visualization, Interactivity, Node-Link Diagrams, Animation, Information Seeking, Exploratory Search.

1 INTRODUCTION

Digital information spaces are not just growing in size, but also in their semantic and structural richness. For example, academic collections not only reference core resources (publications), but they also include a range of additional attributes (e.g., authors, keywords) and relationships (e.g., citations). Similarly, in a movie database, movies are the core resources, with actors and genres as important attributes, and sequels and remakes providing additional relations. While systems often use these attributes and relations to display details for individual resources and to support filtering, we believe these rich connections can fuel deep and engaging explorations of an information space.

We argue that the semantic and structural aspects of data are cur-

rently underutilized in the representation and exploration of information spaces. In particular, data attributes and relations are mainly used for searching for information resources and navigating from one item to another, resulting in information presentations that lack context and feel disconnected. Besides helping to find and browse individual resources, we want to use semantic and structural relations to widen the angle on an information space and convey what resources are about and how they relate. For example, taking a look at the keywords of an author's publications tells us something about her academic interests. Seeing the actors from movies of a given genre gives us a sense for the subset of movies. How can we expose these hidden aspects of information spaces to facilitate open-ended exploration?

There has been considerable research interest in data attributes as *facets*, i.e., multiple classifications that can be used in unison to filter resources. Movies can be filtered, for example, along genre, cast, and rating. Faceted interfaces typically display textual lists of facet values that can be combined to refine a result set. While faceted interfaces utilize facets to filter resources, they do not actually represent how facets relate to one another, or how resources are connected via references. Recent research on visualization investigated the utility of visual representations to make sense of faceted document collections [2, 12].

- Marian Dörk is with University of Calgary, E-mail: mdoerk@ucalgary.ca.
- Nathalie Henry Riche, Gonzalo Ramos, and Susan Dumais are with Microsoft, E-mail: {nath, gonzalo, sdumais}@microsoft.com.

Manuscript received 31 March 2012; accepted 1 August 2012; posted online 14 October 2012; mailed on 5 October 2012.

For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

These visualizations are targeted mainly at high-level analysis and not at open-ended exploration of the information resources and facets.

In this paper, we present *PivotPaths*, a visual interface for exploring faceted and interconnected resources (see Figure 1). *PivotPaths* exposes facets and relations as visual paths that invite the searcher to ‘take a stroll’ through an information space. *PivotPaths* features a visualization canvas that arranges facet values and information resources in spatial association and connects them by subtle curves indicating facet-item-facet relationships. Although we designed *PivotPaths* for ease and aesthetics, it also provides powerful pivot operations for navigating along different visual arrangements that are integrated with each other through a coherent visual language and gradual transitions. The interface is intended for both novices and experts to visually traverse large information spaces in an aesthetically pleasing manner that encourages exploration and serendipitous discoveries.

We deployed *PivotPaths* in a large research institution for two weeks. We report on the feedback we collected and our observations of *PivotPaths* in use. We present two participants’ strolls that illustrate the flexibility of *PivotPaths* and describe the discoveries they made.

2 RELATED WORK

PivotPaths relates to prior work on supporting information seeking through interaction and visualization design. Information seeking is often thought of as the back-and-forth between searching and browsing, somewhat resembling the act of berrypicking [1]. Extending this idea, the notion of *exploratory search* suggests the blending of querying and browsing to pursue higher-level information needs [13]. Our research is motivated by this initiative towards a “richly associated search across heterogeneous spaces” [22]. We find these ideas substantiated in the persona of the information flaneur [6], curiously traversing information spaces by using visualizations to find serendipity [18].

Most of the interfaces offering access to large collections of digital or physical artifacts (e.g., digital libraries, online shops, and image search) rely on metadata to support searching, browsing, and filtering. Orthogonal sets of categories that characterize elements of a collection are referred to as *facets*. Early interfaces designed for navigating resources along facets feature lists of links for facet values [24] and coordinated columns for each facet [23]. While faceted interfaces offer powerful ways of filtering large sets of resources into smaller sets, relationships across facets and resources are seldom represented.

There have been several efforts to support the exploration of faceted information collections through visualization. For example, visual encodings for the relative number of resources per facet value [17] and cross-facet relationships through brushing [5] can guide the exploration of information spaces. More specifically, faceted visualizations of academic paper collections can help analyze temporal and topical patterns in digital libraries [10, 11, 12]. In addition, visualization systems have been developed to support visual analysis of cooccurring entities in large document collections [4, 16, 21]. We wish to extend this line of research that was mainly framed around investigative analysis. In contrast, our goal is to support more casual forms of information navigation utilizing lightweight visualizations. Existing faceted visualizations feature multiple views that represent different entity types as interactively coordinated, yet visually separated boxes. We adopt a ‘softened’ multi-view approach with regions that are flexible in size, visually linked by facet curves, and permeable to moving entities.

Besides facets, many information spaces contain references such as citations that let us think about information seeking through the lens of graph visualization. On one hand, citation patterns can be aggregated to support analysis of scientific research at a high level [3]. On the other hand, instead of representing an entire citation network at once, one can search for a particular node as a starting point from which to traverse a graph iteratively [19]. Furthermore, by positioning nodes based on attribute similarities and distinguishing between incoming and outgoing edges, it becomes possible to reveal the extent of a node’s influence [7]. A recently introduced visualization of citation patterns reveals the genealogy of academic papers through tendrils connecting paper titles organized on side-by-side chronological lists [14]. The system is visually enticing and informative up to a point, yet it does

not permit people to pivot among the database facets.

Several techniques have been introduced to allow for pivoting through a graph along different types of attributes, entities, and structures. For example, a multivariate graph can be rolled up into a meta representation, in which aggregated nodes are positioned on a grid according to their attributes [20]. Facets can also be shown as edges connecting entities that share the same facet values [2]. When dealing with multiple information structures in parallel, visual transitions between them can help maintain context [15]. While animated transitions between these different representations can help understand pivot operations, it remains cognitively difficult to follow and comprehend display changes. To provide more awareness of a sequence of data transformations, a graph can represent the refinement operations [8], however, this system does not provide much for faceted pivoting.

We believe that users struggle to grasp the concept of pivoting, because this trivial notion, in fact, hides the complexity of operations that may not always be reversible. Furthermore, the term pivoting is used for different types of shifting between information structures [15], aggregate representations [20], facet or entity types [2], or refinement sequences [8]. With our research we use a simpler notion of pivoting: gradual moving between sets of resources. Pivot operations in existing systems may be animated, however, they often result in drastically changed layouts. Pivoting from papers to authors or from one hierarchy to another flips the underlying structure of the display. Helping the viewer to grasp these complex operations remains a challenge. Our pivoting operations are designed to not change the type of structure, representation, facets, and entities. Instead we seek to pivot between partially overlapping sets of resources and facet values, where the overlap in elements provides context for the viewer.

While we share many goals with these previous efforts, none of these projects fully captures the notions of casual exploration, faceted relationships, light-weight representations, visual consistency, and graceful transitions. Our work with *PivotPaths* aggregates and extends these notions and techniques to support a type of visual pivoting as a method for ‘strolling’ through a faceted information space.

