Supporting the Evaluation of Intelligent Sources

Dirk Vriens
Radboud University of Nijmegen, The Netherlands

INTRODUCTION

To survive, organizations need to produce and process information about their environment, for instance, about customers, competitors, suppliers, governments, or all kinds of socioeconomic and technological trends. The process of obtaining this information is often called competitive intelligence (cf Fleisher & Blenkhorn, 2001; Kahaner, 1997; Vriens, 2004). An important stage in the competitive intelligence process is the collection stage. In this stage, one has to determine relevant sources, access them, and retrieve data from them (cf Bernhardt, 1994; Kahaner). For each data class, many possible sources are available, and determining the right ones is often difficult. Moreover, accessing sources and retrieving data may require a lot of effort and may be problematic (cf Cook & Cook, 2000; Fuld, 1995; Kahaner, 1997). In this chapter, we present a tool for supporting the effective and efficient use of sources: the source map. In essence, a source map links data classes to sources and contains information about these links. This information indicates the adequacy of sources in terms of ease of access, ease of retrieval, and usefulness of the retrieved data. A source map can support the selection of appropriate sources and it can support the assessment of the overall adequacy of available sources.

BACKGROUND

The process of competitive intelligence is often described as a cycle of four stages (the intelligence cycle; see Kahaner, 1997; Vriens, 2004). This cycle comprises (a) the direction stage (in which the organization determines about what aspects in the environment data should be collected), (b) the collection stage (where sources are determined and data are collected), (c) the analysis stage (in which the data are analyzed to assess whether they are useful for strategic purposes), and (d) the dissemination stage (where the data are forwarded to decision makers; Bernhardt, 1994; Gilad & Gilad, 1988; Herring, 1999; Kahaner, 1997; Sammon, 1986). The collection stage is considered to be the most time-consuming stage (e.g., Chen, Chau, & Zeng, 2002) and if it is not performed carefully, many difficulties arise (e.g., too much time spent on search, collection stage leads to irrelevant data, information overload; see, for example, Cook & Cook, 2000; Chen et al.; Teo & Choo, 2001; Vriens & Philips, 1999). For successfully carrying out collection activities, knowledge about what sources contain what kind of data and knowledge about how to approach these sources (metaknowledge regarding the collection of data) would be very helpful. This chapter presents a tool to structure and deal with this metadata: the source map.

To collect data about the environment one has to

1. identify possible sources,
2. judge the value of the source (in terms of different criteria; e.g., does it contain relevant data? What are the costs of employing this source? Is it reliable?), and
3. use value judgments to select the appropriate sources.

Many authors discuss Step 1 by pointing to a variety of available sources (cf Fuld, 1995; Kahaner, 1997; Sammon, 1986). Typical sources include the Internet, online databases, sales representatives, internal or external experts, CEOs, journals, tradeshows, conferences, embassies, and so forth.

The literature treats the valuation step more implicitly. It discusses distinctions regarding sources, such as open versus closed sources, internal versus external sources, or primary versus secondary sources (Fleisher & Blenkhorn, 2001; Kahaner, 1997). These distinctions implicitly refer to criteria used in the valuation of sources. The distinction of open versus closed sources implicitly refers to, for instance, criteria such as ease in collection or relevance. The distinction of primary versus secondary sources implicitly refers to the criterion of the reliability of the data. In our view, it is possible to value sources more precisely when the valuation criteria are stated explicitly and not implicitly in the form of these distinctions.

The selection step is even more elusive in literature (and practice). This step integrates value judgments to select appropriate sources for collecting the required data. Few methods seem to be designed for source selection.

In this article, we propose a tool to structure and support the valuation and selection of sources: the source map. This tool builds on Fuld's (1995) intelligence maps and knowledge maps (e.g., Davenport & Prusak, 1998).

The purpose of the source map is to help pin down the appropriate sources quickly and detect weaknesses in the available sources.

