



## English version

Dictionary compiled by the ECLECTIK team  
Meaning-Text Linguistics Observatory (OLST)  
University of Montreal

by Marie-Claude L'Homme  
Translated and adapted by Aryane Bouchard and Josh  
Holden

September 2009

## Table of contents

<b>1. The Team .....</b>	<b>3</b>
<b>2. Goals of the DiColInfo Dictionary .....</b>	<b>4</b>
<b>3. Choice of Entries .....</b>	<b>4</b>
<b>4. Originality of the DiColInfo .....</b>	<b>5</b>
<b>5. Corpora and Natural Language Processing .....</b>	<b>6</b>
<b>6. How to Search for Terms in the DiColInfo.....</b>	<b>6</b>
<b>7. Entries.....</b>	<b>8</b>
<b>7.1 The Headword .....</b>	<b>9</b>
<b>7.2 Entry Status .....</b>	<b>10</b>
<b>7.3 Actantial Structure.....</b>	<b>11</b>
7.3.1 Semantic Roles.....	11
7.3.2 Typical Actants .....	12
7.3.3 Linguistic Realizations of Actants.....	13
<b>7.4 The Definition .....</b>	<b>14</b>
<b>7.5 Contexts .....</b>	<b>14</b>
7.5.1 Annotated Contexts .....	14
<b>7.6 Related Terms.....</b>	<b>17</b>
7.6.1 Synonyms and Variants.....	17
7.6.2 List of Paradigmatic and Syntagmatic Lexical Relations .....	17
7.6.3 Lexical Relations and Actantial Structure .....	19
7.6.4 Triple Encoding of the Lexical Relations .....	19
<b>7.7 Additional Information .....</b>	<b>21</b>
<b>7.8 Administrative Section.....</b>	<b>21</b>
<b>8. French language terms .....</b>	<b>22</b>
<b>9. Current State of the DiColInfo .....</b>	<b>22</b>
<b>10. Obtaining the DiColInfo for Research Purposes .....</b>	<b>22</b>
<b>11. To learn more.....</b>	<b>23</b>

## **1. The Team**

The following people participated in the DiColInfo team over various phases by selecting terms, editing articles or refining the Natural Language Processing (NLP) component.

### **Director**

Marie-Claude L'Homme

### **Editors – English-language Version**

Aryane Bouchard (2008)  
Louis-Philippe Dargis (2006-2007)  
Janine Pimentel Mendonça (2008)  
Marie-Ève Paquette (2008-2009)

**Acknowledgements:** We would like to thank Stéfan Popovic, who designed our first Web site. The new online version of the DiColInfo was encoded in XML format in collaboration with Guy Lapalme of the Applied Computational Linguistics Research Group. The XML version and the research prototype were designed by Benoît Alain, Patrick Baril-Robichaud, Guy Lapalme and Vincent St-Amour. Corinne Thirioin-Dupuis and Benoît Robichaud also helped design the online version. The team would also like to thank Alain Polguère, who's explanations helped to clarify many aspects of the DiCo, on which the DiColInfo is based.

The development of the DiColInfo was partially funded by a grant from the Fonds Québécois de la Recherche sur la Société et la Culture (FQRSC). The portion covering the annotation of contexts was funded by Canada's Social Sciences and Humanities Research Council (SSHRC).

## 2. Goals of the DiColInfo Dictionary

DiColInfo is an abbreviation of *Dictionnaire fondamental de l'informatique et de l'Internet* (Dictionary of Computer Science and Internet Terms). As such, it is a specialized dictionary listing and explaining the myriad connections between terms from various domains of computer science as well as Internet usage. For example, it could enable a user to answer the following questions:

- Q: What is the term for the “author” of a computer program?
  - (A: *programmer*)
- Q: What are the words that typically denote an instrument used for browsing the Internet?
  - (A: *browser, navigator*)
- Q: What does one stores on a hard drive?
  - (A: *files, data, software, programs*)
- Q: What are the most common verbs for uses of the Web?
  - (A: *browse, surf*)

The DiColInfo is aimed at readers who wish to improve their mastery of computer-related terminology. Readers have the option of viewing it at two different levels of detail, either without the technical metalanguage (practical for users seeking advice on the use and combination of terms) or displaying this metalanguage, which may be of use to linguists, lexicographers and terminologists seeking direct access to a useful metalanguage which they can apply in their own work.

## 3. Choice of Entries

This dictionary is concerned with the basic terms related to computer science and the Internet. By *term* we intend any lexical unit whose meaning may be associated with one of the chosen specialized domains. We therefore reject lexical items unrelated to computer science, even if they are frequently used in that field. By *basic* we would like to highlight the fact the focus of the DiColInfo is on those terms users are most likely to encounter in computer-related texts. To that end, we avoid terminology used only in narrow sub-fields of computer science.

The terms in the DiColInfo belong to various parts of speech including nouns (COMPUTER, PRINTER, VIRUS), verbs (BROWSE, INFECT, VIEW), adjectives (REMOTE, ROBUST, VIRTUAL) and adverbs (REMOTELY). The DiColInfo repertoire also contains fixed phrases, be they nominal (such as READ/WRITE HEAD), verbal (LOG ON), adjectival or adverbial.

The terms we have retained are diverse, referring to concrete objects (FILE, HARD DISK, LAPTOP), human participants (USER, PROGRAMMER), representations (BIT, DATA), activities (ALLOCATION, CONFIGURE, DELETE), properties (COMPATIBLE, FREE), or units of measure. Finally, the dictionary contains some proper names and acronyms (INTERNET, WEB), which are described as if they were common nouns, i.e. they are used with a determiner and have their own combinatorial properties.

