United States District Court,
N.D. California.

INFORMATICA CORPORATION, a Delaware Corporation,
Plaintiffs.
v.
BUSINESS OBJECTS DATA INTEGRATION, INC., formerly known as Acta Technology, Inc., a Delaware Corporation,
Defendants.
and Related Counterclaim,
and Related Counterclaims.

No. C 02-3378 JSW

Aug. 2, 2005.

<div align="center">

CLAIMS CONSTRUCTION ORDER

</div>

JEFFREY S. WHITE**, District Judge.**

The Court has been presented with a technology tutorial and briefing leading up to a hearing pursuant to Markman v. Westview Instruments, Inc., 517 U.S. 370 (1996). This Order construes the ten claim terms selected by the parties, which appear in the four patents at issue in this case, United States Patent Nos. 6,014,670 ("the '670 patent") and 6,339,775 B1 ("the ' 775 patent"), both entitled "Apparatus and Method for Performing Data Transformations in Data Warehousing," United States Patent No. 6,208,990 B 1 ("the '990 patent") entitled "Method and Architecture For Automated Optimization of ETL Throughput in Data Warehousing Applications," and United States Patent No. 6,044,374 ("the '374 patent") entitled "Method and Apparatus for Sharing Metadata Between Multiple Data Marts Through Object References."

<div align="center">

BACKGROUND

</div>

Plaintiff Informatica Corporation and defendant Business Objects Data Integration, Inc. ("BODI"), formerly known as Acta Technology, Inc., are competitors in the field of business analytics software, which enables enterprises to automate integration, analysis and delivery of key data. Informatica designs and markets software that enables its customers to transform and update data within various data warehouses and data marts, and to share data among enterprise constituencies. BODI designs and markets a real-time data integration platform that intelligently manages and optimizes the performance of an enterprise's data infrastructure.

On July 15, 2002, Informatica filed this suit alleging infringement of three of the patents-in-suit by defendant BODI. Informatica subsequently amended the complaint and added a claim for infringement of the fourth patent-in-suit. BODI answered the first amended complaint and denied infringement of the patents-in-suit. BODI additionally asserted counterclaims of patent invalidity, unenforceability, and non-

infringement. Pursuant to a stipulation reached between the parties, BODI later filed an amended answer and counterclaims, which added counterclaims for unenforceability of the patents-in-suit due to inequitable conduct before the Patent and Trademark Office during the patent prosecution.

There are four patents at issue in this case. The first two, the '670 patent and the '775 FN1 patent, claim an apparatus and method for transforming data in data warehousing applications. The patents additionally disclose a functional specification for a transformation description language, which describes how data is to be manipulated in a data warehousing application. The parties agree that a data warehousing application is software that extracts, transforms and loads data into one or more databases, where the data is useful for decision support, usually across an enterprise. The '990 patent pertains to an apparatus and method for automating optimal throughput of the data extraction/transformation/loading ("ETL") process in data warehousing applications. Finally, the '374 patent relates to a method and apparatus for sharing metadata between multiple data marts through the use of object references. Metadata is data about data. Again, the parties agree that a data mart application is software that extracts, transforms and loads data into one or more databases, where the database(s) contain a subset of corporate data useful for decision support of an aspect of a business. Informatica is the sole assignee of all four patents-in-suit.

FN1. The '775 is a continuation-in-part of and claims the benefit of application Ser. No. 08/966,449, which designated the '670 patent.

ANALYSIS

A. Legal Standard

"It is a bedrock principle of patent law that the claims of a patent define the invention to which the patentee is entitled the right to exclude." Innova/Pure Water, Inc. v. Safari Water Filtration Sys., Inc., 381 F.3d 1111, 1115 (Fed.Cir.2004). The interpretation of the scope and meaning of disputed terms in patent claims is a question of law and exclusively within the province of the court to decide. Markman v. Westview Instruments, Inc., 517 U.S. 370, 372 (1996). The inquiry into the meaning of the claim terms is "an objective one." Innova/Pure Water, Inc., 381 F.3d at 1116. As a result, a court undertaking the construction of disputed terms "looks to sources available to the public that show what a person of skill in the art would have understood the disputed claim language to mean." *Id*. In most cases, the court's analysis will focus on three sources: the claims, the specification, and the prosecution history. Markman v. Westview Instruments, Inc ., 52 F.3d 967, 979 (Fed.Cir.1995) (en banc), *aff'd*, 517 U.S. 370, 372 (1996). However, on occasion reliance on extrinsic evidence regarding the relevant scientific principles, the meaning of technical terms, and the state of the art at the time at the time the patent issued is appropriate.