3 STROLLING THROUGH INFORMATION SPACES

We propose the notion of ‘taking a stroll’ through an information space using visualizations that provide visual cues and interactive pathways. Our fundamental proposition is to explicitly reveal diverse facets and connections between information resources in a way that both elicits and facilitates casual exploration of large information spaces. In the following, we motivate and characterize our ‘strolling’ approach, and articulate research questions, and design goals.

3.1 Search, Browsing, and Visualization

Searching is primarily geared towards finding one or a few relevant resources. As such, the search process is primarily about refining a result set from the very many to the very few. Existing tools for traversing information spaces such as academic publications or movie collections tend to promote a search and narrow behavior, relying on a reductionist approach that displays top lists based on metrics such as citations or sales. At the end of a search, there tends to be the single *find* that ideally satisfies the information need. The search process can be seen as a funnel-like task, where the searcher’s choices are iteratively reduced. In contrast, we are interested in a lateral process that expands the options for navigation by showing tangential information and encouraging the searcher to expand routes through an information space.

During *browsing* the information seeker navigates from one resource to the next via some form of connection or hyperlink. While browsing is much more open-ended than searching, one typically stays at the level of individual resources. When selecting a link, the contents of the display are replaced by the next resource. Abrupt display changes and the low-level (item to item) nature of browsing systems make it difficult to put resources in context with each other and see the attributes they have in common. We want to take the open-ended aspect of browsing and raise it to a higher level in the experience so that it allows the person to use their vision to relate resources and facets.

Visualization allows us to interact with an information space at a much higher level than searching and browsing. For example, a visualization of publication trends or movie sales, allows an analyst to identify global patterns within these collections. This type of representation situates the viewer/analyst at a meta-level from which she can observe an entire collection. While this vantage point allows for observations of overall trends, it can also make it difficult to engage with particular instances in an information space since they are reduced to abstract marks and shapes. We seek to use visualizations to raise the level of search and browsing, but not to the point where one loses contact with individual resources. By providing access to individual resources and facet values, we aim to support exploration within the information space while using visual representation and direct manipulation techniques as vehicles for casual exploration and discovery.

We explore a way to use visualization for information navigation that does not follow the refinement model of search, departs from the abrupt, context-losing aspects of browsing, and operates at a level of (visual) granularity that permits seeing both the forest and the tree.

3.2 Strolling: Casual Exploration using Visualization

We take inspiration from the information flaneur [6], a persona characterizing an information seeker, who casually explores information spaces with the curious mindset of an urban explorer. It is not the literal interpretation of strolling as a physical activity that informs our approach, but rather its experiential aspects. Similar to the smells, sounds, and sights that are enticing cues when exploring a city, facets and relations can be revealed as cues luring the information flaneur into pursuing different paths through an information space.

While drawing elements from search, browsing, and visualization, we propose a new way to traverse an information space that differs from the aforementioned modalities. Instead of trying to find the most relevant resources (search), navigating from one resource to the next (browsing), or analyzing a collection at a high level (visualization), we formulate a hybrid between these approaches and define *strolling* as:

the casual exploration of an information space using interactive visualizations that offer a wider perspective on relations and facets, and provide diverse pathways for gradually pivoting between sets of related resources.

To illustrate what we mean by strolling, let us imagine Fiona, who enjoys watching old movies. Recently she developed a lingering desire to find out more about film noir. Since she talked about this movie period with a cineaste friend, she wanted to learn more about this genre. Once, when she had a slow afternoon at work, she visited a website about movies that offered visualizations. She entered ‘film noir’ which brought up popular movies of the period. The site displayed the main movies in the centre and arranged relevant actors and themes around them. After glancing over the display, she notices an actor who appeared in several popular movies. Taking a look at the movies he appeared in, she is able to see the other actors he worked with and related genres and subjects. She notices that he often played in crime-related movies, a movie subject she is quite fond of. Finally, she returns to the initial ‘film noir’ view and continues exploring the space.

In the previous scenario, Fiona was driven primarily by her curiosity towards an area of interest that she pursued by casually navigating between sets of resources. A range of information activities can be captured by this notion of strolling including egocentric search, familiarizing oneself with a new domain, or simply aimless exploration of an information space. For these types of activities existing interfaces often introduce a focus that constrains the open-ended nature of exploration and the likelihood of serendipitous discovery. The question is: How can we conceive visual interfaces that encourage strolling?

There are difficult challenges related to the design of visual representations and interactive capabilities in support of casual exploration. We seek to find a middle ground between low-level access to individual items and high-level views of entire collections. The interface should both convey relationships and facets, and provide ways to explore resources and pivot between them. One significant limitation of

most exploration tools are abrupt visual changes between interactions. To support a continuous exploration experience, how can we make display changes more gradual and comprehensible?

3.3 Design Goals

The preceding set of aspirations and questions translates into five concrete design goals (DG) that form the basis for the PivotPaths system:

1. *Relate facets and items.* The interface should use visual variables such as color, shape, and position to convey how resources are related to one another and how they share facet values.
2. *Encourage pivoting.* Every visual element in the interface (regardless whether facet value or resource) should provide simple pivoting interactivity for moving between sets of resources. Pivoting steps should be ‘cheap’, in that they should be simple to trigger and reverse.
3. *Make view changes gradual.* Transitions between interface states should not be abrupt, and should explain how and why information changes. Even before a pivot operation is triggered, the interface can indicate some of the changes.
4. *Aim for inviting aesthetics.* The overall appearance of the interface as well as the individual components should follow a consistent visual language that elicits meaning and function, is easily understood, and is ultimately enjoyed.
5. *Situate in Web context.* The interface should be situated in the context where information seeking already happens, i.e., the Web browser. Common uses and features of the Web should be supported to benefit the strolling experience.

4 DESIGNING PIVOTPATHS

To explore how a visual interface could support strolling, we carried out a design process through multiple iterations. These iterations have been accompanied by a range of evaluations: interviews with search, visualization, and interaction experts as well as observational user studies. We also conducted an intranet deployment in which PivotPaths was used by hundreds of people and discuss the feedback we collected. These evaluations provided guiding advice and evidence for the design and development of PivotPaths, from early mockups and prototypes to a fully functional system.

In this section, we present the final design of PivotPaths, especially regarding the representational and interactive aspects, and highlight the decisions that originated from the initial prototypes and usage observations. Note that we provide later in this paper a more detailed discussion of the findings from the deployment and observations, that suggest further improvements to our present design.

4.1 Visualizing Tripartite Information Spaces

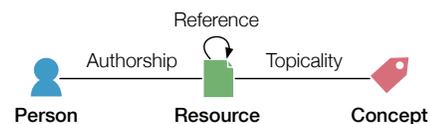


Fig. 2. Many information spaces resemble the structure of a tripartite graph comprising people, resources, and concepts.

Information spaces can be seen as heterogeneous graphs containing different types of edges and vertices. In this sense, PivotPaths is a graph exploration tool that allows the viewer to navigate an information space as a tripartite graph containing people, resources, and concepts (see Figure 2). This schema is based on a common structure that captures the many information spaces involving, for example, people as creators of cultural artifacts that are referenced by other artifacts and characterized by concepts. In particular, academic information spaces follow this kind of structure with researchers as authors, publications as resources, citations as references, and keywords suggesting

the main subjects of publications. To describe the design of PivotPaths, we use publication data retrieved from Microsoft Academic Search via their public API. We will discuss other information spaces later.