For now, we offer no prescriptive judgments about these terms or their use. All terms were considered provided they respected our selection criteria, part of which has been alluded to above.

#### **4. Originality of the DiCoInfo**

The DiCoInfo provides a description of the linguistic behavior of terms, particularly their lexical semantic properties. This dictionary differs from the majority of specialized dictionaries as well as from terminology databases in that it provides rich detail about the meaning and the grammatical and phraseological profile of the terms (actantial structure, combinatorial properties, semantic relations), but avoids listing extensive encyclopedic information about the related complex concepts.

The DiCoInfo therefore features the following innovations:

- No previous computer science dictionary is concerned exclusively with the linguistic behavior of terms.
- Few specialized dictionaries invoke lexical semantics in their entries. The DiCoInfo draws on the Meaning-Text Theory approach to lexicography, specifically the *Dictionnaire explicatif et combinatoire* (Mel'čuk *et al.* 1984-1999, 1995) and its application as the *DiCo* (Polguère 2003) and the *DiCouèbe* (Jousse et Polguère 2005). Here, however, the model is applied specifically to terms, i.e. those lexical units whose meanings are related to a particular specialized domain.
- Few specialized dictionaries provide a similarly exhaustive description of the lexico-semantic properties of the terms. The DiCo describes the Actantial structure of each term, the lexical relations between the headword and semantically related words. This covers paradigmatic lexical relations (synonyms, antonyms, etc.) as well as the syntagmatic relations, including the headword's collocates. The DiCoInfo not only lists these related terms, but additionally provides a short description of the precise lexical relation between the head word and the related term.

## **5. Corpora and Natural Language Processing**

The description of the terms portrayed in the DiColInfo is derived primarily from an examination of their behavior in a specialized corpus. A 1 million word corpus was compiled to that end. This corpus comprises texts about the Internet, computer networks and programming, micro-computing and operating systems. Most of the texts are from the genre of didactic materials. The lexicographers then supplemented their initial description of the terms by consulting previous specialized dictionaries or documents available online.

Some parts of our methodology rely on natural language processing. Here particularly worthy of mention is our approach to the selection of terms, partly based on the calculation of lexical specificity (see Drouin 2003 and Lemay et al. 2005) and our analysis of certain semantic relations between terms (see Claveau et L'Homme 2004, 2005a, 2005b ; 2006).

## **6. How to Search for Terms in the DiColInfo**

The DiColInfo can be consulted in one of two ways. The first option is to browse alphabetically ordered lists of terms, as in Figure 1 below. The user may click on the terms to directly access their description. This version is updated daily.

### **English Terms**

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [Z](#)

A

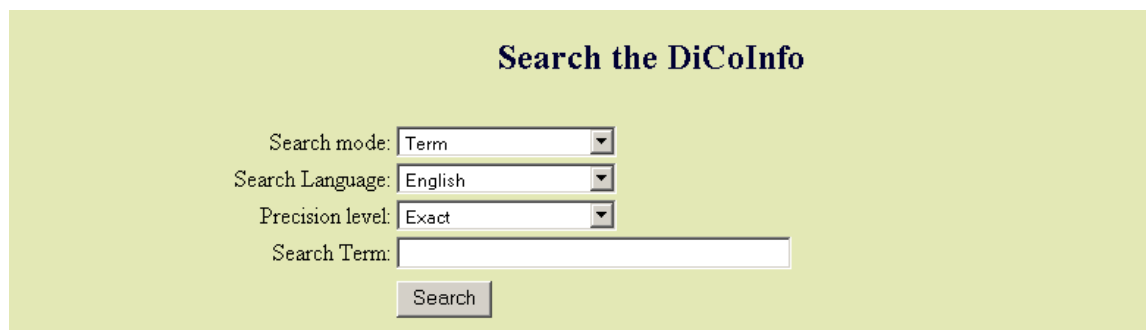
[abort](#) [access](#) [activate](#) [address](#) [administrator](#) [algorithm](#) [allocate](#) [allocation](#) [analog](#)  
[anchor](#) [architecture](#) [attach](#) [attached document](#) [attached file](#) [attachment](#)  
[attachment document](#) [attachment file](#) [attack](#) [attacker](#)

[Top of page](#)

[List of English terms](#)

Figure 1. Alphabetical Presentation of DiColInfo Terms.

A search engine is also available. This window (reproduced below in Figure 2) offers several advanced search options which we will now review.



The screenshot shows a search interface titled "Search the DiCoInfo". It features four input fields and a search button. The first three are dropdown menus: "Search mode" is set to "Term", "Search Language" is set to "English", and "Precision level" is set to "Exact". The fourth is a text input field labeled "Search Term:" which is currently empty. Below the text field is a button labeled "Search".

Figure 2. The DiCoInfo Search Parameters.

The search terms are not case-sensitive but are sensitive to accent marks (important when searching for French terms).

The search engine allows the user to search by term or by lexical relation. A term-based search matches the character sequence entered by the user with the headword section of the entries. This method also looks for synonyms and variants. For example, the user may reach the entry for BROWSER 1 by entering BROWSER. The same entry also comes up if the user enters NAVIGATOR, which appears in the entry as synonym. The lexical relation-based search, on the other hand, associates the sequence of characters entered with the lexical units found in the list of lexical relations.

The search parameter “Search language” allows the user to choose the language(s) that the requested terms may be from. Choosing “French” will limit the results to French-language terms while “English” will generate the English-language results. “Bilingual” will return terms in either language.