The starting point of the claim construction analysis is an examination of the specific claim language. "[T]he analytical focus must begin and remain centered on the language of the claims themselves, for it is that language that the patentee chose to use to particularly point out and distinctly claim the subject matter which the patentee regards as his invention." Innova/Pure Water, Inc., 381 F.3d at 1116 (internal quotations and citations omitted). Indeed, in the absence of an express intent to impart a novel meaning to a term, an inventor's chosen language is given its ordinary meaning. York Prods., Inc. v. Cent. Tractor Farm & Family Center, 99 F.3d 1568, 1572 (Fed.Cir.1996); *see also* Invitrogen Corp. v. Biocrest Mfg., L.P., 327 F.3d 1364, 1367 (Fed.Cir.2003) ("Claim language generally carries the ordinary meaning of the words in their normal usage in the field of invention."). The terms used within a claim bear "a 'heavy presumption' that they mean what they say and have the ordinary meaning that would be attributed to those words by persons skilled in

the relevant art." Texas Digital Sys., Inc. v. Telegenix, Inc., 308 F.3d 1193, 1202 (Fed.Cir.2002); *see also* Renishaw PLC v. Marposs Societa' per Azioni, 158 F.3d 1243, 1248 (Fed Cir.1998) (recognizing that "the claims define the scope of the right to exclude; the claim construction inquiry, therefore, begins and ends in all cases with the actual words of the claim"). The court's final construction, therefore, must accord with the words chosen by the patentee to mete out the boundaries of claimed invention.

The court should also look to intrinsic evidence, including the written description, the drawings, and the prosecution history, if included in the record, to provide context and clarification regarding the intended meaning of the claim terms. Teleflex, Inc. v. Ficosa N. Am. Corp, 299 F.3d 1313, 1324-25 (Fed.Cir.2002). The claims do not stand alone, rather, "they are part of 'a fully integrated written instrument.' " Phillips v. AWH Corp., Nos. 03-1269, 03-1286, --- F.3d. ----, 2005 WL 1620331, (Fed.Cir. July 12, 2005) (en banc) (quoting *Markman*, 52 F.2d at 978). The specification "may act as a sort of dictionary, which explains the invention and may define the terms used in the claims." Markman, 52 F.3d at 979-80. The specification can also indicate whether the patentee intended to limit the scope of a claim, despite the use of seemingly broad claim language. SciMed Life Sys., Inc. v. Advanced Cardiovascular Sys., Inc., 242 F.3d 1337, 1341 (Fed.Cir.2001) (recognizing that when the specification "makes clear that the invention does not include a particular feature, that feature is deemed to be outside the reach of the claims of the patent, even though the language of the claims, read without reference to the specification, might be considered broad enough to encompass the feature in question").

Intent to limit the claims can be demonstrated in a number of ways. For example, if the patentee "acted as his own lexicographer," and clearly and precisely set forth a definition of the disputed term in either the specification or the prosecution history, the court will defer to that definition. CCS Fitness, Inc. v. Brunswick Corp., 288 F.3d 1359, 1366 (Fed.Cir.2002). All that is required to so limit the claims is for the patentee to set out the alternative meaning in the specification "in a manner sufficient to give one of ordinary skill in the art notice of the change from the ordinary meaning." Innova/Pure Water, Inc., 381 F.3d at 1117. In addition, the court will adopt an alternative meaning of a term "if the intrinsic evidence shows that the patentee distinguished that term from prior art on the basis of a particular embodiment, expressly disclaimed subject matter, or described a particular embodiment as important to the invention." CCS Fitness, Inc., 288 F.3d at 1367. Likewise, the specification may be used to resolve ambiguity "where the ordinary and accustomed meaning of the words used in the claims lack sufficient clarity to permit the scope of the claim to be ascertained from the words alone," Teleflex, 299 F.3d at 1325.