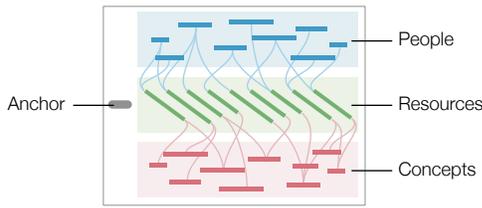


Fig. 3. The interface of PivotPaths is separated into three main regions for people, resources, and concepts. Curved edges between facet values and resources indicate faceted relationships.

The PivotPaths interface consists of an interactive canvas for navigating such tripartite information spaces. The canvas has three main regions accommodating people, resources, and concepts (see Figure 3). In the current design, the PivotPaths system treats people and concepts as the facets of resources, however, other perspectives are possible, for example, it is feasible to view resources as facets associated with people. The central area can be seen as a sample of resources that the viewer can inspect, while the facet labels above and below the resources provide contextual information about them.

The sample of resources is based on one or two anchors. Every entity—person, resource, or concept—can be an anchor that selects the set of resources displayed in the middle. If the set of resources is too big to fit the display a subset of resources is selected based on one of three orderings: most recent, most cited, or random. In cases where an anchor selects more items than can be displayed, adjusting the type of the ordering changes the nature of the subset. We show the complete metadata for the sample in the middle area, but do not include facet values of the hidden resources.



Fig. 4. An element has three main components: icon, label, and control.

As illustrated in Figure 4, every element in the visualization has an icon indicating its type, a short label, and a control for setting or removing it as an anchor. While every resource element has the same size, the sizes of facet elements vary to indicate the relative number of resources they are associated to. For example, the author ‘G G Grinstein’ at the top of Figure 7 is one of the largest labels, since it is associated with the most resources in the current view.



Fig. 5. The appearance of elements changes according to their state from least to most salient with increasing significance.

The goal of designing the appearance of the elements was to make the visual arrangements and elements inviting for interaction and visually appealing (DG 2 and 4). Individual elements are displayed according to their entity type and interaction state (see Figure 5). We have chosen tones of blue, green, and red to consistently display people, resources, and concepts. An element can be in one of four states:

default, brushed, hovered, or anchor. When no interaction affects the element it is displayed in its default state. Moving the mouse pointer over an element turns it into a hovered state, and related elements into a brushed state. By clicking on the label it is activated as an anchor.

These differing styles should reflect the viewer’s interest in the displayed elements at a given moment. We have designed slightly differing styles for each state, to convey a visual ranking between the states: high contrast and darker tones are used to convey higher significance, i.e., higher interest by the viewer. The current anchor is the darkest as it is primary for selecting the resources and thus the associated facet values. Furthermore, the hovered state is the second most salient element as it represents temporary interest of the viewer. Based on which element is hovered at a given moment, the brushed elements indirectly represent possibly interesting elements and thus have a more subtle highlight to contrast with the remaining elements.

4.2 Positioning

The positioning of elements is designed to convey how facets and resources are interconnected (DG 1). Element positions are determined differently depending on their type and the anchor(s). Resources are positioned in the middle of the canvas sorted horizontally according to the selected ordering. The facet elements (people and concepts) are positioned according to the resources they are associated with. Horizontally, facet elements are positioned based on the average position of the associated resources. Vertically, facet values are positioned to indicate connectivity; facet elements that are associated with more resources are placed further away from the resource versus those that have fewer resources. In the case of similar connectivity values, the vertical ordering is secondarily also based on total number of associated resources beyond the current sample. The effect is that highly connected people rise to the top, while highly connected keywords sink to the bottom. For example, one of Catherine Plaisant’s more regular collaborators, Ben Shneiderman, is positioned on the top of the people area, and she publishes many articles about ‘Information Visualization’, which is positioned on the bottom of the concept area (see Figure 7).

It is important to note, however, that the positioning of elements is not thought to be the basis of dedicated analysis, but rather in support of more casual exploration. Arranging facet values according to their association with resources can give a general sense how they are related, which in turn can inform subsequent navigation steps.

In addition to the relatively loose spatial association between facet values and resources, the visualization contains curves that communicate the relationships between facets and resources as well as across facets (DG 1). By default, the curves are displayed in a faint tone of the corresponding facet type’s color. However, when a facet element is in the brushed state and in another canvas region of the hovered element, the color of the respective curve changes to a darker tone to indicate how the brushed and hovered elements relate. This way it is possible to explore how certain facet values relate not just to resources, but also to values of the respective other facet.

4.3 Three Layouts for Relating and Anchoring Elements

The PivotPaths interface features three layouts based either on: a facet anchor, a resource anchor, or a pair of facet anchors (see Figure 6). Our intention behind the design of these layouts is to allow the viewer to relate resources along facet values and resource references (DG 1), and support pivoting between different anchors (DG 2). The anchors are used as filters selecting subsets of resources.



Fig. 6. PivotPaths offers three layouts that are based on either a single anchor (left and middle) or a pair of anchors (right).