THE SOURCE MAP AS A TOOL FOR ASSESSING SOURCES

What is a Source Map?

A source map links data (or classes of data) to sources in such a way that the (most) appropriate sources can be selected for the collection of the requested data. If viewed as a matrix, the column entries may refer to data classes (e.g., products under development by competitor X) and the row entries to possible sources. Each column then indicates what sources may be used to gather the requested data (e.g., a patent database, economic journals, or the Internet site of competitor X). To determine what sources are (most) appropriate, the source map needs to contain information about criteria for appropriateness and their valuation. The cells in the source map (connecting the data classes to sources) should contain this information. To get this information, it should be clear (a) what the relevant criteria are, (b) how they can be given a value, and (c) how to integrate them into an overall judgment of the appropriateness of the sources. The next two sections deal with these issues.

Note that we treat the source map as a tool for supporting and structuring collection activities *given* the data classes. We assume that the data (classes) are already defined in the direction phase (the first phase of the intelligence cycle).

Criteria and Scores for Judging Sources

The criteria for assessing the appropriateness of sources link up with the three activities required to deal with sources. These activities are the following.

1. Accessing the source. Accessing means determining the exact location and approaching the source to prepare retrieval.
2. Retrieving (in interaction with the source) the data from the source.
3. Using the retrieved data in further processing (i.e., for the production of intelligence).

Referring to these activities, the appropriateness of sources depends on four dimensions: (a) ease of access, (b) ease of retrieval, (c) usefulness of the content of the retrieved data and processing ease, and (d) cost effectiveness. Below, we discuss criteria in these dimensions.

Criteria for Access and Retrieval

To assess the appropriateness of sources regarding access and retrieval, barriers in employing a source can function as criteria (cf Fuld, 1995; Davenport & Prusak, 1998). Examples of these barriers are as follows.

- A language barrier.
- A cultural barrier (i.e., a difference in culture between collector and source).
- An institutional barrier. In some (bureaucratic) organizations, it may be very hard to locate and approach certain people and documents.
- A personal barrier. Personal characteristics can make it difficult to approach and interact with someone.
- A geographical barrier. Some sources need to be dealt with on location.
- A technological barrier. Accessing some sources and retrieving data from them may sometimes be possible only by means of specific information and communications technology, requiring specific knowledge or skills.
- A fee barrier. For accessing some sources and/or retrieving data, a fee may be charged.
- A time barrier. For some sources, the response time may be very slow.
- A clarity barrier. This barrier refers to the effort one has to give to make sense of the data from the source. Factors that increase this barrier are the use of specific jargon and the lack of (requested) structure in the data.
- A stability barrier. This barrier refers to the stability of access to the source (some sources may cease to exist, some are not available at the expected moment, others may decide to stop providing their services, etc.).

In our view, these criteria can also be used to express the costs associated with using a particular source. We therefore prefer to deal with the above criteria, instead of using cost estimates that may be derived from them, because (a) it is difficult to translate the criteria into costs and (b) if only cost estimates are used, one loses information about the appropriateness of sources.

Using a barrier as a criterion to assess appropriateness, it can be scored on a five-point Likert scale where 1 means *very problematic* and 5 means *nonexistent*.

Criteria for the Use of Data

There are four criteria for assessing the appropriateness of sources regarding the use of the data for the production of intelligence. One of them is a processing criterion and three of them are content criteria.

The processing criterion refers to the ease of processing. This can be determined by the format in which the data are delivered; that is, does the source deliver the data in a format that can be used directly for the purposes of the collector or does it need reformatting? One may score this criterion on a five-point scale ranging from 1, *much reformatting needed*, to 5, *right format*.

The content criteria are completeness, reliability, and timeliness (cf O'Brien, 1998, for a summary of these criteria). When applied to the value of sources, these criteria mean the following.

