9 Science Visualization and Discursive Knowledge

Loet Leydesdorff

Introduction

There is a rich tradition in the visualization of scientific developments and historical events, accelerated recently by the growth and availability of large-scale datasets, software, and computational approaches (see Börner's [2010] *Atlas of Science* for a visual chronicling of this history and Börner, Chen, & Boyack, 2003, for a review). Science visualizations are often predicated on a map metaphor, so much so that the term *science visualization* has become interchangeable with *mapping science*. However, unlike the geographic map, science has no natural baselines (see Day, chapter 4, this volume). Scientific domains are not bounded like nations or states, particularly in interdisciplinary areas (Small & Garfield, 1985). Given the complexity of knowledge organization and interaction, there is always some degree of reductionism that must occur in order to project the knowledge space onto a two- or three-dimensional landscape. Furthermore, if the variable of time is included (e.g., if a scholar wishes to animate evolving dynamics), additional care must be taken to stabilize the representation so that the results can be captured as a mental map (Liu & Stasko, 2010; Misue, Eades, Lai, & Sugiyama, 1995).

The intellectual space of science can be mapped in terms of words (e.g., title words) and authors, and co-occurrences of these variables (Callon, Courtial, Turner, & Bauin, 1983; White & Griffith, 1982; White & McCain, 1998). At a higher level of aggregation, journal-journal citation relations—available from the *Science Citation Index*—have been used since the mid-1980s for mapping developments in and among disciplines (Doreian & Fararo, 1985; Leydesdorff, 1986; Tijssen, De Leeuw, & Van Raan, 1987). Small and others further developed the mapping of cocitations (e.g., Garfield, 1978; Small & Sweeney, 1985).

In this chapter, I argue that observable network relations organize the sciences under study into historical instantiations that can be statically visualized. The development of scholarly discourse, however, can be considered self-organizing in terms of fluxes of communication along the various dimensions that operate within different (e.g., disciplinary) codes. Over time, this adds evolutionary differentiation to the historical integration; a richer structure can process more complexity. Latent Semantic Analysis (LSA) focuses on these latent dimensions in textual data, and social network analysis (SNA) on the networks of observable relations. However, the two coupled topographies of information processing in the network space and meaning processing in the vector space operate with different (nonlinear) dynamics.

Multidimensional Scaling

Computer-aided visualization of multivariate data predated the advent of the personal computer and the Internet. Based on Kruskal (1964), scholars in psychometrics developed spatial representations of sets of variables by multidimensional scaling (MDS) (e.g., Kruskal & Wish, 1978; Schiffman, Reynolds, & Young, 1981). Among other forms of output, MDS can generate a two-dimensional map. The first large-scale MDS program ALSCAL ("alternating least square analysis") is still available in current versions of statistical packages such as SPSS.

Table 9.1 provides distances in terms of flying mileages among 10 American cities (SPSS, 1993; Leydesdorff & Vaughan, 2006). MDS enables us to regenerate the map from which these distances were obtained by minimizing the stress (S) in the projection (figure 9.1). Feeding this data into ALSCAL, for example, leads not surprisingly to an almost perfect fit ($S = 0.003$).

These data measure *dissimilarity*, because the larger the numbers, the further apart the cities are—that is, the more "dissimilar" they are in location. One can also use similarity measures for mapping, such as correlation coefficients. Options that might be added to a next generation of such maps include:

1. The ability to visualize the network of connections among the cities
2. Measures of distance other than Euclidean ones—for example, correlations in a multidimensional (vector) space provide a different topology
3. Groupings of nodes using different colors based on attribute values
4. The ability to scale nodes and links with the values of attributes; etc.

A large number of current network visualization and analysis programs provide these features and can be downloaded from the Internet.

Table 9.1
Flying mileages among 10 American cities

	Atlanta	Chicago	Denver	Houston	Los Angeles	Miami	New York	San Francisco	Seattle	Washington, D.C.
Atlanta	0	.	.	.	.	.	.	.	.	.
Chicago	587	0	.	.	.	.	.	.	.	.
Denver	1212	920	0	.	.	.	.	.	.	.
Houston	701	940	879	0	.	.	.	.	.	.
Los Angeles	1936	1745	831	1374	0	.	.	.	.	.
Miami	604	1188	1726	968	2339	0	.	.	.	.
New York	748	713	1631	1420	2451	1092	0	.	.	.
San Francisco	2139	1858	949	1645	347	2594	2571	0	.	.
Seattle	2182	1737	1021	1891	959	2734	2408	678	0	.
Washington, D.C.	543	597	1494	1220	2300	923	205	2442	2329	0