The parameter specifying precision level defines how closely the sequence entered must correspond to the file contents. Choosing “Exact”, will generate only the terms with the exact same spelling as the search term entered by the user. Therefore, if USER is entered, the results will be limited to that sequence only. If the user enters the sequence USE, he or she will get no results. Choosing “Starting with” will generate the terms beginning with the character sequence entered, so a search for the sequence USER will obtain USER and USER-FRIENDLY. “Containing” generates the terms that contain the sequence even as a part of the word, so the search for US will obtain entries for USER, USER-FRIENDLY, USB, as well as VIRUS.

## 7. Entries

Each entry corresponds to a specific sense. This sense must be linked to the areas of computer science and Internet use and technology, as defined by lexicosemantic criteria.

Entries are divided into ten subsections according to their content.

All files contain the subsections *Entry*, *Grammatical information*, *Actantial structure*, *Written by*, and *Last update*. The section *Definition* only appears with the 0-status entries (see section on Entry Status below). The sections *Synonym(s)*, which also contains *Variant(s)* and *Féminin* 'feminine form' (in French), and the section *Lexical relations* only appear in the entries whose keywords have such related words. Finally, the section *Additional information* is optional.

Some sections are displayed automatically as part of the default settings. These include the keyword, grammatical information, the entry's development status, the actantial structure and the definition. If a term has synonyms, a feminine form (in French) or variants, the sections Synonyms, Feminine Form and Variant(s) are also displayed by default. Some of the files contain equivalent terms in French or Spanish (see the section French-language Entries below). Finally the administrative sections *Written by*, and *Last update* are included, as in the example below.

**configure<sub>1</sub>**, vt

Status: 2

Actantial structure: configure: Agent{user 1} ~ s Patient{hardware 1, software 1}

Linguistic realizations of actants

Definition [not yet available since it is a 2-status entry]

Context(s)

Lexical relations

*French*: [configurer<sub>1</sub>](#)

*Spanish*: [\\_configurar](#)

Written by: LPD MCLH JP

Last update: 29/07/2008



**Web<sub>1</sub>**, n

Status: 2

Actantial structure: the Web: ~ used by Agent{internaut 1} to act on Patient{information 1, site 1}

Linguistic realizations of actants

Variant(s): web

Synonyms: World Wide Web, WWW

Context(s):

Lexical relations

*French*: Web<sub>1</sub>

*Spanish*: Web

Written by: LPD AB MCLH

Last update: 05/12/2008

The subsections *Linguistic realizations of actants*, *Context(s)*, *Lexical relations* and *Additional information* are displayed optionally.

## 7.1 The Headword

Each entry keyword is followed by a number indicating the sense.

EMAIL <sub>1</sub>

EMAIL <sub>2</sub>

EMAIL <sub>3</sub>

This number is always displayed, even if the term has only one specialized sense. The sense number is used in all the other sections in which the term appears. If no sense number appears with a word in some section, it means that the term in question is not the keyword of a DiCoInfo entry.

In general, senses are not arranged in a hierarchical order and the sense numbers do not have a particular meaning except to differentiate them. There are, however, some exceptions to this principle.

1. Nominalizations from verbs carry the same sense number as their corresponding verbs. Moreover, the senses indicating process and result are ordered 1 and 2 respectively, as the two nominalizations of 'to download' below.

DOWNLOAD <sub>1</sub>    DOWNLOAD <sub>1.1</sub>

                  DOWNLOAD <sub>1.2</sub>

2. Some very closely related senses are distinguished by letters (a, b), as the transitive and intransitive senses of 'to boot' and related senses of 'virtual' below.

BOOT <sub>1a</sub>	([A machine] boots)
BOOT <sub>1b</sub>	(Someone causes [a machine] to boot)
COMPILE <sub>1a</sub>	(Something, like a program, compiles)
COMPILE <sub>1b</sub>	(An Instrument, for example a compiler, compiles something, like a program)
COMPILE <sub>1c</sub>	(Someone causes something to compile using an Instrument)

Finally, each entry keyword has a label indicating the part of speech to which it belongs. If the keyword is a verb, it is marked as transitive, intransitive or pronominal. In French, if it is a noun, the part of speech is followed by the gender, unless it is a term that can be of either masculine or feminine gender.

VIEW<sub>1</sub>, vt  
 BOOT<sub>1a</sub>, vi  
 COMPUTER<sub>1</sub>, n  
 VIRTUAL<sub>1a</sub>, adj  
 REMOTELY, adv

## 7.2 Entry Status

The DiCoInfo dictionary is still under construction, and its entries are at different points of the editing process. The entry status is a number quantifying the degree of completion of the entry relative to the other entries, so the user can understand what may be missing.

- 0-status = Editing completed, meaning all obligatory subsections have been filled out and their encoding is fixed.
- Status 1 = Advanced editing of the sections Actantial structure, *Context(s)*, and a list of lexical relations is available. However, the definition is not yet provided, nor is the encoding of the lexical functions (see section 5.6).
- Status 2 = Some advanced editing (semantic differences elicited, and the Actantial structure, contexts and a short list of lexical relations are already provided, but the definition and encoding of lexical functions are not yet available.

### 7.3 Actantial Structure

This section is divided into two parts. First, a subsection called Actantial structure lists the semantic actants of the headword and describes their role with regard to that predicate term, as in the example below.

BROWSE<sub>1</sub>, vt

Actantial structure: browse: Agent{internaut} ~s Location{internet 1}  
with Instrument{browser 1}

This description shows first of all that BROWSE has three semantic actants and illustrates its grammatical role with regard to these actants. The semantic actants themselves are represented by a system of standard labels which describe their semantic role with regard to the predicate keyword. The DiCoInfo uses approximately 12 standard semantic role labels with specific and consistent meanings.

#### 7.3.1 Semantic Roles

The most frequently used labels for the semantic roles of actants found in the DiCoInfo are AGENT, PATIENT, DESTINATION, SOURCE, INSTRUMENT and LOCATION, defined below. These labels have been selected for the specific aims of this dictionary, so their definitions here may differ slightly from definitions used for these labels elsewhere in the linguistics literature.