However, limitations from the specification (such as from the preferred embodiment) may not be read into the claims, absent the inventor's express intention to the contrary. Id. at 1326; *see also* CCS Fitness, 288 F.3d at 1366 ("[A] patentee need not describe in the specification every conceivable and possible future embodiment of his invention."); Virginia Panel Corp. v. MAC Panel Co., 133 F.3d 860, 866 (Fed.Cir.1997) ("[I]t is well-settled that device claims are not limited to devices which operate precisely as the embodiments described in detail in the patent."). To protect against this result, the court's focus should remain on understanding how a person of ordinary skill in the art would understand the claim terms. Phillips, 2005 WL 1620331, at *15.

If the analysis of the intrinsic evidence fails to resolve any ambiguity in the claim language, the court may then turn to extrinsic evidence, such as expert declarations and testimony from the inventors. Intel Corp. v. VIA Techs., Inc., 319 F.3d 1357, 1367 (Fed.Cir.2003) ("When an analysis of intrinsic evidence resolves any ambiguity in a disputed claim term, it is improper to rely on extrinsic evidence to contradict the meaning so ascertained."). When considering extrinsic evidence, the court should take care not to use it to vary or

contradict the claim terms. Rather, extrinsic evidence is more appropriately relied upon to assist in determining the meaning or scope of technical terms in the claims. Vitrionics Corp. v. Conceptronic, Inc., 90 F.3d 1576, 1583 (Fed.Cir.1996),

Dictionaries also may play a role in the determination of the ordinary and customary meaning of a claim term. The Federal Circuit recently reiterated that "[d]ictionaries or comparable sources are often useful to assist in understanding commonly understood meanings of words...." Phillips, 2005 WL 1620331, at *15. But the *Phillips* court admonished that district courts should be careful not to allow dictionary definitions to supplant the inventor's understanding of the claimed subject matter. "The main problem with elevating the dictionary to ... prominence is that it focuses the inquiry on the abstract meaning of the words rather than on the meaning of the claim terms within in the context of the patent." *Id*. at * 14. Accordingly, dictionaries necessarily must play a role subordinate to the intrinsic evidence.

B. Claim Construction

The '670 and '775 Patents

The '775 patent is a continuation-in-part of and claims the benefit of application Ser. No. 08/966,449, which designated the '670 patent. The specifications for the '670 and '775 patents are essentially identical.

1. "port"

The term "port" appears in numerous claims of the '670 patent (claims 1-7, 9-11, 18-19, 21-23, 30-34, 36-37, 40-42, 46, 51). Informatica asserts that the plain meaning of the term is understood by those in art and proposes the term be construed to mean "An abstraction for describing the inputs and/or outputs of sources, targets or transformation objects." BODI offers the construction:

A port is analogous to a column of a table and provides the primary means of parameterized dataflow between various objects in a mapping. A port must have a name, data type it holds, and its data flow type (*i.e.*, in, out, or in/out). A port must always be defined within the definition of a source, target, or transformation object; thus it would be meaningless to have a stand-alone definition of a port. Ports provide the means for transferring data between sources, targets, and transformation objects.

A port has the following specification:

<port_def>::=<port_name><data_type_def><port_type>

<port_name>::=<string>

<port_type>::=INOUTINOUT

<data_type_def>::=<data_type>[(<precision>[,<scale>])]

<precision>::=<integer>

<scale>::=<integer>

BODI asserts that the term port has no generally accepted meaning within the industry; accordingly, it refers

the Court to the patent specification and pulls its proposed construction from this source. Informatica responds that BQDI's proposed construction impermissibly imports a limitation from the specification into the claims. Moreover, Informatica asserts that the patent as a whole claims an apparatus and method for performing data transformations in data warehousing; the functional specification for a transformation description language ("TDL")-the basis for BODI's construction is-merely a preferred embodiment disclosed in the specification.

Claim construction requires courts to walk the fine line between interpreting the claims in light of the specification, Markman, 52 F.3d at 979, and impermissibly reading a limitation into the claim from the specification, Comark Communications, Inc. v. Harris Corp., 156 F.3d 1182, 1186-87 (Fed.Cir.1998). The balance may be found by determining how the specification characterizes the claimed invention. In other words, does the specification discuss a limitation only as part of some of the possible embodiments or does the specification read as a whole imply that "the very character of the invention requires the limitation be a part of every embodiment." Alloc, Inc. v. Int'l Trade Common, 342 F.3d 1361, 1371 (Fed.Cir.2003), *cert. denied,* 124 S.Ct. 2390 (2004); *see also* Nazomi Communications, Inc. v. ARM Holdings, PLC, 403 F.3d 1364, 1369 (Fed.Cir.2005) (discussing that claims may embrace "different subject matter than is illustrated in the specific embodiments in the specification").