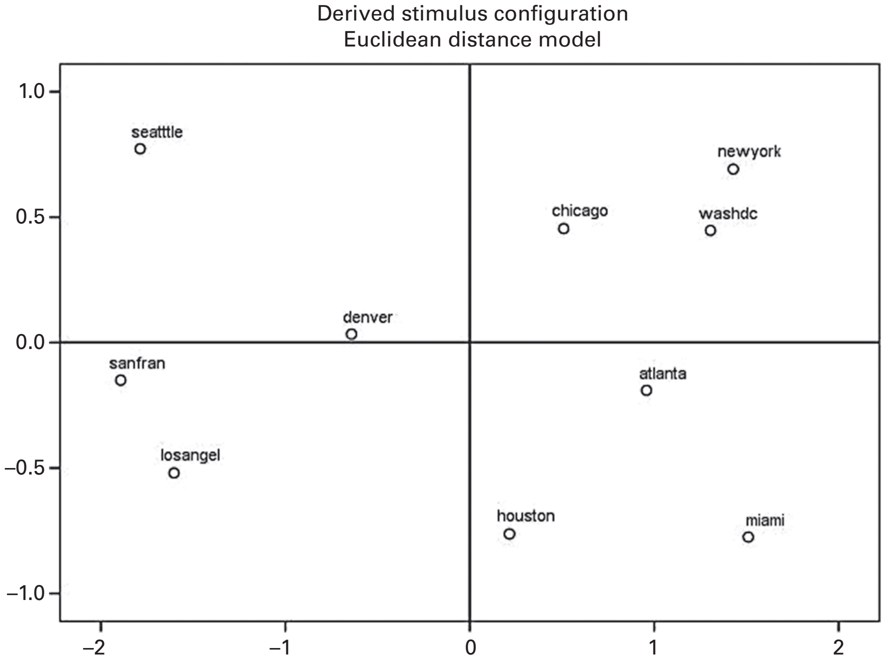

Figure 9.1

MDS mapping (ALSCAL) of 10 American cities using the distance matrix in table 9.1 (normalized raw stress = 0.003).

Graph Theory and Network Analysis

During the 1980s, graph theory emerged as a theoretical basis for network analysis. In the original programs (such as GRADAP) the links had to be drawn by hand. UCINet 2.0 (1984) provided the first network analysis program that integrated a version of MDS (MINISSA),[1] but the number of variables was at the time limited to 52: 26 uppercase and 26 lowercase characters could be indicated (Freeman, 2004). These programs allowed for the use of similarity measures other than Euclidean distances. For example, Leydesdorff (1986) used Pearson correlations to visualize factor structures in aggregated journal-journal citation matrices using UCINet 2.0.

Graphical interfaces became available during the 1990s with the further development of Windows (Windows 95) and the Apple computer. Pajek followed as a visualization and analysis tool for large networks in 1996 (De Nooy, Mrvar, & Batagelj, 2005). Pajek also allows for non-Western characters such as Chinese and Arabic (Leydesdorff & Jin, 2005).[2]

Pr

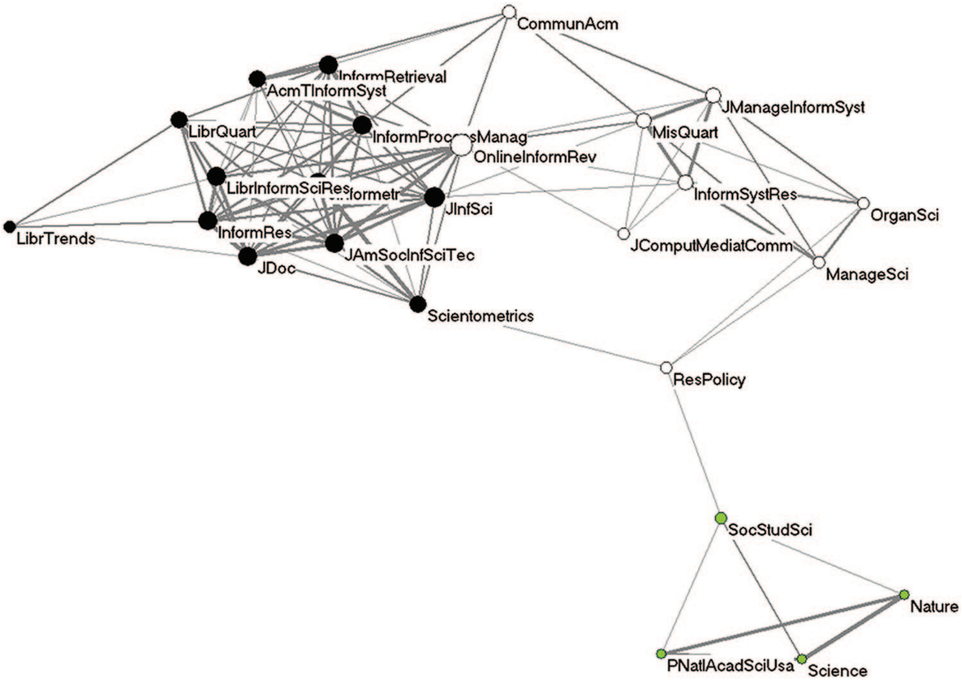

Figure 9.2

Twenty-five journals most cited by authors in *JASIST* during 2010; Kamada and Kawai (1989) used for the layout; node sizes proportional to degree centrality; node colors according to modularity (Q = 0.328); edge width proportional to *cosine* values (*cosine* > 0.2).

Figure 9.2 provides an example of the current state of the art: the aggregated citation network of the *Journal of the American Society for Information Science and Technology* (*JASIST*) as mapped in 2010. (These 25 journals are cited in *JASIST* to the extent of more than 1% of its total citations.) The matrix is analyzed using both Pajek and Gephi;[3] links are indicators of *cosine* similarities between the citing patterns of these journals; the vertices are colored according to the modularity algorithm (Q = 0.328; Blondel, Guillaume, Lambiotte, & Lefebvre, 2008), and sized according to their degree centrality (De Nooy et al., 2005).

Research Policy, positioned between the three components in this map, has accordingly the highest betweenness centrality (0.305). Although different in some details, both the factor analysis[4] and the modular decomposition classify *Research Policy* as belonging to the information systems group of journals (within this context). The visualization adds a network of relations among the nodes. As noted, one is able to use attributes of nodes and links in order to further enrich the visual.

Relational and Positional Maps of Science

Using MDS, one visualizes the variables as a system (e.g., a word-document matrix). In spatial terms, the words attributed to documents are considered as vectors that are vector-summed into a vector space (Salton & McGill, 1983). Given parameter choices (such as the similarity measure), the projection of the variables in MDS is deterministic. For example, the Euclidean distance between San Francisco and New York does not change depending on the intensity of the network relations (e.g., flights) between these two cities.