- a. AGENT:** the actant that originates the action designated by the headword term, or the actant responsible for the existence or the use of the entity designated by the term, as in the example below.

CONFIGURE<sub>1</sub>, vt: AGENT ~s PATIENT (possible realizations of the agent: *administrator, user*)

COMPILE<sub>1c</sub>, vt: AGENT ~s PATIENT with INSTRUMENT (possible realization of the agent: *programmer*)

- b. PATIENT:** the actant which undergoes the action expressed by the term, or the element created, used, or affected by the agentive actant of the action designated by the term, as in these examples.

BOOT<sub>1a</sub>, vi: PATIENT ~s (possible realizations of the patient: *computer, machine, operating system*)

KEYBOARD<sub>1</sub>, n: ~ used by AGENT to act on PATIENT (possible realizations of the patient: *character, data, information*)

- c. **DESTINATION**: the actant designating the target place or position of an entity or process designated by the term, an action initiated by the agent or which is a typical function of an entity.

INSTALL<sub>2</sub>, vt: AGENT ~s PATIENT on DESTINATION (possible realizations of the destination: *computer, hard disk, PC*)

LOG ON<sub>1</sub>, vi: AGENT ~ to DESTINATION (possible realizations of the destination: *computer, host, Internet*)

- d. **SOURCE**: the actant referring to the place from which an activity originates, or referring to the initial element from which the typical function of an entity begins.

DOWNLOAD<sub>1</sub>, vt: AGENT ~S PATIENT from SOURCE to DESTINATION by AGENT (possible realizations of the source: *computer, Internet, network*)

DEBUGGING<sub>1</sub>, n: ~ of SOURCE by AGENT with INSTRUMENT to remove ASSAILANT (possible realization of the source: *program*)

- e. **INSTRUMENT**: the actant referring to the entity used by the agent to carry out the action expressed by the term, or to create or use the entity expressed by the term, as in the examples below.

CLICK<sub>1</sub> vi: AGENT ~s on PATIENT with INSTRUMENT (possible realization of the instrument: *mouse*)

TYPE<sub>1</sub>, vt: Agent ~s PATIENT with INSTRUMENT (possible realization of the instrument: *keyboard*)

- f. **LOCATION**: the actant referring to the place where the action denoted by the term takes place, or to the place where the typical function of an entity unfolds.

BROWSER<sub>1</sub>, n: ~ used by AGENT to go in LOCATION (possible realizations of the location: *Internet, Web*)

Other labels are also used to identify the semantic roles of actants, but they are much less frequent than the ones listed above.

### 7.3.2 Typical Actants

The actants, already labeled according to Actantial roles (see section 7.3.1), are also accompanied by examples of the most commonly used actants. These examples immediately follow the semantic role label, as in the following example.

CONFIGURE 1, vt

Actantial structure: configure: Agent{user 1} ~ s Patient{hardware 1, software 1}

The typical actant corresponds to one of the linguistic expressions of the actant. These examples are supposed to facilitate the user's comprehension of the term described in the entry. The user should therefore be able to read the actant structure in one of two ways:

EDIT 1, vt

Actantial roles: AGENT ~s PATIENT with INSTRUMENT

Typical actants: user 1 ~s file 1 with editor 1

The typical actant is chosen by applying a combination of the following criteria:

1. The typical actant is the one most naturally used in the definition of the term. So while *user* and *programmer* are both words that can correctly refer to the agent of the verb *edit*, the term *user* is felt to be the most appropriate term to describe someone carrying out this activity.
2. The typical actant is the most common word, from a descriptive point of view, of several used in the context of the term being described. For example, *site* is a typical destination of the verb *visit*. Other words such as *blog* and *page* could be used to indicate the same thing, but speakers use them much less often than *site*.
3. The typical actant is often the generic term covering a wider semantic range than the other more specific quasi-synonyms. This last criterion explains the choice of actants such as *data*, *file*, *program* and *hardware* as typical.

### 7.3.3 Linguistic Realizations of Actants

On demand, the user can access a list of linguistic realizations of the actants which were observed in the resources consulted during the composition of the entries. These lists recall the frames illustrating the semantic roles in the Actantial structure section of the entry, as in the following example entry for *abort*.

**abort**<sub>1</sub>, vt

Status: 2

Actantial structure: abort: AGENT{user 1} ~ s PATIENT{task 1}

[Linguistic realizations of actants](#)

<b>Agent</b>
user 1
<b>Patient</b>
job, process, task 1, transfer 1

If one or more of the linguistic realizations of actants are themselves keywords of DiCoInfo entries, the user may click on them as links to their respective entries.

## 7.4 The Definition

Definitions appear in the 0-status entries. In English, no entry has a definition. However, a number of French entries include this data category (refer to French documentation for this data category).

## 7.5 Contexts

The term's contexts, displayed on demand, illustrate how the term is used concretely in current specialized texts. The contexts below are those used for the term INFECTION 1.

**infection**<sub>1</sub>, n

Status: 2

Actantial structure: infection: ~ of DESTINATION{file 1; hardware 1} by ASSAILLANT{virus 1}

[Linguistic realizations of actants](#)[Context\(s\)](#):

*With an antivirus program installed, the likelihood of infection is slim. (Source:CALL THE PC DOCTOR) / The hoax asks you to look out for defects in image quality, which are said to be symptomatic of infection. (Source:VIRUSINFO02) / When you listen to the news, you hear about many different forms of electronic infection. (Source:HOW COMPUTER VIRUSES WORK)*

The contexts are brief extracts from the texts that constitute the corpus for the DiCoInfo. If a minor correction is made to the quoted source, the source name is preceded by *According to*.