The written description of the '670 patent focuses almost exclusively on the TDL process, thereby lending support to BODI's argument. Nevertheless, the specification does not state that TDL is the actual invention, rather it recites that TDL, in the context of the preferred embodiment, is the process created for describing the data definitions, manipulations, and other types of transformations in the data warehousing domain. The Federal Circuit has instructed that even in cases where the specification describes only one embodiment, that embodiment only rarely limits the scope of the claim term. *See, e.g.,* Innova/Pure Water, Inc., 381 F.3d at 1117 (stating that "even where a patent describes only a single embodiment, claims will not be read restrictively unless the patentee has demonstrated a clear intention to limit the claim scope using words or expressions of manifest exclusion or restriction" (internal quotation omitted)); SRI Int'l v. Matsushita Elec. Corp of Am., 775 F.2d 1107, 1121 n. 14 (Fed.Cir.1985) ("That a specification describes only one embodiment does not require that each claim be limited to that one embodiment."). It is the claims that measure the invention, not the specification. Interactive Gift Express, Inc. v. Compuserve, Inc., 256 F.3d 1323, 1331 (Fed.Cir.2001) ("In construing claims, the analytical focus must begin and remain centered on the language of the claims themselves, for it is that language that the patentee chose to particularly point out and distinctively claim the subject matter which the patentee regards as his invention." (internal quotation omitted). The specification here does not contain any statements of limitation or disclaimers that limit the definition of port to something other than the plain meaning.

The doctrine of claim differentiation also undermines BODI's proposed construction. The Federal Circuit has recognized that, while not a hard and fast rule of construction, the doctrine of claim differentiation creates a presumption that each claim in a patent has a different scope. Comark Communications, 156 F.3d at 1187. This presumption "is especially strong when the limitation in dispute is the only meaningful difference between an independent and dependent claim, and one party is urging that the limitation in the dependent claim should be read into the independent claim. Sunrace Roots Enters. Co. v. SRAM Corp., 336 F.3d 1298, 1303 (Fed.Cir.2003). Normally, "limitations stated in dependent claims are not to be read into the independent claims from which they depend." Karlin Tech., Inc. v. Surgical Dynamics, Inc ., 177 F.3d 968, 972 (Fed.Cir.1999).

A review of all the claims at issue, supports the conclusion that BODI's proposed construction is untenable.

Dependent claim 19 adds to independent claim 18 only the clause: "wherein the ports are comprised of a name, a data file, and a flow type." Similarly, dependent claim 36 adds to independent claim 30 only the clause: "wherein the ports are comprised of a name, a data file, and a flow type." The juxtaposition of the independent claims lacking any reference to name, data file, and flow type with the dependent claims that add this limitation provides strong support for Informatica's argument that the independent claims were not intended to require the presence of the details contained in BODI's construction. Adopting BODI's proposed construction would import the nuances of the dependent claims into the independent claims; the dependent claims thereby would be rendered entirely redundant. BODI has not come forward with a persuasive argument to rebut the presumption that the independent claims should not be so limited as to render the dependent claims superfluous.

The Court adopts Informatica's proposed construction and construes the term "port" to mean: **An abstraction for describing the inputs and/or outputs of sources, targets or transformation objects.**

2. "transformation object"

The term "transformation object" appears in numerous claims of the '670 patent (claims 1-5, 7-18, 20-35, 38-39) and three claims of the '775 patent (claims 1, 9, 11). Informatica points to the plain meaning of the term, as understood by one skilled in the art, and proposes the term be construed to mean "A reusable component for processing data according to predefined instructions." BODI proffers the construction "A transformation object is an object which transforms (*i.e.,* modifies) data according to some predefined behavior or rule. An object has two parts. The first part has a declaration for the ports and attributes of the object, and the second part has a definition of the behavior of the object. A transformation object has input and output ports. Different transformation objects perform different, unique functions, and are selected according to the functions which need to be performed in order to generate the desired transformations."