In network analysis, one is often as interested in a representation that uses the intensity of the relations as the distance on the map. For instance, two authors who frequently coauthor should be positioned next to each other in a coauthorship map. In this case, it is not the *correlations* among the distributions, but the *relations* among the nodes that are used for the mapping. Graph-analytic algorithms (e.g., Kamada & Kawai, 1989) optimize the network in terms of relations. The choice of starting point can be random, and each run may lead to a somewhat different outcome.

Let us compare the two approaches to optimizing the vector space versus the network topology. In figures 9.3 and 9.4, 43 title words are included that occurred more than 10 times among the 455 titles in the 2010 and 2011 volumes of *JASIST*. A five-factor solution in the underlying data matrix is used for coloring the nodes in the vector space (figure 9.3) and the network space (figure 9.4), respectively.

Factor 1, for example, is composed of the words *impact*, *factor*, *journal*, *citation,* and *source*. These (green-colored) words are grouped in both figures: they not only entertain strong relations to one another (figure 9.4), but also co-occur in similar *patterns* among the other title words in the sample (figure 9.3). Factor 4, however, with primary factor loadings for the words *effect*, *image*, *study*, *online*, and *behavior*, can more easily be distinguished in figure 9.3 than in figure 9.4. These words co-occur with other words in the set more diffusely, yet they form a latent dimension of the data.

In other words, there is no necessary relationship between co-occurrences in the observable network of relations, and correlations among co-occurrence patterns. The co-occurrence patterns can be mapped using the correlation coefficients among the distributions, whereas the values of co-occurrence relations provide us with a symmetrical (affiliations) matrix that can be visualized directly. In the latter case, one visualizes the network of observable *relations*, whereas in the former, one visualizes the latent structure in these data. For example, two synonyms may have (statistically) similar *positions* in a semantic map, but they will rarely co-occur in a single title.

Pr

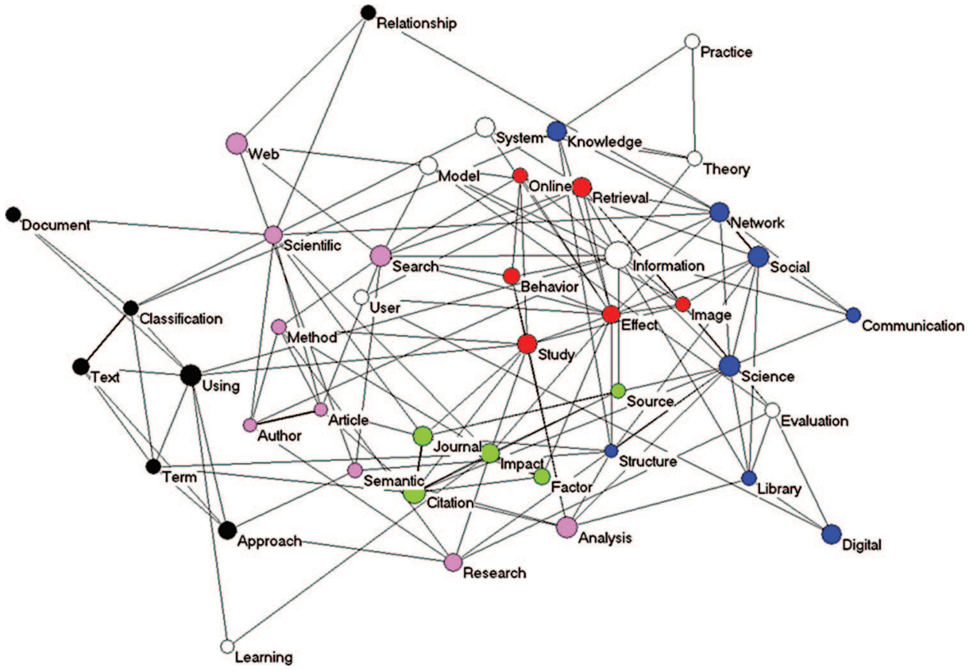

Figure 9.3

Cosine-normalized map of 43 words occurring more than 10 times during 2010 and 2011 in titles of *JASIST*. (*Cosine* ≥ 0.1; Kamada & Kawai, 1989.) The nodes are colored according to the five-factor solution of this network (Varimax rotated; SPSS), and scaled in accordance to their degree centrality.

These two perspectives on the data have led to two different research traditions in textual analysis and social network analysis, respectively. As noted, LSA focuses on the latent dimensions in the data, while SNA focuses on the observable relations in networks. In SNA, for example, eigenvector centrality—that is, factor loading on the first factor—can be used as an attribute of the nodes, whereas in LSA the factors (eigenvectors) in different directions organize the semantic maps (Landauer, Foltz, & Laham, 1998). The factor-analytic approach has been further developed using Singular Value Decomposition (SVD), whereas graph theory has provided an alternative paradigm for developing algorithms in SNA.