### 7.5.1 Annotated Contexts

Some files contain a section called Annotated Contexts, for example the entry for FORWARD below.

**forward<sub>1</sub>**, vt

Status: 2

Actantial structure: forward: AGENT{user 1} ~ s PATIENT{message 1} to DESTINATION {recipient 1}

[Linguistic realizations of actants](#)

[Context\(s\)](#):

[Annotated contexts](#)

*The email goes on to ask you to forward this email to as many people as possible. (Source:INTERNET) / Similarly, the application gateway accepts all incoming e-mail, reformats it to use the proprietary system, and forwards each message to the appropriate recipient using the proprietary e-mail system. (Source:COMER) / One in six (about 15 percent) told Pew that private communications—IM logs, e-mails, or text messages—had been posted publicly by someone else or forwarded around. (Source:CYBERBULLYING)*

This section provides additional information about the linguistic behaviour of these terms<sup>1</sup>, including:

1. Access to a larger number of contexts (up to 20 different contexts).
2. The highlighting of contexts and actants that are used with the term, with their semantic roles (as defined in subsection 7.3.2) clearly labelled.
3. The highlighting of the surrounding contexts, i.e. the non-obligatory participants often associated with the term. These participants are also labelled with their semantic role.
4. A table summarizing the semantic roles in the various contexts, their frequencies, as well as the syntactic groups and functions associated with each role. The table also includes the realization of the actants as they are most commonly observed in the corpus.

Below we have reproduced two examples of annotated contexts, taken from the file for the term FORWARD.

The email goes on to ask *you* to FORWARD *this email to as many people as possible*. [INTERNET 0 JP 07/05/2009]

Similarly, *the application gateway* accepts all incoming e-mail, reformats it to use the proprietary system, and FORWARDS *each message to the appropriate recipient using the proprietary e-mail system*. [COMER 0 JP 07/05/2009]

In the annotated contexts, the elements are graphically marked in the following ways:

1. The keyword term to which the annotation belongs is capitalized.

---

<sup>1</sup> The methodology used for the annotation of these contexts is largely drawn from the work of the team responsible for the FrameNet project (Fillmore 1977, 1982; Fillmore and Atkins 1992; Ruppenhofer et al. 2006).

2. The surrounding actants are coded by color. Each color corresponds to a different semantic role.
3. Actants themselves appear in bold.
4. The realizations of the actants appear in italics.
5. The information following the annotated contexts is a sequence containing the source, the annotation status (only 0-status contexts are displayed), the editor's code, and the date last updated.

The summary table follows the annotated contexts. Part of one such table, for the term FORWARD, is reproduced below.

FORWARD 1		
<b>Actants</b>		
<b>Agent</b>	Indirect link (NP) (4) Subject (NP)	you (2) gateway mail transport agent service
<b>Patient</b>	Object (NP) (6) Object (PRO) (2)	message (2) mail (2) email it
<b>Destination</b>	Complement (PP-to) (5)	people (2) recipient anyone friend
<b>Others</b>		
<b>Mode</b>	Complement (Clause)	using the proprietary e-mail system
<b>Means</b>	Indirect link (NP)	protocol

The summary table is divided into two parts. The first is dedicated to the Actantial information, while the second presents information about the other syntactic groups observed in the environment around the term which is the keyword of the entry.

The first column of the table recalls the semantic roles from the annotated contexts. The second column presents the syntactic functions associated with each of these semantic roles as well as the syntactic group related to a specific function. If this semantic group is a prepositional phrase, the preposition is also indicated. The frequency of each semantic group fulfilling a given function is also provided. The final column summarizes the example linguistic expressions taken from the annotated contexts.



## 7.6 Related Terms

Related terms (except synonyms) are described in a list containing the related term followed by an explanation of its relationship with the keyword.

### 7.6.1 Synonyms and Variants

If a given term has synonyms or spelling variants, these come after the linguistic realizations of the actants (or after the definition, if the term's entry is already updated to 0-status). Two examples are provided below.

#### **browser**<sub>1, n</sub>

Status: 2

Actantial structure: a browser: ~ used by AGENT{internaut} to go in LOCATION{Internet 1}

[Linguistic realizations of actants](#)

Synonym(s): browsing software, Internet browser, navigator, web browser, Web navigator, WWW browser

#### **program**<sub>1, n</sub>

Status: 2

Actantial structure: program: ~ created by AGENT{programmer 1} in MATERIAL{language} to act on PATIENT{task}

[Linguistic realizations of actants](#)

Variant(s): programme

Synonym(s): computer program

### 7.6.2 List of Paradigmatic and Syntagmatic Lexical Relations

The final and most important subsection is dedicated to the description of lexical relations between the keyword and other terms, many of which also appear in the DiCoInfo. This description, displayed by default, appears in two columns.

The example below shows some of the terms related to the keyword term PARTITION<sub>1</sub>.

**partition**<sub>1, n</sub>

Status: 2

Actantial structure: partition: ~ on SUPPORT{hard disk 1} created by AGENT{user 1} to act upon PATIENT1{data 1} or PATIENT2{software 1}

[Linguistic realizations of actants](#)

[Context\(s\)](#)

[Lexical relations](#)

[Actantial roles](#)

<b>Explanation - Typical term</b>	<b>Related term</b>
Which is divided into logical partitions	extended ~
On a p.	on a ~
The user creates a p.	<a href="#">create</a> <sub>1a</sub> ~
The user removes a p.	<a href="#">delete</a> <sub>1a</sub> ~
The user prepares a p.	<a href="#">format</a> <sub>1a</sub> ~
Size	~ size

The right column lists the terms that have a semantic link with the keyword. They are accompanied by a sense number which allows the user to consult the entries for these terms in the dictionary. If no sense number is present, then the word does not have a DiCoInfo article. Related terms that contain a clickable link to a file may continue to be developed elsewhere in the dictionary.