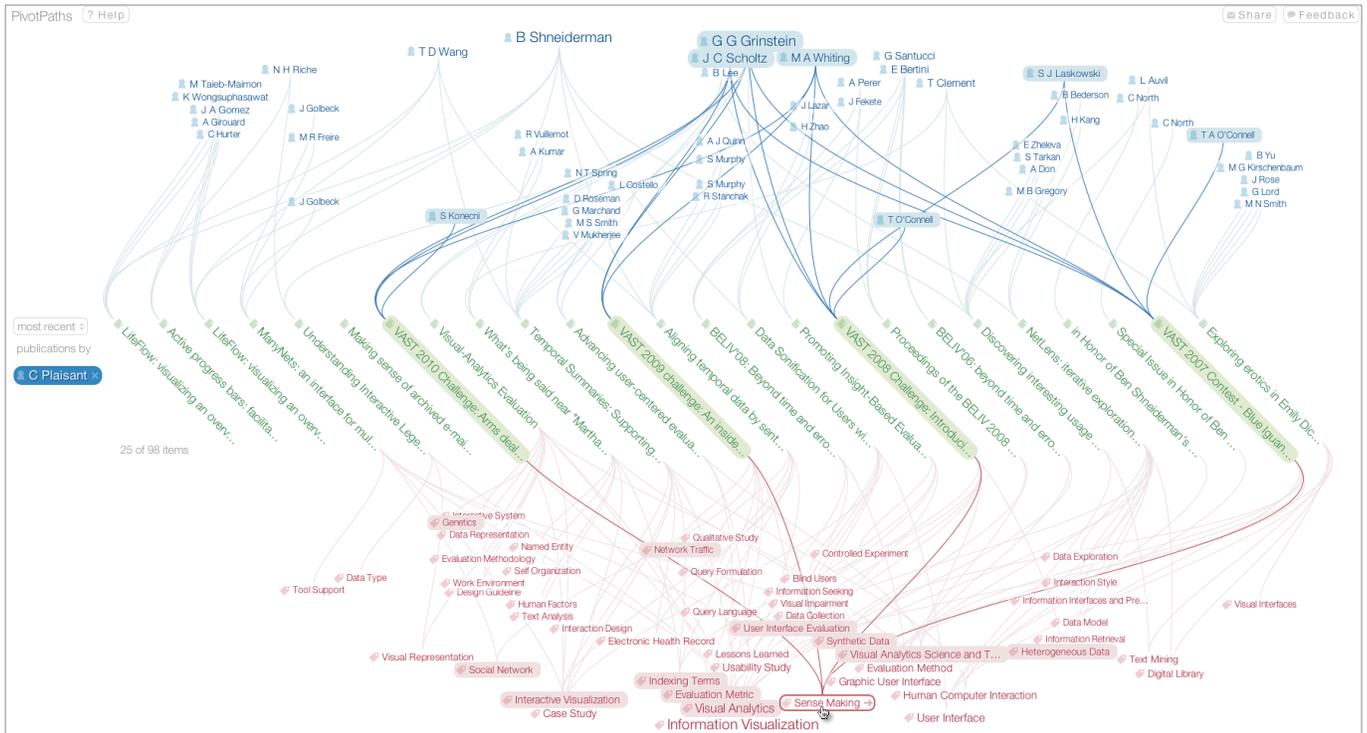


Fig. 7. Single Facet Layout: The facet anchor Catherine Plaisant selects her most recent publications and their co-authors and keywords.

Single Facet Layout. The most basic layout uses a single facet value as an anchor. The anchor is positioned to the left of the resources that it retrieves. For publications, this layout would either display a set of papers by a given author or on a given research topic. Figure 7 shows the most recent publications by Catherine Plaisant, with the people facet visualizing her co-authors (above) and the concept facet summarizing the main keywords (below). One can also select a concept as the anchor displaying the resources associated with a given keyword.

Resource Layout. Selecting a resource as an anchor retrieves resources that are linked by references. In the case of many information spaces, references are directed, such as an academic reference within one publication pointing to another. Based on the directionality of the reference, the resources in the middle area of the canvas are split into two groups. The resource-based anchor is positioned in the middle between these two sets of resources. In the case of publications, the publications referenced by the anchored publication are situated on the left, while the publications citing the anchored publication are positioned on the right (see Figure 1). Distinguishing between a publication’s references and citations suggests its impact in the context of other resources, authors, and topics. Before the resource anchor view was implemented, several people in our feedback sessions mentioned their interest in pivoting from one paper to the next. By iteratively pivoting along resources within this layout, it is possible to gradually move from a current paper towards earliest publications.

Comparison Layout. The third layout juxtaposes two facet anchors that can be values of either of the people or concept facets. The idea of the positioning is to provide a comparison view that is loosely based on Venn diagrams as illustrated in Figure 6, right. Resources that share both facet values are positioned in the middle between the anchors, and the resources that have only one of the two anchored facet values are positioned to the side of the respective anchor. Partitioning shared and distinct resources along two anchors allows us to compare, for example, two authors along the publications they have collaborated on. Furthermore, by positioning facet values (co-authors and keywords) according to their association with the partitioned resources, one can see what the anchored facet values have in common and how they are distinct respectively. For example, selecting Mary Czerwinski and

George Robertson as the two anchors, the keywords associated with the selected resources suggest complementary interests in research on ‘User Study’ and ‘User Interface’ (see Figure 8, bottom left and right).

4.4 Interactivity for Strolling

The interactivity of PivotPaths is designed to facilitate strolling through faceted information spaces (DG 2). In the following, we describe the main interactive capabilities for adjusting the set of retrieved resources and accessing individual elements. Every element displayed on the exploration canvas of PivotPaths offers three ways of interacting with it: setting or removing it as an anchor, displaying details, and highlighting connected elements. Furthermore, the sorting of the resources can be changed with a widget and the home screen provides a search box to find facet values and resources.

Anchoring elements is the most crucial to the concept of strolling through information spaces: every element (facet or resource) can become an anchor that selects resources. In order to set an element as an anchor, one simply clicks on the label, after which the previously anchored element loses its anchor state and returns to the default state. The newly anchored element moves to the left side of the canvas (if it is a facet value) or to the middle (if it is a resource). In contrast to these single-anchor layouts, the comparison view requires two facet anchors. When the current view is either in the single facet or comparison layout, hovering over another facet value reveals an arrow to the right of the label (see Figure 4); clicking this arrow will position the newly anchored element to the right side of the canvas.

When there are already two facet anchors, anchoring a new facet element with the arrow will reset the left anchor, move the right anchor to the left, and add the new anchor to the right.

The *sorting* of the resources also determines which resources are selected and which are capped. Above each anchor is a drop-down menu that allows the viewer to choose between most recent, most referenced, and random. We included a random sorting to facilitate more playful exploration that could lead to unexpected discoveries. A random selection also gives a more representative sample of resources, than with selections based on recency or citation counts. Many more orderings are feasible. For example, in the case of academic publications, we

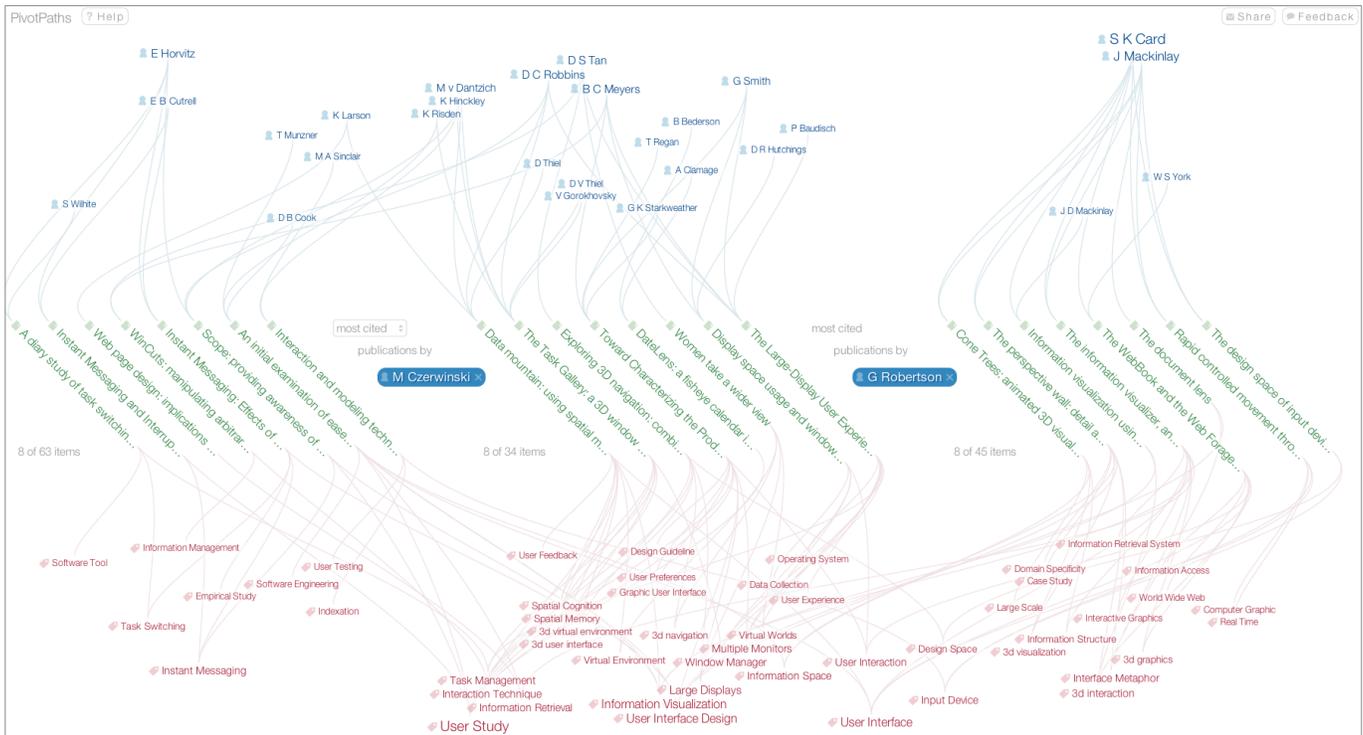


Fig. 8. Comparison Layout: Two anchors from the people facet, M Czerwinski and G Robertson, partition the resources into three groups.

added the sorting ‘Most cited (per year)’ as a time-normalized sorting that is less prone to favor old publications.