A star in a graph can be in the center of the multidimensional space, and therefore not load strongly on any of the dimensions. In figure 9.4, for example, the word *information*, which occurs 94 times in this set (followed by *citation*, 44 times), did not load positively on any of the five factors distinguished; this variable is factor-neutral and

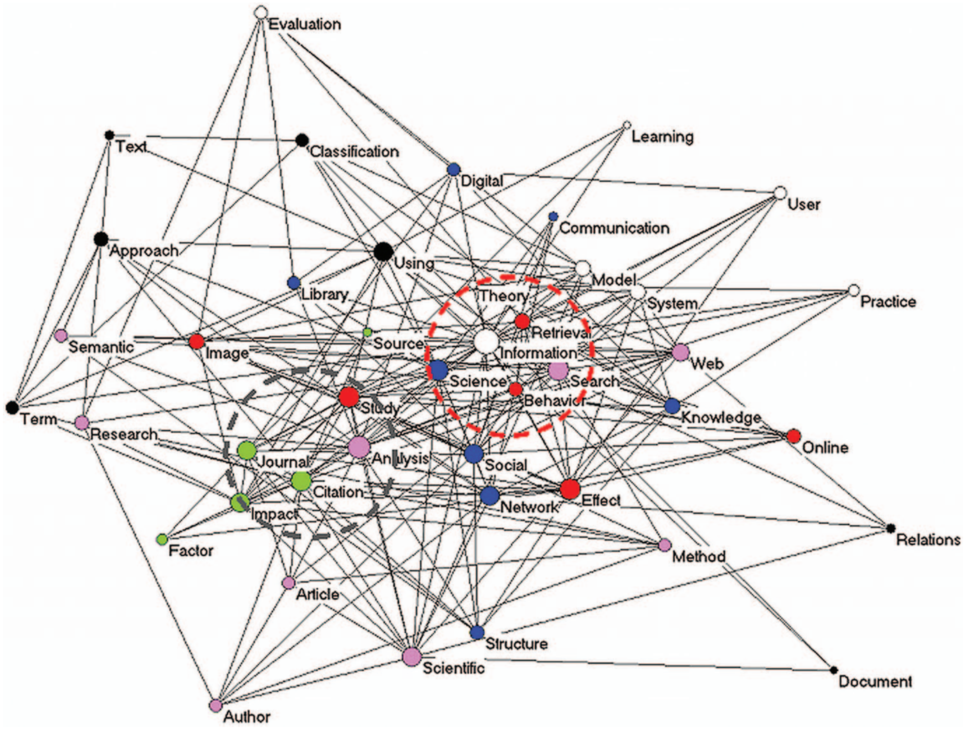

Figure 9.4
Co-occurrence map of 43 words occurring more than 10 times during 2010 and 2011 in titles of *JASIST*. (Co-occurrence values ≥ 2; Kamada & Kawai, 1989.) The nodes are colored according to the five-factor solution of this network (Varimax rotated; SPSS), and scaled in accordance to their degree centrality.)

therefore colored white. However, using the degree distribution for sizing the nodes in figure 9.4, i*nformation* has the highest degree, co-occurring with 37 of the 43 title words, followed by *analysis* with a degree of 33. A core set of words surrounding *information* (circled red in figure 9.4) belongs to the center of the field of the information sciences. *Citation* (Factor 1) and *analysis* (Factor 3) are part of a secondary grouping of the relations (gray circled).

Interpreting Science Visualizations

When a network is spanned in terms of relations, this process shapes an architecture in which all components have a position. The analysis of this architecture (that is, the set of relations) enables us to specify what the relations mean in the network as a system. For example, the word *information* was most central in this network (figure 9.4), but it was not colored in terms of having meaning in any of the relevant dimensions indicated at the systems level. Yet the word as a variable carries Shannon-type information (uncertainty; Shannon, 1948).

The graph-analytic approach informs us, as analysts, about the network of relations, but not about what these relations mean in terms of the discourse(s) under study. However, graph-theoretic concepts such as centrality also have meaning in social network analysis. The analyst's (meta) discourse can be distinguished from the communication among the words under study. The latter communication can represent scholarly discourses, political discourses, and media information.

Within each of these discourses, codes of communication can span dimensions that provide the communicated words with meaning. Both the developments in the observable networks (vectors) and the hypothesized dimensions (eigenvectors) can be theorized. The relations among nodes can be considered attributes of the nodes, but the dimensions of the communication are attributes of the links. SNA focuses on the positions of nodes in terms of vectors, whereas LSA focuses on the positions of links in terms of these next-order structures.

This scheme can be generalized: the relations among authors can also be considered as a system of links and therefore another semantic domain. Any system that can position its components as a system provides itself and its elements with meaning (Maturana, 1978). A discourse, for example, provides meaning to the words that are communicated.

The two perspectives of meaning processing and information processing can be considered feedback mechanisms operating on each other. The shaping of the networks of relations causes structures that can provide feedback evolutionarily as a next-order system on the networks of relations from which they emerge. Meaning is provided from the perspective of hindsight, but with reference to other possibilities ("horizons of meaning;" cf. Husserl, 1929/1973). The next-order meaning processing cannot continue without information processing; otherwise, the systems would no longer be historical. The historical instantiation can from this perspective be considered a retention mechanism of the semantic systems that evolve over time (Leydesdorff, 2011a).

The Network and the Vector Space

The multidimensional (vector) space can be regarded as a *system* of relations including interaction terms, and the network space as an *aggregate* of observable relations among nodes. One can also call the network relations first-order (being observable) and the vector space second-order because the latent dimensions of the system are not given but hypothesized—for example, in a factor-analytic model. Whereas observable variation is stochastic, latent structure is deterministic. The deterministic selection mechanism(s), however, can be expected to be further developed over time in parallel to the networks of relations because of the feedback mechanisms involved.