The semantic links described are paradigmatic or syntagmatic. Most of the terms bearing a paradigmatic relation with the keyword appear only in the right column. Most of the terms related syntagmatically with the keyword are accompanied by an example utterance showing how they can be used with the keyword term.

To describe the semantic relation between the keyword term and the terms in the right-side column, we have largely adopted the lexical functions framework of Explanatory and Combinatorial Lexicology. However, to facilitate understanding of these semantic relations among a wider readership, the explanation in the left column is simplified, popular versions of scientific definitions, modeled partially after the everyday-language versions of the definitions found in the *DiCo* (Polguère 2003) and in the *Lexique actif du français* (Mel'čuk et Polguère 2007).

Finally, the related terms are arranged in the following order:

1. Quasi-synonyms, related meanings and generics (i.e. *laptop: computer, notebook, palmtop*)
2. Antonyms and contrastives (i.e. *local: remote*)
3. Links or relations between the keyword and often regular morphological derivations (i.e. *attach: attached, attachment*)

4. The specific or modified “kinds of” a generic keyword (i.e. *file: empty ~, HTML ~, binary ~, XML ~, attachment*)
5. Verbal collocates (and nominal or adjectival derivations from these verbs) (i.e. *file: create a ~, delete a ~, save a ~*)
6. Other, less regular paradigmatic (i.e. *computer: memory, processor, manufacturer*)

### 7.6.3 Lexical Relations and Actantial Structure

Many explanations of lexical semantic relations between a keyword and another word refer to the Actantial structure, as in the following examples.

**Internet**<sub>1, n</sub>

Status: 2

Actantial structure: the Internet: ~ used by Agent{user 1} to act on Patient{information, site}

The user starts using the I.	connect to the ~
The user uses the I.	browse the ~

### 7.6.4 Triple Encoding of the Lexical Relations

A user can access three different levels of explanation of the lexical relations described, if he or she so wishes.

The first-level explanations appear by default when the user displays the lexical relations. The explanations here refer to the notation of the “typical” actants.

**site**<sub>1</sub>, n

Status: 2

Actantial structure: site: ~ created by AGENT{author} to offer PATIENT{information} to RECEIVER{space}

[Linguistic realizations of actants](#)

Synonym(s): internet site, web site, website

[Context\(s\)](#)

[Lexical relations](#)

[Actantial roles](#)

<b>Explanation - Typical term</b>	<b>Related term</b>
That is used to share ideas regularly	<a href="#">blog</a> <sub>1</sub>
That is created by the author for commercial purposes	commercial ~
That uses a specific protocol	FTP ~
That is identical to a s. but placed on a different computer	mirror ~
On a s.	on a ~
The author creates a s.	<a href="#">create</a> <sub>1</sub> a ~
The author creates a s.	construct a ~
The author updates a s.	<a href="#">update</a> <sub>1</sub> a ~
The author makes sure a s. can still be used by the user	maintain a ~
The s. has information	the ~ contains...

The second-level explanation is offered when the user clicks on Actantial roles. The explanations here refer to the standard notation of Actantial roles, as in the following example.

**site**<sub>1, n</sub>

Status: 2

Actantial structure: site: ~ created by AGENT{author} to offer PATIENT{information} to RECEIVER{space}

[Linguistic realizations of actants](#)

Synonym(s): internet site, web site, website

[Context\(s\)](#)

[Lexical relations](#)

[Actantial roles](#)

<b>Explanation – Actantial role</b>	<b>Related term</b>
That is used to share ideas regularly	<a href="#">blog</a> <sub>1</sub>
That is created by the agent for commercial purposes	commercial ~
That uses a specific protocol	FTP ~
That is identical to a « Key word » but placed on a different computer	mirror ~
On a « Key word »	on a ~
The agent creates a « Key word »	<a href="#">create</a> <sub>1 a</sub> ~
The agent creates a « Key word »	construct a ~
The agent updates a « Key word »	<a href="#">update</a> <sub>1 a</sub> ~
The agent makes sure a « Key word » can still be used by the receiver	maintain a ~
The « Key word » has patient	the ~ contains...

The third and final level of explanation is offered when the user clicks on Lexical functions. The explanations here closely follow the notation proposed for Explanatory and Combinatorial Lexicology (Mel'čuk et al. 1984-1999, 1995). This level of explanation is available only for the 0-status articles.

## 7.7 Additional Information

Certain entries contain a subsection whose main function is to direct the user to Web sites containing useful information about the term being described.

## 7.8 Administrative Section

The final subsection record administrative information: the identity of the editors and the date when the entry was last updated.

**programming**<sub>1, n</sub>

Status: 2

Written by: LPD MEP MCLH  
Last update: 09/04/2009

## **8. French language terms**

The DiColInfo contains many more entries for French terms than for English terms. In addition, many French-language entries have 0-status or 1-status. At the time of writing, all the English-language entries were updated to status 2, and are therefore less fully elaborated than the French entries.

A section called “Equivalent” is available for many files. If the equivalent in another language of the term is available in the database, this word will be a clickable link taking the user to that entry.

## **9. Current State of the DiColInfo**

At the time of writing, the DiColInfo contained almost 1900 entries for French and approximately 1000 entries for English (remembering that each entry corresponds to a sense, as described in section 7.1). Many French entries (more than 1000) are complete and available at the Web site <http://olst.ling.umontreal.ca>. In English, approximately 500 entries are online.