- Completeness: The source can deliver all the data required to gain insight into the data class for which the source is used. This can, for instance, be measured in terms of the number of times the source was unable to deliver the requested data and/or the number of aspects of a data class for which the source could not provide data.
- Reliability: This refers to the reliability of the data from the source. It can be measured, for instance, in terms of the number of times the data from the source proved to be incorrect.
- Timeliness: The data from the source is up to date. It can be measured in terms of the number of times that the source delivered obsolete data.

In literature, one often finds relevance as an additional criterion to assess the content of data. Relevance then refers to the suitability of the data in gaining insight into the data class for which the source was used. However, relevance can be adequately expressed in terms of completeness, reliability, and timeliness. Completeness links the data provided by a source to the required data defined by the data class. Given the completeness, the data should further be correct and up to date to be relevant. Relevance, therefore, can be treated as an overarching concept, referring to the other three content criteria.

The content criteria can, again, be scored on a five-point scale, where 1 means *very incomplete*, *very unreliable*, and *very obsolete*, respectively, and 5 means *very reliable*, *very complete*, and *very timely*.

Content of Source Map Cells

The criteria for the appropriateness of sources and their scores should be put in the source map. To this end, each cell in a source map contains the following information (see also Figure 1).

1. General information about the source, consisting of the name of the source, the data-carrier (human, data or electronic) and (if known) the exact or default location.
2. Scores on the criteria for access, retrieval, content and processing of the (data from the) source.
3. Information about what data could not be delivered if the source was incomplete. This is useful for analyzing the appropriateness of the sources (see next section).
4. Remarks concerning one of the above aspects.

Figure 1. Content of cells in a source map (the shaded areas are not applicable)

Name:
Carrier:
Location:

	Language barrier	Cultural barrier	Institut. barrier	Personal barrier	Geogr. barrier	Techno. barrier	Time barrier	Fee barrier	Clarity barrier	Stability barrier
Access	1…5									
Retrieval										

Content:
Completeness: 1…5 If incomplete: What data could not be delivered?
Timeliness: 1…5
Reliability: 1…5

Process/format: 1…5

Remarks: …

Using the Source Map

A source map allows for two different uses. First, it is used to find appropriate sources for a particular data class. Second, it is used to assess the overall adequacy of the sources. For both types of use, it is necessary to compare the sources. In this section, we discuss how to compare sources and how to use this method for comparison for the two different uses.

Comparing Sources

Sources can be compared using a single criterion (e.g., which source scores highest on completeness?). It is also possible to integrate the values of (several) individual criteria and compare these integrated scores. To integrate these values into an overall score, we propose the following procedure.

1. Define two classes of criteria: efficiency criteria and effectiveness criteria. The class of efficiency criteria consists of the access criteria, the retrieval criteria, and the ease-of-processing criterion. The class of effectiveness criteria consists of the criteria completeness, reliability, and timeliness.
2. Estimate weights for the criteria in the two classes. A possible way of determining the weights of the individual criteria is to have CI professionals produce a rank order of the criteria (in each of the two classes) expressing their ideas about the relative relevance of the criteria. Next, one could discuss the results and produce one rank order for each class. (This procedure could be supported by groupware, such as Group Systems; cf Nunamaker, Dennis, Valacich, Vogel, & George, 1991).
3. Compute, for each source, the overall scores for the two classes. For both classes, we suggest taking the weighted average score for the given criteria. To compute these scores, the scores on the individual criteria should be available. These scores might be obtained initially by asking CI professionals. From that moment on, they should be evaluated every time a source is used and updated when necessary.

Finding Appropriate Sources

The most straightforward use of the map is to find out what sources are available for a particular data class. A step beyond merely enumerating available sources is to give a judgment about the appropriateness of these sources in terms of the criteria presented in the previous section. To this end, we use the efficiency and effectiveness scores of the sources. For a particular data class, all the available sources can be plotted regarding these two scores (see Figure 2).

The figure states that Source 5 scores best on effectiveness, Source 4 best on efficiency, and Source 1 scores lowest on both classes of criteria.