Accordingly, the systems view of MDS is deterministic, whereas the graph-analytic approach can also begin with a random or arbitrary choice of a starting point. Using MDS, the network is first conceptualized as a multidimensional space that is then reduced stepwise to lower dimensionality. At each step, the stress increases; Kruskall's stress function is formulated as follows:

$$S = \sqrt{\frac{\sum_{i \neq j} (\|x_i - x_j\| - d_{ij})^2}{\sum_{i \neq j} d_{ij}^2}} \qquad (1)$$

In this formula $||x_i - x_j||$ is equal to the distance on the map, while the distance measure d_{ij} can be, for example, the Euclidean distance in the data under study. As noted, one can use MDS to illustrate factor-analytic results (in tables), and in this case the Pearson correlation obviously provides the best match.

Spring-embedded or force-based algorithms can be considered a generalization of MDS but were inspired by the above-mentioned developments in graph theory during the 1980s. Kamada and Kawai (1989) were the first to reformulate the problem of achieving target distances in a network in terms of energy optimization. They formulated the ensuing stress in the graphical representation as follows:

$$S = \sum_{i \neq j} s_{ij} \quad \text{with } s_{ij} = \frac{1}{d_{ij}^2} (\|x_i - x_j\| - d_{ij})^2 \qquad (2)$$

Equation 2 differs from equation 1 by taking the square root in equation 1, and because of the weighting of *each* term in the numerator with $1/d_{ij}^2$ in equation 2. This weight is crucial for the quality of the layout but defies normalization with $\sum d_{ij}^2$ in the denominator of equation 1; hence the incomparability between the two stress values.

The ensuing difference at the conceptual level is that spring embedding is a graph-theoretic concept developed for the topology of a network. The weighting is achieved for each individual link. MDS operates on the multivariate space as a system, and hence refers to a different topology. In the multivariate space, two points can be close to each other without entertaining a relationship (Granovetter, 1973). For example, they can be close or distant in terms of the correlation between their *patterns* of relationships (cf. Burt, 1992).

Pr

In the network topology, Euclidean distances and geodesics (shortest distances) are conceptually more meaningful than correlation-based measures. In the vector space, correlation analysis (factor analysis, etc.) is appropriate for analyzing the main dimensions of a system. The *cosines* of the angles among the vectors, for example, build on the notion of a multidimensional space. In bibliometrics, Ahlgren, Jarneving, and Rousseau (2003) have argued convincingly in favor of the *cosine* as a nonparametric similarity measure because of the skewedness of the citation distributions and the abundant zeros in citation matrices. Technically, one can also input a *cosine*-normalized matrix into a spring-embedded algorithm. The value of (1 – *cosine*) is then considered a distance in the vector space (Leydesdorff & Rafols, 2011). In sum, there are a wealth of possible combinations in a parameter space of clustering algorithms and similarity criteria.

The Visualization of Heterogeneous Networks

The two coupled topographies of information processing in the network space and meaning processing in the vector space operate with different (nonlinear) systems dynamics (Luhmann, 1995). The historical dynamics of information processing in instantiations organizes the system, and thus interfaces with and tends to integrate, the (analytically orthogonal) dynamics along each eigenvector. The systems dynamics, however, can be considered self-organizing in terms of fluxes along the various dimensions—used as codifiers of the communication—and with potentially different speeds. This development over time adds evolutionary differentiation to the historical integration; a richer structure can process more complexity.

Integrating retention can be organized in dimensions other than differentiating expansion. For example, archives and reflexive authors historicize and thus stabilize the volatile networks of new ideas, metaphors, and concepts. Relations among words can be regarded as providing us with access to the variation, whereas cited references anchor new knowledge claims in older layers of texts (Lucio-Arias & Leydesdorff, 2009). Authors and institutions may provide historical stability because differences are reflected and locally integrated in communicative actions.

The textual domain provides us with options to combine these different layers in visualizations and animations. The sciences evolve as heterogeneous networks of words, references, authors, and at different levels of aggregation. The composing subdynamics, for example, of specialties and disciplines are not organized neatly in terms of specific variables, but in terms of configurations of variables, such as specific resonances among cognitive horizons (paradigms), social identities, and corpora of literature. The human beings involved (and their organizations) cannot be reduced to literature, and

$$\begin{array}{l|cccc} & \text{au}_1 & \text{au}_2 & \dots & \text{au}_m \\ \hline \text{doc}_1 & a_{11} & a_{21} & \dots & a_{m1} \\ \text{doc}_2 & a_{12} & \dots & \dots & a_{m2} \\ \text{doc}_3 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \text{doc}_n & a_{1n} & \dots & \dots & a_{mn} \end{array} + \begin{array}{l|cccc} & \text{w}_1 & \text{w}_2 & \dots & \text{w}_k \\ \hline \text{doc}_1 & b_{11} & b_{21} & \dots & b_{k1} \\ \text{doc}_2 & b_{12} & \dots & \dots & b_{k2} \\ \text{doc}_3 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \text{doc}_n & b_{1n} & \dots & \dots & b_{kn} \end{array} = \begin{array}{l|cccc} & \text{v}_1 & \text{v}_2 & \dots & \text{v}_{(m+k)} \\ \hline \text{doc}_1 & c_{11} & c_{21} & \dots & c_{(m+k)1} \\ \text{doc}_2 & c_{12} & \dots & \dots & c_{(m+k)2} \\ \text{doc}_3 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \text{doc}_n & c_{1n} & \dots & \dots & c_{(m+k)n} \end{array}$$

Figure 9.5
Two matrices for *n* documents with *m* authors and *k* words can be combined to a third matrix of *n* documents vs. (*m* + *k*) variables.

cognitive development can be considered a latent dimension emerging in networks of texts and people (Leydesdorff, 1998). This thesis of the heterogeneity of the technosciences was first proposed by authors in the semiotic tradition (Callon et al., 1983).