Projects aimed at extending the DiColInfo to descriptions of terms from other languages (namely, Spanish) are currently being developed.

## **10. Obtaining the DiColInfo for Research Purposes**

It is possible to obtain an original XML version of the DiColInfo for research purposes by making a request to Marie-Claude L’Homme ([mc.lhomme@umontreal.ca](mailto:mc.lhomme@umontreal.ca)).

## 11. To learn more...

### On terminology:

L'Homme, M.C. (2004a). *La terminologie : principes et techniques*, Montréal: Presses de l'Université de Montréal.

### On Explanatory and Combinatorial Lexicography:

Jousse, A.L. and A. Polguère (2005). *Le DiCo et sa version DiCouèbe. Document descriptif et manuel d'utilisation*. Version du rapport 1.0 – 19 avril 2005, Montréal: Observatoire de linguistique Sens-Texte (OLST).

Mel'čuk, I., A. Clas, A. and A. Polguère (1995). *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve (Belgique): Duculot / Aupelf - UREF.

Polguère, A. (2003). "Collocations et fonctions lexicales : pour un modèle d'apprentissage", In F. Grossmann and A. Tutin (éd.). *Les collocations. Analyse et traitement*, Coll. Travaux et recherches en linguistique appliquée, Paris: Éditions de Werelt, pp. 117-142.

### On FrameNet and Frame Semantics:

Fillmore, C.J. (1968). "The case for case", In Bach, E. and R.T. Harms (eds.). *Universals in Linguistic Theory*, New York: Holt, Rinehard and Winston, pp. 1-88.

Fillmore, C.J. (1977): "Scenes-and-frames semantics, Linguistic Structures Processing", In Zampolli, A. (ed.). *Fundamental Studies in Computer Science*, No. 59, North Holland Publishing, pp. 55-88.

Fillmore, C.J. (1982). "Frame Semantics", In The Linguistic Society of Korea (ed.). *Linguistics in the Morning Calm*. Seoul: Hanshin, pp. 111-137.

Fillmore, C.J. and B.T.S. Atkins. (1992). "Towards a Frame-based Organization of the Lexicon: The Semantics of RISK and its Neighbors", In Lerher, A. and E. Kittay (eds.). *Frames, Fields and Contrast. New Essays in Semantics and Lexical Organization*. Hillsdale: Lawrence Erlbaum, 75-102.

Fillmore, C.J., C. R. Johnson and M.R.L. Petruck (2003). "Background to FrameNet", In Fontenelle, T. (ed.). *FrameNet and Frame Semantics. Special Issue of the International Journal of Lexicography* 16(3), pp. 235-250.

Ruppenhofer, J., M. Ellsworth, R.L.M. Petruck, C. Johnson and J. Scheffczyk. 2006. *FrameNet II: Extended Theory and Practice* ([http://framenet.icsi.berkeley.edu/index.php?option=com\\_wrapper&Itemid=126](http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper&Itemid=126)) (Consulted 12 September 2006).

### On the usefulness of Explanatory and Combinatorial Lexicology in describing terms:

Frawley, W. (1988). "New forms of Specialized Dictionaries", *International Journal of Lexicography* 1(3), pp. 189-213.

L'Homme, M.C. (2004c). "A Lexico-semantic Approach to the Structuring of Terminology", In *Computerm 2004*, dans le cadre de Coling 2004, Université de Genève, Genève (Suisse), 29 August 2004, pp. 7-14.

L'Homme, M.C. (2007). "Using Explanatory and Combinatorial Lexicology to Describe Terms", In Wanner, L. (ed.). *Selected Lexical and Grammatical Topics in the Meaning-Text Theory*. In Honour of Igor Mel'cuk, Amsterdam/Philadelphia: John Benjamins.

Other dictionaries and databases compiled using similar methods as the DiCoInfo:

Binon, J., S. Verlinde, J. Van Dyck and A. Bertels (2000). *Dictionnaire d'apprentissage du français des affaires. Dictionnaire de compréhension et de production de la langue des affaires*, Paris: Didier.

Descamps, J.L. (1976). *Dictionnaire contextuel de français pour la géologie : essai de classement de concordances de français scientifique et étude critique*, Paris : Didier.

Dicouèbe. *Dictionnaire en ligne de combinatoire du français* (<http://olst.ling.umontreal.ca/dicouebe/>) (Accessed 31 July 2007).

FrameNet (<http://framenet.icsi.berkeley.edu/>) (consulté le 11 décembre 2006).

*The Kicktionary. A Multilingual Electronic Dictionary of Football (Soccer) Language* (<http://www.kicktionary.de/>) (Accessed 30 September 2006).

Mel'čuk, I. et al. (1984-1999). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques 1-IV*, Montréal: Les Presses de l'Université de Montréal.

Mel'čuk, I. et A. Polguère (2007). *Lexique actif du français*, Bruxelles : Duculot.

Picoche, J. et J.-C. Rolland (2001). *Dictionnaire du français usuel*, Bruxelles: De Boeck / Duculot.

*Unified Verb Index*. University of Colorado (<http://verbs.colorado.edu/verb-index/index.php>) (Accessed 31 July 2007)

Verlinde, S. A. Bertels, J. Binon, N. Nouwen, S. Ostyn, G. Petit, J. Van Dyck, F. Schaeffler, D. Tribout and L. Perrier (éd.). *DAFLES. Dictionnaire d'Apprentissage du Français Langue Étrangère ou Seconde* (<http://www.kuleuven.ac.be/dafles/>) (Accessed 20 September 2005).

On the lexicographic methodology used to compile the DiCoInfo:

Jousse, A.L. and M. Bouveret (2003). "Lexical Functions to Represent Derivational Relations in Specialized Dictionaries", *Terminology* 9(1), pp. 71-98.