Figures like the above can help in analyzing the appropriateness of a source for a particular data class. As a general heuristic for ranking the sources, we suggest that sources in the upper right quadrant should be preferred to those in the lower right quadrant, and these should be preferred to the ones in the upper left quadrant. Sources in the lower left quadrant should probably be discarded.

Sources that come up as appropriate should be checked for completeness. If they are complete, they can be added to the list with preferred sources. If they are incomplete, it is necessary to find out if there are sources that can compensate for this lack. To this end, information about what data the source is unable to deliver can be used. This information directs the search for an appropriate compensating source.

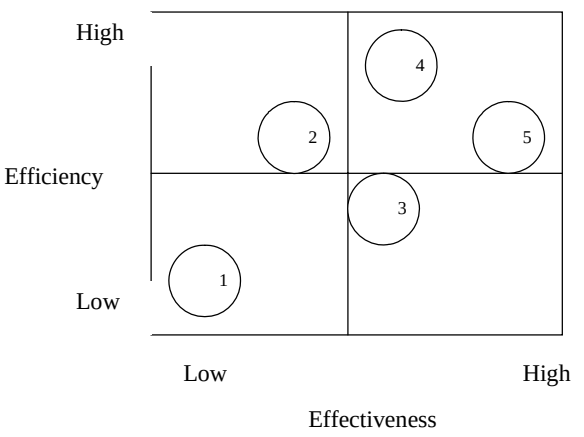
For sources that score high on effectiveness but low on efficiency (lower right quadrant), it should be examined (a) whether the relevance of the data class makes the effort (and costs) worthwhile and/or (b) whether measures can be taken to make the use of the source more efficient, for example, the efficiency of scale in gathering data (a subscription to an often-used online database).

For sources that score in the two left quadrants, it can be established what exactly causes the score. Dependent on the outcome of this investigation, it may be decided to stop using the source.

Assessing the Adequacy of Sources

To judge the overall adequacy of the sources, the map may help in answering the following questions.

Figure 2. Scores of five sources regarding their appropriateness for a particular data class (see text)



1. Do the sources cover all data classes?
2. Do we have adequate sources for the required data classes? If some data classes only have sources that have scores in the lower left quadrant of Figure 2, problems may arise. If a rank order of the data classes regarding their relevance exists, one can also establish whether the most relevant data classes are covered by appropriate sources.
3. Do we have enough different sources for the (most important) data classes? This question refers to the flexibility in collecting data. If a source is suddenly unavailable, one needs to have adequate alternatives. It is also useful to have different sources for the purpose of validating the data.

Answers to these questions help intelligence officers to identify weaknesses in the available sources and direct their efforts to repair them.

Implementing a Source Map

To build, maintain, and use a source map does not require exceptional resources. IT applications for implementing the map range from sophisticated applications to simple solutions. An example of a simple solution is an implementation of the map by means of Microsoft Excel sheets. However, it is also possible to use more sophisticated application, for instance, Web-based applications of the map. Making the map available via an intranet, for instance, can enhance its use and maintenance. In addition to these technological issues, it is important to define and allocate tasks and responsibilities regarding maintenance and use of the map. Finally, data collectors should be motivated to use the map to define their search strategies. In our experience, data collectors see the benefits of a good map and will be inclined to use and maintain it.

FUTURE TRENDS

To aid intelligence officers in their task to evaluate sources, the source map was introduced. Given the increasing need for organizations to collect data about their environment, it can be expected that the need for tools to evaluate sources (like the source map) will also increase. In order to deal with this, information technology tools may be tailored to support the implementation of source maps and the process of keeping them up to date (see, for instance, Philips, 2004).

CONCLUSION

To produce actionable intelligence, the efficient and effective use of sources is imperative. However, up until now, little attention has been paid to supporting the selection of sources. In this paper, we deal with this omission by presenting the source map as a support tool. Properly implemented source maps can be valuable instruments in the support of collection activities. In our view, they can aid in both the everyday use of sources and in the assessment of the overall adequacy of available sources.