Because the different dynamics at interfaces within and between knowledge-based systems (such as science, technology, and innovation) are documented in texts, the texts can provide us with access to the different dimensions. In SNA, for example, these various dimensions of the data can be mapped as modalities. Another option for mapping hybrid networks was suggested by Leydesdorff (2010). All relevant variables can be attributed to (sets of) documents as units of analysis. The various asymmetrical matrices of *n* documents versus, for example, *k* words and *m* authors can be aggregated as visualized in figure 9.5.

The resulting matrix can be factor-analyzed or—using matrix algebra—transformed into a symmetrical affiliations matrix. In figure 9.6, 33 of the 36 coauthors of these same documents are positioned in a semantic map (as in figure 9.3). (Three other authors were not connected at *cosine* > 0.1.) I added a dashed circle around the coauthorship network of Mike Thelwall as an example. Other variables (e.g., cited references, institutional addresses, country names) can be made equally visible, and colored or sized accordingly.

Animation of the Visualizations

Can the maps for different years (or other time intervals) also be animated? Several network visualization programs are available that enable the user to smooth the transitions based on interpolations among the solutions at different moments in time. The dynamic problem is then reduced to a comparatively static one: the differences among maps for different years are assumed to provide us with a representation of the

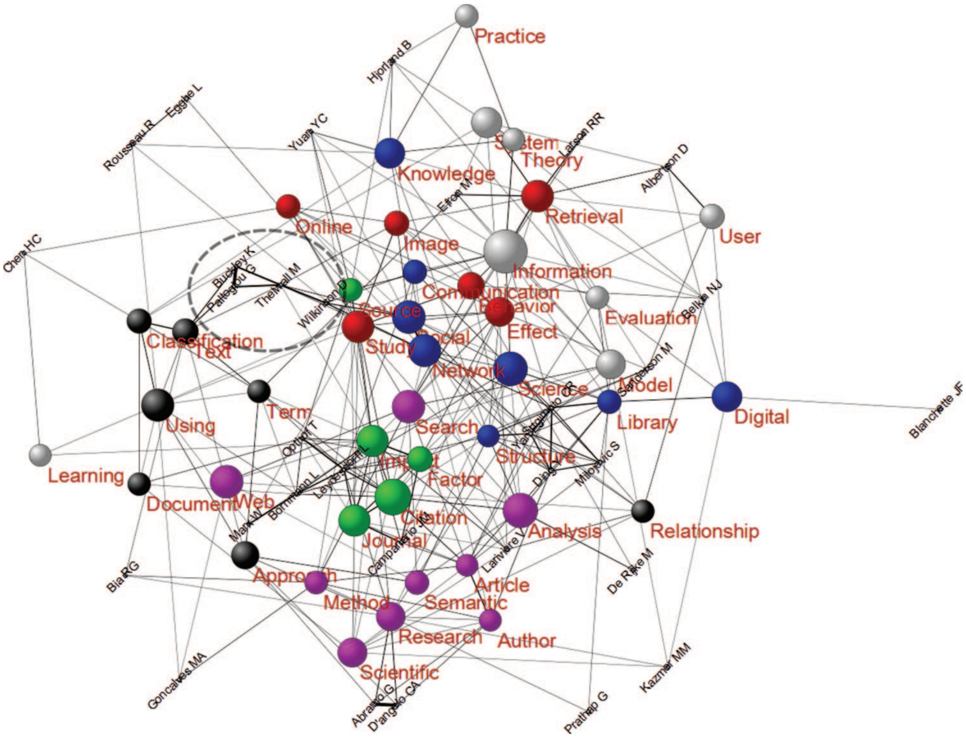

Figure 9.6

43 words (from figure 9.3) and 33 authors related at *cosine* > 0.1.

evolution of the system. However, the solution for each year is already an optimization of a higher-dimensional configuration into the two-dimensional plane. It can thus be difficult to distinguish between the development of the system and error.

An analytic solution of the system of partial differential equations provided by all the changing vectors and eigenvectors is impossible, and a numerical one computationally too intensive. Using MDS, however, Gansner, Koren, and North (2005) proposed minimizing not the stress, but the *majorant* of the stress, as a computationally more effective and methodologically more promising optimization. Baur and Schank (2008) extended this algorithm to layout dynamic networks (cf. Erten, Harding, Kobourov, Wampler, & Yee, 2004). The corresponding dynamic stress function is provided by the following equation:

$$S = \left[\sum_t \sum_{i \neq j} \frac{1}{d_{ij,t}{}^2} (\|x_{i,t} - x_{j,t}\| - d_{ij,t})^2 \right] + \left[\sum_{1 \le t < |T|} \sum_i \omega \|x_{i,t} - x_{i,t+1}\|^2 \right] \qquad (3)$$

In equation 3, the term on the left is equal to the static stress (in Equation 2), while the term on the right adds the dynamic component, namely the stress over subsequent years. This dynamic extension penalizes drastic movements of the position of node *i* at time *t* ($\vec{x}_{i,t}$) toward its next position ($\vec{x}_{i,t+1}$) by increasing the stress value. Thus, stability is provided in order to preserve the mental map between consecutive layouts (Liu & Stasko, 2010).