A simple *search* widget is provided on an empty canvas when no anchor is selected. The search automatically provides suggestions as soon as a few letters are typed in (see Figure 9). By typing the name of familiar person, resource, or concept, the viewer can quickly gain access to the information space with the selected entity as a starting point. The sorting of the results follows the three-way grouping between people, resources, and concepts on the visualization canvas. An interesting alternative to offering only the search functionality on the home screen, would be to make searching for entities optional by already displaying the most recent or popular resources and their faceted connections. This way the viewer could starting exploring the information collection without having to come up with any search terms.

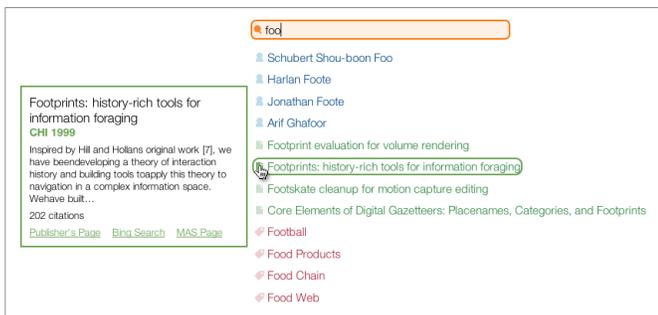


Fig. 9. When no anchor is selected, one can search for familiar people, resources, or concepts by their name.

The *details* of an element can be displayed by clicking on the icon positioned on the left side of the element’s label (see Figure 4, left). Originally, details were only shown when hovering over the icon, however, during the first observational study sessions several participants noted that they wished that they could pin the detail overlays. Making the details less transient also allowed to add links that could then be selected. The information that is displayed in the detail overlay

depends on the type of element. For example, an author’s detail overlay includes her full name, affiliation, photo, main research interests, and links to relevant websites (see Figure 1). Furthermore, the detail overlay includes the number of resources that an element is associated with. For example, the detail overlays of publications includes the title, conference/journal, year, abstract, links, and number of citations (see Figure 9). In the interviews, several researchers noted that they appreciated immediate access to the respective publication files in order to pursue their interest in a given topic and read it about in more depth. The database that we used exposed these source links, so we included them in the detail overlay. The detail overlays of keywords feature the first paragraphs from the corresponding Wikipedia articles.

By *hovering* the mouse pointer over an element, related elements in the display are highlighted. For example, in Figure 7, the mouse pointer hovers over ‘Sense Making’ which is associated with seven authors via four publications, as well as other keywords. The curves that link the currently hovered element to resources and then to the other facet’s values are highlighted. We decided not to highlight curves in the same facet region to not trigger too many display changes. When hovering over a resource, other referenced resources (e.g., by citations) are highlighted as well as curves going to the facet values associated with the hovered resource.

4.5 Making Pivot Transitions Comprehensible

Changing the selected anchor affects the displayed set of resources and facet values. We call these changes *pivoting* because there is always some overlap between subsequent screens. A minimal overlap of elements is ensured because the next anchor, at a given moment, is based on the current selection of resources. In this sense, one can compare the pivoting between anchors to using a pair of compasses on a map. Even though there is overlap between pivot operations, there can be many elements that are added, removed, and changed. The goal when designing the animated transitions was to make display changes as gradual, comprehensible, and predictable as possible (DG 3).

The animated transition between two visualization states is separated into four main stages: hovering, removal, change, and addition:

1. *Hovering*: Hover highlights can be seen as a preview of a pivot

operation. Depending on the type of sorting the elements that are brushed before an element is anchored remain in the display. They keep their highlight until after the change transition is over.

2. *Removal*: All resources and facet values that are not visible in the subsequent view are faded out by decreasing the opacity and then removing the (now invisible) element from the canvas.
3. *Change*: Elements included in both views are gradually transitioned into their new positions and sizes. If the previous anchor is included as a regular element in the subsequent view, it is transitioned into its new position and size.
4. *Addition*: Resources and facet values that are added anew are faded into the display in an animated transition.

Initially, these transitions (except hovering) were carried out at the same time, making it difficult to comprehend which element left, stayed, and appeared. In the interviews, participants suggested separating the transitions into stages to make them more comprehensible [9]. Currently the curves between entities do not animate, but fade out during the removal stage and fade back in after the addition stage.

4.6 Utilizing the Web Context

PivotPaths utilizes the Web as the primary context for exploring information spaces (DG 5). In particular, we considered the browser's features, the Web as the data source, and context for sharing.

PivotPaths encodes the visualization state in the URL and in human-readable form also in the browser title. This way we can use the browser's history to allow easy backtracking to a previous state. It is feasible to also incorporate elements from previous visualization states into the visualization canvas. In addition to the browser's back button, a viewer can save a given view for later revisits by simply bookmarking it with their browser. We added a sharing function, that essentially populates an email with an embedded link to the current view.

PivotPaths incorporates the larger context of the Web as it provides links to resources of web-based information spaces that can be followed within the same environment. For example, the academic collection contains links to authors' personal websites and full-text links to publications that can be accessed from within the browser. Furthermore, situating PivotPaths in the Web allows for straightforward sharing of visualization states in communities like Facebook and Twitter.

In terms of its technical realization, PivotPaths builds on the latest technological developments around HTML5, in particular by using JavaScript for client-side interaction logic, browser-native vector graphics (SVG) for the curves, and animated transitions for pivot changes (CSS). Furthermore, we designed PivotPaths for a range of devices and display sizes. We have made the number of resources that are selected at a time and the font-sizes of elements dependent on the browser window size. Having used standards-based Web technologies for graphics and interaction without relying on proprietary extensions, makes PivotPaths more portable by default.

The architecture of the PivotPaths system includes a server-side component that is implemented in JavaScript using the Node.js environment. While the layouts are generated dynamically in the browser, the server provides access to a custom MySQL database. However, it would be conceivable to run PivotPaths against an external API, given that it supports Boolean queries, returns sufficient faceted metadata for the resources, and has low latency to ensure high responsiveness.