L'Homme, M.C. (1998). "Définition du statut du verbe en langue de spécialité et sa description lexicographique", *Cahiers de lexicologie* 73(2), pp. 61-84.

L'Homme, M.C. (2002). "Fonctions lexicales pour représenter les relations sémantiques entre termes", *Traitement automatique des langues (TAL)* 43(1), pp. 19-41.

L'Homme, M.C. (2003). "Capturing the Lexical Structure in Special Subject Fields with Verbs and Verbal Derivatives: A model for specialized lexicography", *International Journal of Lexicography* 16(4), pp. 403-422.

L'Homme, M.C. (2004b). "Sélection des termes dans un corpus d'informatique : comparaison de corpus et critères lexico-sémantiques", In *Euralex International Congress. Proceedings*. Lorient (France), pp. 583-593.

L'Homme, M.C. (2005). "Conception d'un dictionnaire fondamental de l'informatique et de l'Internet : sélection des entrées", *Le langage et l'homme* 40(1), pp. 137-154.

L'Homme, M.C. (2005). "Sur la notion de terme", *Meta* 50(4), pp. 73-107.



L'Homme, M.C. (2008). "Le DiColInfo. Méthodologie pour une nouvelle génération de dictionnaires spécialisés", *Traduire* 217, pp. 78-103.

#### On the other-language versions of the DiColInfo:

Bae, H.S. and M.C. L'Homme (2008). "Converting a Monolingual Lexical Database into a Multilingual Specialized Dictionary", In Boers, F., J. Darquennes, K. Kerremans and R. Temmerman (eds.). *Multilingualism and Applied Comparative Linguistics*, Volume 2, Cambridge: Cambridge Scholars Publishing, pp. 225-255.

L'Homme, M.C. and H.S. Bae (2006). "A Methodology for Developing Multilingual Resources for Terminology", *LREC 2006. Language Resources and Evaluation. Proceedings*, Genoa (Italy).

#### On the automatic methods employed to compile to DiColInfo:

Alain, B (2006). *Le DiColInfo. De l'idée à l'impression*. Rapport de stage. Département d'informatique et de recherche opérationnelle, Université de Montréal

Bolle, P. (2006) *Projet actants et dictionnaires d'apprentissage : proposition d'un modèle de représentation adapté au grand public*. Rapport de stage, Département de linguistique et de traduction, Université de Montréal.

Claveau, V. and M.C. L'Homme (2004). "Discovering Specific Semantic Relationships between Nouns and Verbs in a Specialized French Corpus", In *Computerm 2004. Proceedings*, Coling 2004, Genève (Switzerland), pp. 39-46.

Claveau, V. and M.C. L'Homme (2005a). "Apprentissage par analogie pour la structuration de terminologie – Utilisation comparée de ressources endogènes et de ressources exogènes", In *Actes. Terminologie et intelligence artificielle, TIA 2005*, Université de Rouen, Rouen (France).

Claveau, V. and M.C. L'Homme (2005b). "Structuring Terminology by Analogy Machine Learning", In *Terminology and Knowledge Engineering. TKE 2005*, Copenhagen Business School, Copenhagen (Denmark).

Claveau V. and M.C. L'Homme (2006). "Discovering and Organizing Noun-Verb Collocations in Specialized Corpora Using Inductive Logic Programming", *International Journal of Corpus Linguistics* 11(2), pp. 209-243.

Drouin, P. (2003). "Term Extraction Using Non-technical Corpora as a Point of Leverage", *Terminology* 9(1), pp. 99-115.

Lemay, C., M.C. L'Homme and P. Drouin (2005). "Two Methods for Extracting "Specific" Single-word Terms from Specialized Corpora: Experimentation and Evaluation", *International Journal of Corpus Linguistics* 10(2), pp. 227-255.

#### On the contextual annotations:

Bae, H.S., M.C. L'Homme and G. Lapalme (2008). "Semantic Roles in Multilingual Terminological Descriptions: Application to French and Korean Contexts", *Multilingual and Comparative Perspectives in Specialized Language Resources. Proceedings of the Workshop. Language Resources and Evaluation, LREC 2008*, Marrakech, Morocco.

Hadouche, F., M.C. L'Homme, G. Lapalme and A. Le Serrec (2009). "Intégration d'informations syntaxico-sémantiques dans les bases de données terminologiques : méthodologie d'annotation et perspectives d'automatisation", In *International Workshop in Terminology and Lexical Semantics (TLS'09)*, Université de Montréal, Montréal.

## Other dictionaries consulted:

Collin, S.M.H., F. Laurendeau and B. Mouget (1996). *Le bilingue de l'informatique : dictionnaire français-anglais, anglais-français*. Coll. "Peter Collins", Middlesex: Peter Collin.

*Le dictionnaire des développeurs* (<http://dico.developpez.com/html/>) (Accessed 31 July 2007).

Ginguay, M. (1998). *Dictionnaire français-anglais d'informatique: bureautique, télématique, micro-informatique*, 6<sup>th</sup> Ed., Paris: InterEditions.

*Le Grand dictionnaire terminologique* (<http://www.granddictionnaire.com>) (Accessed 31 July 2007)

Meynard, I. (2000). *Internet. Répertoire bilingue de combinaisons lexicales spécialisées français-anglais*, Montréal: Linguattech.

Office de la langue française (2001). *Dictionnaire d'Internet, de l'informatique et des télécommunications. Technologies de l'information anglais-français*, Québec: Gouvernement du Québec.

*Oxford Dictionary of Computing* (2004). Fifth Edition. Oxford: Oxford University Press.

*Termium Plus* (<http://www.termiumplus.com>) (Accessed 15 December 2007).