In other words, the configuration for each year can be optimized in terms of the stress in relation to the solutions for previous years and in anticipation of the solutions for following years. In principle, the algorithm allows us (and the dynamic version of *Visone*—available at http://www.leydesdorff.net/visone—enables us) to extend this method to more than a single time step. Using a single year in both directions, Leydesdorff and Schank (2008) animated, for example, the aggregated journal-journal citations in "nanotechnology" during the transition of this field at the end of the 1990s.[5]

Note that this approach is different from taking the solution for the previous moment in time as a starting position for a relative optimization. The nodes are not repositioned given a previous configuration, but the previous and the next configurations are included in the algorithmic analysis for each year. More recently, Leydesdorff (2011b) further elaborated this approach by projecting the eigenvectors as constructs among the variables into the animations.[6] Thus, one can make visible not only the evolution of observable variables, but also the evolution of latent structures. In principle, it would be possible to decompose the resulting stress into dynamic and static components.

Conclusion and Future Directions

The relations between semantic maps and social networks have been central to my argument because when visualizing the sciences as bodies of knowledge, the multimodal network of words, authors, etc., has to be specified. Discursive knowledge is communicated, and thus a network visualization is possible in different dimensions. However, knowledge can be considered a latent dimension of meaning processing in a network: discursive knowledge emerges in configurations of words, authors, references, etc., and can then be codified and institutionalized, for example, in journals, specialties, departments, and disciplines. The self-organization of the sciences in latent dimensions conditions and enables the observable relations in networks of authors, words, and citation relations.

The sciences are first shaped by the communicating agents, but textual communications can then develop a dynamic of their own as the communications are further

Pr

codified by theorizing. The sciences develop as systems of rationalized expectations in this codified dimension. However, the development of ideas leaves footprints in the texts (Fujigaki, 1998). The dynamics of texts and authors are different, and the dynamics of communication are (co)determined by the feedback from emerging knowledge dimensions. In figure 9.2, for example, the knowledge dimension was operationalized as three groups of journals belonging to different specialties.

The visualization of the sciences as a research program thus requires distinguishing among semantic maps, social networks, and the latent sociocognitive structures that can emerge on the basis of the interactions among people and texts. Three layers (people, texts, cognitions) coevolve in terms of observable variables and latent eigenvectors. Because of the next-order organization, the variables can be expected to interact among themselves and to shape and reproduce structures that can both recur on previous states and anticipate further developments of the system(s) (Luhmann, 1995; Maturana, 1978).

Visualization and animation of the sciences constitute an active research front in the development of the information sciences and bibliometrics. In the future, animations using multiple perspectives can be expected to replace models of multivariate analysis in which independent factors explain the data. Configurations of variables generate different synergies (Leydesdorff, Rotolo, & De Nooy, 2013). These implications follow from considering not only the communication of information, but also its meaning (Krippendorff, 2009; Leydesdorff, 2010). Horizons of meaning can be expected to generate redundancy—that is, new and more possibilities that change the value of existing ones.[7]

Animations enable us to capture different perspectives analogously as visualizations capture different arrangements of variables. The development of animations in the coupled layers of information and meaning processing can be expected to raise new questions for the further development of bibliometrics, network analysis, statistics, and relevant neighboring specialties.

Acknowledgments

I am grateful to Katy Börner for comments on a previous version, and to Thomson Reuters for access to relevant data.

Notes

1. MINISSA is an acronym for "Michigan-Israel-Nijmegen Integrated Smallest Space Analysis"; it became available around 1980 (Schiffman, Reynolds, & Young, 1981).

Pr

2. Pajek is a freeware program for network visualization and analysis available at http://vlado.fmf.uni-lj.si/pub/networks/pajek.

3. Gephi is an open-source program for network analysis and visualization, available at https://gephi.org.

4. Three factors explain 49.2% of the variance in this matrix.

5. Available at http://www.leydesdorff.net/journals/nanotech.

6. See http://www.leydesdorff.net/eigenvectors/commstudies.

7. The mutual information in three dimensions (μ*; cf. Yeung, 2008, pp. 59–60) among the three main factors structuring the coword network (figure 9.3) is –122.2 mbits, whereas this redundancy virtually disappears when the 33 coauthors are added to the network: $\mu^* = -7.0$ mbit (figure 9.6). For the social network among the 36 coauthors, this value of μ is positive. In other words, the coauthor network itself does not communicate meaning in this case (Leydesdorff, 2010, 2011b; Leydesdorff & Ivanova, in press).

References

Ahlgren, P., Jarneving, B., & Rousseau, R. (2003). Requirement for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*, *54*(6), 550–560.

Baur, M., & Schank, T. (2008). *Dynamic graph drawing in Visone*. Technical University Karlsruhe, Karlsruhe. Retrieved from http://i11www.iti.uni-karlsruhe.de/extra/publications/bs-dgdv-08.pdf.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics*, *8*(10), 10008.

Börner, K. (2010). *Atlas of science: Visualizing what we know*. Cambridge, MA: MIT Press.

Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science & Technology*, *37*(1), 179–255.

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.

Callon, M., Courtial, J.-P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Sciences Information. Information Sur les Sciences Sociales*, *22*(2), 191–235.

De Nooy, W., Mrvar, A., & Batagelj, V. (2005). *Exploratory social network analysis with Pajek*. New York: Cambridge University Press.

Doreian, P., & Fararo, T. J. (1985). Structural equivalence in a journal network. *Journal of the American Society for Information Science American Society for Information Science*, *36*, 28–37.

Pr

Erten, C., Harding, P. J., Kobourov, S. G., Wampler, K., & Yee, G. V. (2004). GraphAEL: Graph animations with evolving layouts. In G. Liotta (Ed.), *Graph drawing*. *Lecture Notes in Computer Science 2912*, 96–110. Berlin: Springer Verlag.

Freeman, L. C. (2004). *The development of social network analysis: A study in the sociology of science*. North Charleston, SC: BookSurge.