5 PIVOTPATHS IN USE

Evaluating an interactive system designed to support open-ended exploration and encourage users to take a stroll in large information spaces is challenging for multiple reasons. The persona of the information flaneur [6] we are designing for does not lend itself well to controlled conditions often imposed by traditional quantitative user studies. The notion of completing specific tasks is inappropriate in this context and the very idea of asking people to use the system at a particular time and in particular settings may get in the way of the natural curiosity and casual information seeking PivotPaths aims at fostering.

In addition, it is difficult to define quantitative metrics of success for aspirations such as providing an aesthetic and enjoyable experience, exhibiting relations between facets and facilitating pivoting. For these reasons, we opted to deploy PivotPaths in a research institution and study it in two complementary ways. First, we logged *in situ* usage of PivotPaths and received free-form feedback. Second, we conducted observations and semi-structured interviews with a small number of people as they used PivotPaths in their offices. This provided richer insights about the reasons for some of the behaviors and assumptions about system capabilities.

5.1 Overview

We deployed PivotPaths as a web-based system at a research organization via their intranet. The sample dataset comprised 160,000 articles, 180,000 authors, and 20,000 keywords from four research areas in computer science: human-computer interaction, computer graphics, world wide web, and information retrieval. To recruit participants, we sent email about the availability of PivotPaths to several hundreds of people working in areas related to the dataset.

The intranet study attracted at most 290 participants (based on web cookies). If a participant used different browsers or machines they would count several times. We excluded sessions that did not involve any interactions with the system (possibly due to technical difficulties), resulting in a total of 211 actual-use sessions. On average, sessions lasted 7.1 minutes and included 5 pivot changes. We received eleven feedback emails from participants who used the system.

We also conducted a small observational study with semi-structured interviews. We observed four participants, from the same population described above, using PivotPaths. These observations allowed us to capture how visual encoding was perceived and gain a better understanding of their exploration patterns.

5.2 Results

In the following, we discuss the results from the deployment and observational study with regard to our primary design goals (DG 1-4).

Relating facets and items. Based on the usage during the deployment, participants used all three entities—publications (42%), authors (33%), and keywords (25%)—to navigate the information space (percentages of single-anchor views). The interviews confirmed that PivotPaths integrated these facets into coherent views. Numerous comments throughout the interview sessions highlighted that the interface provided “overview” and “context” for the papers, that the links showed the connections between the information, and that the highlighted groups of authors or keywords gave a good sense of key collaborators and general research areas. P4 did a comparison with his favorite digital library explaining how PivotPaths showed a lot of important information that was difficult to extract from other interfaces. He stated that seeing both authors and keywords related (on mouse over) gave even more context on the facets themselves.

Pivoting vs refining. Based on feedback emails, participants used PivotPaths to find colleagues (ranging from previous acquaintances to current collaborators), themselves, and the main papers and authors on a research topic. However, there was some confusion between filtering and pivoting. For example, email respondent E9 “wasn't quite sure of when it was filtering versus pivoting”. Similarly, all participants of the observational study first expected that PivotPaths was a search and refine interface, allowing them to narrow down to a smaller set of papers. In the beginning of the interview, P3 commented that it was frustrating to re-initiate the searches (pivot) instead of iteratively narrowing down the subset. However, as P3 led us through her exploration, she never expressed the need for such refinement and, on the contrary, she commented multiple times on views that appeared too “empty”. We observed that for 3 of the 4 participants, reaching a view that had few elements was somehow disappointing. Towards the end of her session, P3 reflected on PivotPaths: “I thought of it as a search tool but it is really more of a learning tool”. P4 also captured the essence of PivotPaths by expressing that pivoting (instead of filtering) ensured that he had some elements of interest as well as tangential ones he could investigate further.

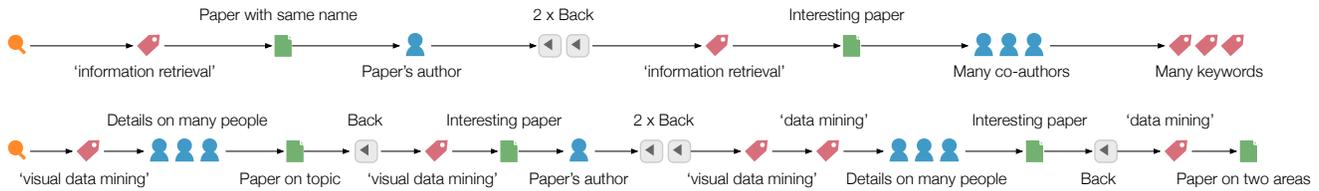


Fig. 10. P1's explorations includes regular use of the back button and successive selections of keywords and authors.

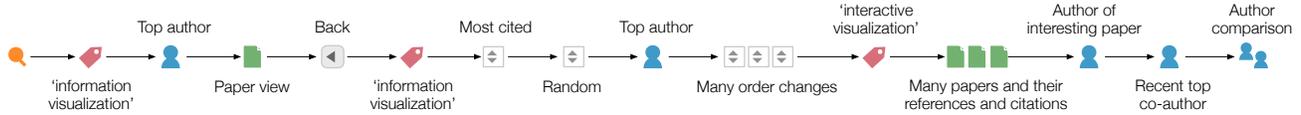


Fig. 11. P4's exploration with PivotPaths included different anchor types to learn about the different aspects of the information space.

Smooth navigation. All participants (deployment and observation) used all three facets to pivot from subset to subset. During the interviews, participants commented many times on the animations. P1 used animations to “follow where authors went” from one anchor view to another. P4 also used them to follow authors between two orderings. P1 and P4 clearly used the browser’s back button as a key exploration tool. They went back to anchor views, pivoting to different authors, papers, or keywords. All participants expressed that the interface offered an interesting way to navigate, and often commented on the different ways they could go about exploring information in one particular view.

Aesthetic qualities. Regarding the visual aesthetics of PivotPaths, the email feedback was the most interesting. Most participants who shared their experience with PivotPaths used enthusiastic language such as “awesome”, “cool”, and “lovely” [F1, F3, F5]. In particular, the combination of the visual arrangements and transitions between pivot operations were described to be very helpful [F7] and “Aesthetically fantastic [and] tuned perfectly” [F4]. On the flip side, participants also noted difficulties with the arrangement of elements (“it looks a bit messy” [F8]). Four of the eleven email respondents shared problems with some design decisions, particularly, regarding the legibility of angled and truncated paper labels.

In summary, our design goals were largely met: our participants were able to relate facets and items, appreciated pivoting after familiarizing themselves with it, and commented positively on the animated transitions and visual aesthetics of the PivotPaths interface. Besides several addressable design issues around legibility and layout, the results of the study also raise some deeper, more difficult questions. In particular, the established familiarity with search interfaces and their refinement model challenged our intent to support more open-ended, lateral, and casual exploration styles that we described as strolling. Despite this challenge our participants pursued exploratory strolls—two of which we describe in detail in the following.

5.3 Two Strolls

Figures 10 and 11 illustrate two strolls in the data performed by P1 and P4. We asked these two participants to learn about a new area they were interested in and identify a possible mentor they would want to work with. These two strolls illustrate how PivotPaths supports a variety of explorations styles.