REFERENCES

Bernhardt, D. C. (1994). I want it fast, factual, actionable: Tailoring competitive intelligence to executive needs. *Long Range Planning*, *27*(1), 12-24.

Chen, H., Chau, M., & Zeng, D. (2002). CI-spider: A tool for competitive intelligence on the Web. *Decision Support Systems, 34*, 1-17.

Cook, M., & Cook, C. (2000). *Competitive intelligence.* London: Kogan Page.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge.* Boston: Harvard Business School Press.

Fleisher, C. G., & Blenkhorn, D. L. (2001). *Managing frontiers in competitive intelligence.* Westport: Quorum, CT.

Fuld, L. M. (1995). *The new competitor intelligence.* New York: Wiley.

Gilad, B., & Gilad, T. (1988). *The business intelligence system.* New York: Amacon.

Herring, J. P. (1999). Key intelligence topics: A process to identify and define intelligence needs. *Competitive Intelligence Review, 10*(2), 4-14.

Kahaner, L. (1997). *Competitive intelligence.* New York: Touchstone.

O'Brien, J. (1998). *Introduction to information systems* (2nd ed.). New York: McGraw-Hill.

Nunamaker, J. F., Dennis, A. R., Valacich, J. S., Vogel, D. R., & George, J. F. (1991). Electronic meetings to support group work. *Communications of the ACM, 34*(7), 40-61.

Philips, E. A. (2004). Building a competitive intelligence system: An infrastructural approach. In D. Vriens (Ed.), *Information and communication technology for competitive intelligence* (pp. 227-247). Hershey, PA: IRM Press.

Sammon, W. L. (1986). Assessing the competition: Business intelligence for strategic management. In J. R. Gardner, R. Rachlin, & H. W. Sweeny (Eds.), *Handbook of strategic planning* (pp. 4.12-4.19). New York: Wiley.

Teo, T. S. H., & Choo, W. Y. (2001). Assessing the impact of using the Internet for competitive intelligence. *Information & Management, 39*, 67-83.

Vriens, D. (2004). *Information and communication technology for competitive intelligence.* Hershey, PA: IRM Press.

Vriens, D., & Philips, E. A. (1999). Business intelligence als informatievoorziening voor de strategievorming. In E. A. Philips & D. Vriens (Eds.), *Business intelligence*. Deventer, Netherlands: Kluwer, 11-44.

KEY TERMS

Collection Stage: Stage of the intelligence cycle. In this stage, sources regarding the required environmental data are located and accessed, and the data are retrieved from them.

Competitive Intelligence: In the literature, two definitions are used: a product definition and a process definition. In the product definition, competitive intelligence is defined as information about the environment, relevant for strategic purposes. The process definition highlights producing and processing this environmental information. Process definitions often refer to the intelligence cycle.

Intelligence Cycle: Cycle of four stages (collections of intelligence activities). The stages are direction (also referred to as planning, in which the strategic information requirements are determined), collection (determining sources and retrieving data), analysis (assessing the strategic relevance of data), and dissemination (of the intelligence to strategic decision makers).

Source: Something or someone containing data and from which the data can be retrieved. Many distinctions regarding sources are given in the competitive intelligence literature, for instance, open versus closed sources, primary versus secondary sources, internal versus external sources, and a distinction referring to the carrier of the data (human, electronic, or paper).

Source Evaluation: The process of assessing the efficiency and effectiveness of a source or several sources, given certain criteria. The result of this process can be (a) a judgment about the usefulness of a particular source for collecting data and/or (b) an insight into the relative usefulness of all available sources. See also "Source map."

Source Identification: Identifying suitable sources (i.e., efficient and containing the relevant data) given a certain data need. See also "Source map."

Source Map: A source map is a matrix linking data classes to sources. In the cells of the matrix, the sources are valued according to different criteria (e.g., accessibility, costs, timeliness of the data, etc.).