Fujigaki, Y. (1998). Filling the gap between discussions on science and scientists' everyday activities: Applying the autopoiesis system theory to scientific knowledge. *Social Sciences Information. Information Sur les Sciences Sociales*, *37*(1), 5–22.

Gansner, E. R., Koren, Y., & North, S. (2005). Graph drawing by stress majorization. In J. Pach (Ed.), *Graph drawing Lecture notes in computer science* (pp. 239–250). Berlin: Springer.

Garfield, E. (1987). Launching the *ISI Atlas of Science*: For the new year, a new generation of reviews. *Current Contents,* (1), 3–8.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Husserl, E. (1973). *Cartesianische Meditationen und Pariser Vorträge* [Cartesian meditations and the Paris lectures]. The Hague: Martinus Nijhoff. (Original work published 1929)

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, *31*(1), 7–15.

Krippendorff, K. (2009). Information of interactions in complex systems. *International Journal of General Systems*, *38*(6), 669–680.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*(1), 1–27.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2), 259–284.

Leydesdorff, L. (1986). The development of frames of references. *Scientometrics*, *9*(3–4), 103–125.

Leydesdorff, L. (1987). Various methods for the mapping of science. *Scientometrics*, *11*, 291–320.

Leydesdorff, L. (1998). Theories of citation? *Scientometrics*, *43*(1), 5–25.

Leydesdorff, L. (2010). Redundancy in systems which entertain a model of themselves: Interaction information and the self-organization of anticipation. *Entropy*, *12*(1), 63–79.

Leydesdorff, L. (2011a). "Meaning" as a sociological concept: A review of the modeling, mapping, and simulation of the communication of knowledge and meaning. *Social Sciences Information. Information Sur les Sciences Sociales*, *50*(3–4), 1–23.

Leydesdorff, L. (2011b). "Structuration" by intellectual organization: The configuration of knowledge in relations among scientific texts. *Scientometrics*, *88*(2), 499–520.

Pr

Leydesdorff, L., & Ivanova, I. A. (in press). Mutual redundancies in inter-human communication systems: Steps towards a calculus of processing meaning. *Journal of the American Society for Information Science and Technology*.

Leydesdorff, L., & Jin, B. (2005). Mapping the Chinese Science Citation Database in terms of aggregated journal-journal citation relations. *Journal of the American Society for Information Science and Technology*, *56*(14), 1469–1479.

Leydesdorff, L., & Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics*, *5*(1), 87–100.

Leydesdorff, L., Rotolo, D., & De Nooy, W. (2013). Innovation as a nonlinear process, the scientometric perspective, and the specification of an "innovation opportunities explorer." *Technology Analysis and Strategic Management*, *25*(6) 641–653.

Leydesdorff, L., & Schank, T. (2008). Dynamic animations of journal maps: Indicators of structural change and interdisciplinary developments. *Journal of the American Society for Information Science and Technology*, *59*(11), 1810–1818.

Leydesdorff, L., & Vaughan, L. (2006). Co-occurrence matrices and their applications in information science: Extending ACA to the web environment. *Journal of the American Society for Information Science and Technology*, *57*(12), 1616–1628.

Liu, Z., & Stasko, J. T. (2010). Mental models, visual reasoning and interaction in information visualization: A top-down perspective. *IEEE Transactions on Visualization and Computer Graphics*, *16*(6), 999–1008.

Lucio-Arias, D., & Leydesdorff, L. (2009). The dynamics of exchanges and references among scientific texts, and the autopoiesis of discursive knowledge. *Journal of Informetrics*, *3*(2), 261–271.

Luhmann, N. (1995). *Social systems*. Stanford, CA: Stanford University Press.

Maturana, H. R. (1978). Biology of language: The epistemology of reality. In G. A. Miller & E. Lenneberg (Eds.), *Psychology and biology of language and thought: Essays in honor of Eric Lenneberg* (pp. 27–63). New York: Academic Press.

McCain, K. W. (1990). Mapping authors in intellectual space: A technical overview. *Journal of the American Society for Information Science American Society for Information Science*, *41*(6), 433–443.

Misue, K., Eades, P., Lai, W., & Sugiyama, K. (1995). Layout adjustment and the mental map. *Journal of Visual Languages and Computing*, *6*(2), 183–210.

Salton, G., & McGill, M. J. (1983). *Introduction to modern information retrieval*. New York: McGraw-Hill.

Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Introduction to multidimensional scaling: Theory, methods, and applications*. New York: Academic Press.

Pr

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal,* 27, 379–423, 623–656.

Small, H., & Garfield, E. (1985). The geography of science: Disciplinary and national mappings. *Journal of Information Science*, *11*(4), 147–159.

Small, H., & Greenlee, E. (1986). Collagen research in the 1970s. *Scientometrics*, *19*(1–2), 95–117.

Small, H., & Sweeney, E. (1985). Clustering the Science Citation Index using co-citations I. A comparison of methods. *Scientometrics*, *7*(3–6), 391–409.

SPSS. (1993). *SPSS professional statistics 6.1*. Chicago: SPSS.

Tijssen, R., De Leeuw, J., & Van Raan, A. F. J. (1987). Quasi-correspondence analysis on square scientometric transaction matrices. *Scientometrics*, *11*(5–6), 347–361.

White, H. D., & Griffith, B. C. (1982). Authors as markers of intellectual space: Co-citation in studies of science, technology and society. *Journal of Documentation*, *38*(4), 255–272.

White, H. D., & McCain, K. (1998). Visualizing a discipline: An author cocitation analysis of information science, 1972–1995. *Journal of the American Society for Information Science American Society for Information Science*, *49*, 327–355.

Yeung, R. W. (2008). *Information theory and network coding*. New York: Springer.

Pr