P1 showcases how PivotPaths’ features can be discovered during its use and how they impact the style of exploration (see Figure 10). P1 first sought a key topic and a key paper in the area he was interested in as he would probably have done in a traditional search and refine interface. He found the people that had the most publications in these views and reviewed their respective authored papers. As he read titles, his pointer hovered over papers, highlighting related authors and keywords. As he realized that these clusters of keywords could give him a better idea of research areas that could be interesting, he changed his exploration style, pivoting from keywords to keywords (instead of authors) with the goal of identifying interesting papers as he went. Unsatisfied by his findings, P1 typed in a new search term and selected the

corresponding keyword. But this time, instead of engaging in multiple rapid pivoting operations, he spent more time exploring the present view, hovering over the different elements, accessing details for authors and papers. After carefully considering the entire view, he used it as a branching point for exploring a few papers of interest. Through this process, he finally identified a paper at the intersection of his two initial search terms and noted their authors as possible mentors.

P4 provides a good example of how people perform open-ended explorations with PivotPaths, seeking to answer multiple successive questions in a single session (see Figure 11). P4 started the session with a rather *random* exploration, selecting authors and papers as he went, modifying the paper selection criteria to find more salient elements to pivot from. When he discovered an author he knew of, the intent of his exploration became to *learn about his research*. He used the selection criteria to learn about what this person worked on in general and his most recent work. He identified a particularly interesting cluster of keywords. Using one of these keywords as an anchor, he aimed at identifying more precise ‘*lines of research*’. To do so, he proceeded in successively explorations of papers and their citations/references. He commented that such views gave him a good idea of “what is this line of research on”. Once he declared himself interested in a particular line of research, he quickly identified two researchers. To select *which author was most promising mentor*, he compared them in a side by side view. From this view, he noticed that the two researchers were in fact in two distinct areas and picked the most relevant one.

6 APPLICATIONS

In this section, we briefly describe applications of PivotPaths to information spaces other than academic publications: movies, video clips, and news. Our intent here is to explore the applicability of strolling in other domains and discuss limitations with our design of PivotPaths.

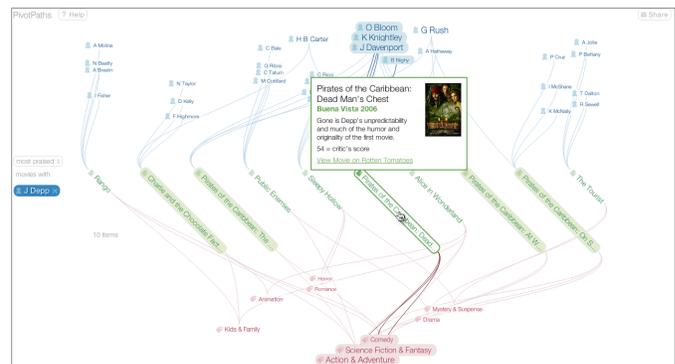


Fig. 12. Top grossing movies with Johnny Depp as one of the main actors. The detail overlay shows studio, release year, poster, critics’ consensus and score, and a link to the source page of this movie.

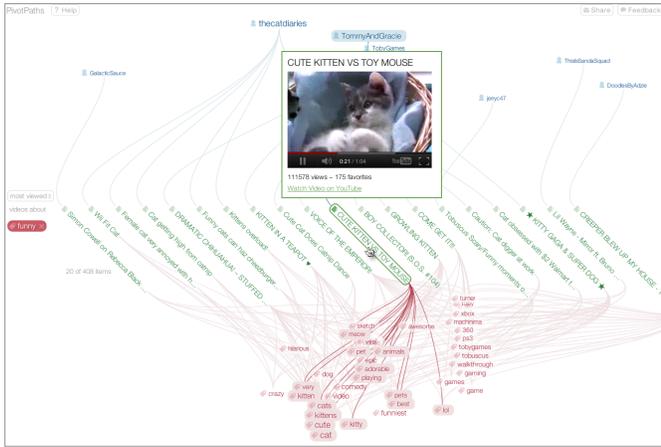


Fig. 13. Exploring ‘funny’ YouTube videos reveals the primacy of cat content and the respective content creators.

Compared to academic publications, movie collections are primarily for leisurely enjoyment making it an adequate data set for strolling. We chose the top grossing movies from the Internet Movie Database and retrieved movie data from the movie rating site Rotten Tomatoes. We use actors and genres as the facets, and sequels and remakes as the references among movies. Most websites offering information about movies consist of linked pages of top lists and movie details. In contrast, the PivotPaths canvas arranges and associates multiple movies, authors, and genres in a common view, making it possible to explore commonalities across movies in terms of the casts and genres as well as how they reference each other. For example, selecting the movies for a specific actor allows one to see the main types of movies they were in, e.g., ‘Action & Adventure’ for Johnny Depp (see Figure 12, bottom). While not the most effective encoding for rigorous analysis, the varying font sizes of labels make prevalent genres more salient, but still provide access to more unique terms, such as ‘Horror’ and ‘Romance’. These subtle differences in saliency can be seen as visual cues for an information flaneur to pivot towards resources that have more or less in common with the current selection.

Compared to the conventional production and distribution of Hollywood movies, the video community YouTube constitutes a parallel universe, where people upload, annotate, and consume video content with little (or no) access restrictions. Video clips can be accessed via search, user channels, and top lists of most popular videos. To explore YouTube videos with PivotPaths, we used the user names, tags, and responses—on YouTube, community members can upload videos as responses to another video. In addition, YouTube allows viewers to share videos as embedded elements into any Web page. We use this feature to integrate the video content into the detail overlays. Exploring ‘funny’ videos with PivotPaths reveals to our surprise that there are many videos about cats on YouTube (see Figure 13). When using PivotPaths for a while it also becomes apparent that something was missing from the PivotPaths interface, which uses primarily text and shows only visual content in the detail overlays. An open question is how a faceted visualization for strolling could better accommodate the contents of visual resources such as videos and images. Furthermore, it shows that the tags used on YouTube do not follow a certain taxonomy, and thus are much more ‘messy’ than the movie genres.

For news data, we used The New York Times newswire API that exposes a rich stream of news articles together with their authors, content facets, and links to other related articles. Using PivotPaths to explore these news stories, offers a qualitatively different experience from the typical layout of a newspaper or news website. First, by exposing the authors of news stories as one of the two facets in the PivotPaths interface it is possible to take a perspective on news that is rarely taken. We can identify the writers who have been covering certain issues, and see their articles in context of evolving news coverage. Second, articles concerned with diverse issues may appear on the same screen in Piv-

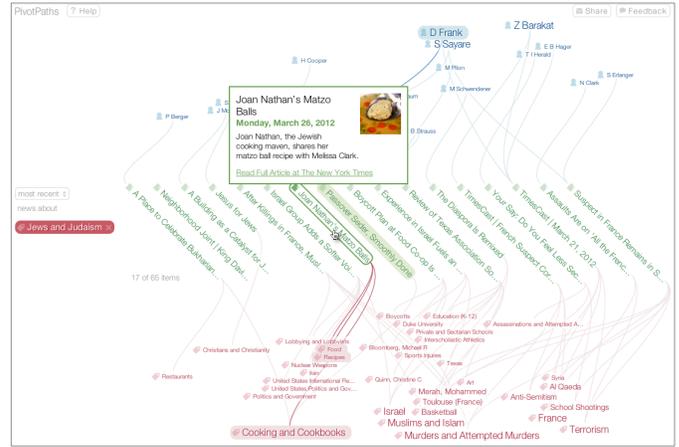


Fig. 14. Traversing articles from the New York Times. The anchor ‘Jews and Judaism’ (left) selects articles about cooking and crime (bottom).

otPaths when sharing a connection (see Figure 14). Newspapers typically categorize articles into separate sections that make it difficult to see such cross-relations. Compared to the video clips from YouTube, PivotPaths thrives on the semantic richness of the newswire API provided by the New York Times. The API actually distinguishes between different facets, such as news topic and geographic place, which we merged, but it would be interesting to give the viewer more choice in the selection of facets. Compared to the first page of newspapers and news websites, the initial screen of PivotPaths is more or less blank. Providing an entrance point that already selects some articles based on recency, popularity, or randomness can be beneficial for strolling.

7 CONCLUSION

As information spaces grow in size and richness, conventional interfaces for searching and browsing fall short of inviting for casual navigation, exposing how resources relate, and making display changes comprehensible. We explored how information visualization can be used to widen the angle on faceted collections and support *strolling*.

In contrast to visualization aimed at data analysis, we proposed strolling as the use of interactive visualizations for casual information navigation and open-ended exploration. *PivotPaths* is the result of an iterative design process aimed at supporting this notion of strolling. PivotPaths features a visualization canvas that positions and connects facets and resources to indicate rich data relationships and encourage gradual pivoting along them. Three layouts arrange resources and their facets around a given facet value, resource, or a pair of facet values. The visual representations are designed to encourage the viewer to explore the information space, make comparisons across facets and resources, and follow animated transitions between pivot operations.

We deployed PivotPaths in a large research institution for two weeks. Our observations show that we largely met our design goals towards strolling, and confirm that the system supports different exploration styles. Participants positively commented on the integration of multiple facets into coherent views, and we collected evidence that such views fueled their explorations. Participants who described their experience with PivotPaths as insightful, commented that the system led them to resources matching their interest and engaged them to pursue the exploration of more tangential information. The study also raised questions about the role of our established familiarity with search-and-refine interfaces that may impede the support of casual, more lateral forms of information exploration. Long-term studies of strolling interfaces are needed to investigate these issues further.

We applied PivotPaths to movies, video clips, and news to illustrate the generalizability of this technique to information spaces other than academic publications. In the future, we aim at deploying PivotPaths on a broader scale and a variety of information spaces. Our goal is to improve the strolling experience and gain a better understanding of how to support exploratory and casual navigation styles.

ACKNOWLEDGMENTS

We are very grateful to our colleagues at Microsoft Research for their advice and assistance throughout this project. We would also like to thank the referees for their constructive feedback and suggestions.

REFERENCES

- [1] M. J. Bates. The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, 13(5):407–424, 1989.
- [2] N. Cao, J. Sun, Y. Lin, D. Gotz, S. Liu, and H. Qu. FacetAtlas: Multifaceted visualization for rich text corpora. *TVCG: Transactions on Visualization and Computer Graphics*, 16(6):1172–1181, Nov/Dec 2010.
- [3] C. Chen. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3):359–377, 2006.
- [4] C. Collins, F. B. Viégas, and M. Wattenberg. Parallel tag clouds to explore and analyze faceted text corpora. In *VAST 2009: IEEE Symposium On Visual Analytics Science And Technology*, pages 91–98. IEEE, 2009.
- [5] M. Dörk, S. Carpendale, C. Collins, and C. Williamson. VisGets: Coordinated visualizations for web-based information exploration and discovery. *TVCG: Transactions on Visualization and Computer Graphics*, 14(6):1205–1212, Nov/Dec 2008.
- [6] M. Dörk, S. Carpendale, and C. Williamson. The information flaneur: A fresh look at information seeking. In *CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1215–1224. ACM, 2011.
- [7] M. Dörk, S. Carpendale, and C. Williamson. Visualizing explicit and implicit relations of complex information spaces. *Information Visualization*, 11(1):5–21, 2012.
- [8] C. Dunne, N. H. Riche, B. Lee, R. Metoyer, and G. Robertson. Graphtrail: Analyzing large multivariate, heterogeneous networks while supporting exploration history. In *CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012.
- [9] J. Heer and G. G. Robertson. Animated transitions in statistical data graphics. *TVCG: Transactions on Visualization and Computer Graphics*, 13(6):1240–1247, Nov/Dec 2007.
- [10] H. Kang, C. Plaisant, B. Lee, and B. Bederson. Netlens: iterative exploration of content-actor network data. *Information Visualization*, 6(1):18–31, 2007.
- [11] B. Lee, M. Czerwinski, G. Robertson, and B. Bederson. Understanding research trends in conferences using paperlens. In *CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1969–1972. ACM, 2005.
- [12] B. Lee, G. Smith, G. G. Robertson, M. Czerwinski, and D. S. Tan. FacetLens: exposing trends and relationships to support sensemaking within faceted datasets. In *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1293–1302. ACM, 2009.
- [13] G. Marchionini. Exploratory search: From finding to understanding. *Communications of the ACM*, 49(4):41–46, 2006.
- [14] J. Matejka, T. Grossman, and G. Fitzmaurice. Citeology: Visualizing paper genealogy. In *CHI 2012 Extended Abstracts (alt.chi): Extended Abstracts on Human Factors in Computing Systems*, pages 181–190, 2012.
- [15] G. Robertson, K. Cameron, M. Czerwinski, and D. Robbins. Polyarchy visualization: visualizing multiple intersecting hierarchies. In *CHI '02: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 423–430. ACM, 2002.
- [16] J. Stasko, C. Görg, Z. Liu, and K. Singhal. Jigsaw: Supporting investigative analysis through interactive visualization. In *VAST 2007: IEEE Symposium On Visual Analytics Science And Technology*, pages 131–138. IEEE, 2007.
- [17] M. Stefaner, T. Urban, and M. Seefeldler. Elastic lists for facet browsing and resource analysis in the enterprise. In *DEXA 2008: conference on Database and Expert Systems Application*, pages 397–401. IEEE, 2008.
- [18] A. Thudt, U. Hinrichs, and S. Carpendale. The Bohemian Bookshelf: Supporting serendipitous book discoveries through information visualization. In *CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1461–1470. ACM, 2012.
- [19] F. van Ham and A. Perer. "search, show context, expand on demand": Supporting large graph exploration with degree-of-interest. *TVCG: Transactions on Visualization and Computer Graphics*, 15(6), Nov/Dec 2009.
- [20] M. Wattenberg. Visual exploration of multivariate graphs. In *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 811–819. ACM, 2006.
- [21] C. Weaver. Multidimensional visual analysis using cross-filtered views. *VAST 2008. IEEE Symposium on Visual Analytics Science and Technology*, pages 163–170, 2008.
- [22] R. W. White, B. Kules, S. M. Drucker, and m.c. schraefel. Supporting exploratory search: Introduction. *Communications of the ACM*, 49(4):36–39, 2006.
- [23] M. Wilson, A. Russell, D. Smith, et al. mSpace: improving information access to multimedia domains with multimodal exploratory search. *Communications of the ACM*, 49(4):47–49, 2006.
- [24] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *CHI '03: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 401–408. ACM, 2003.