As with the previous term, BODI advocates for a narrow construction of this term and relies on language in the intrinsic record. None of the indicia of a patentee's intent to limit the claims is present here. The intrinsic record provides no indication that the patentee acted as its own lexicographer during the prosecution and explicitly defined the term "transformation objects" to differ from its otherwise ordinary meaning. CCS Fitness, 288 F.3d at 1366. Nor is there any evidence before the Court establishing that the patentee distinguished this term from prior art on the basis of a particular embodiment or expressly disclaimed subject matter. *Id*. Finally, as discussed above, the specification does not state that TDL is the actual invention, nor is there any clear and unmistakeable disavowal of the claim scope, "Absent a clear disclaimer of particular subject matter, the fact that the inventor may have anticipated that the invention would be used in a particular way does not mean that the scope of the invention is limited to that context." Liebel-Flarsheim Co. v. Medrad, Inc., 358 F.3d 898, 909 (Fed.Cir.2004) (internal quotation omitted). The Court declines to limit the claim to the preferred embodiment disclosed in the specification.

The Court adopts Informatica's proposed construction and construes the phrase "transformation object" to mean: **A reusable component for processing data according to predefined instructions.**

3. "sequence transformation object"

The term "sequence transformation object" appears in three claims of the ' 670 patent (claims 13, 25, 39) and three claims of the '775 patent (claims 1, 9, 11). Information proposes the term be construed as "A transformation object that generates keys from an initial value." BODI proposes the term be construed as:

A transformation object that is used for creating unique keys for records as they are processed in a mapping. Each instance of a sequence transformation is created with an initial value, which is used at the start of an execution, an increment value to compute the values of subsequent indexes, and an end value. Default values will be used if any of these parameters are omitted. This transformation has two predefined output ports, curval and nextval, that contain the current value and the next value of the sequence index, respectively.

A sequence has the following specification:

CREATE Sequence <sequence_name> (curval INTEGER OUT, nextval INTEGER OUT, [Startvalue [<integer>] ATTR,] [Increment [<integer>] ATTR,] [Endval [<integer>] ATTR,] [Cycle [<boolean>] ATTR,] [Cache [<boolean>] ATTR,] [Reset [<boolean>] ATTR,])

Once again, BODI's proposed construction comes directly from the specification of the 670 and '775 patents, and the parties present essentially the same arguments regarding the appropriate use of the limitations disclosed in the specification. For the reasons stated above, the Court declines to import those limitations into the claim language. The Court adopts Informatica's proposed construction and construes the phrase "sequence transformation object" to mean: **A transformation object that generates unique keys from an initial value.**

The '990 Patent

4. "transformation component"

The term "transformation component" appears in several claims of the '990 patent (claims 1-7, 9-15). Informatica contends that the term "transformation component" is synonymous with the previously defined "transformation object," from the '670 and '775 patents, and thus proposes the two terms be construed to have the same meaning. Informatica, proposes: "A reusable component for processing data according to predefined instructions." BODI proposes: "A transformation component is the same as a transformation object, which is a software binary file acting as an individual unit that possesses built-in autonomy and encapsulates the functionalities of a transformation. The autonomy of each transformation software component is manifested in two dimensions. The first dimension of component autonomy relates to how each transformation component implements staging (storing) the incoming data fields as it processes these data fields. The second dimension of component autonomy relates to how each transformation component automatically selects is own mode of manipulating data input/output. Each transformation component either pushes data to another transformation component, or performs a push/pull operation on the data."

The Detailed Description of the '990 patent specifically states that "[i]n the following detailed description of the present invention, some of the interchangeable key terms relating to the present invention are collected and defined in the section below to resolve possible ambiguity and to facilitate future reference." The precise term at issue here is listed immediately below this statement. "A transformation component is the same as a transformation object, which is a software binary file acting as an individual unit that possesses built-in autonomy and encapsulates the functionalities of a transformation." These statements evince the patentee's intention to depart from the ordinary meaning of transformation component. The introductory language makes clear that it is discussing the components of "present invention." The statements in the specification about the meaning of a transformation component are not merely exemplary, rather they define the patentee's understanding and use of the term transformation component in light of the claimed invention. "Because the inquiry into the meaning of claim terms is an objective one, a patentee who notifies the public

that claim terms are to be limited beyond their ordinary meaning to one of skill in the art will be bound by that notification, even where it may have been unintended." Innova/Pure Water, Inc., 381 F.3d at 1117; *see also* Watts v. XL Sys., Inc., 232 F.3d 877, 883 (Fed.Cir.2000) (concluding that the written description limits the invention by "stating that 'the present invention utilizes [the] feature' ").

The Court concludes that the Detailed Description includes a clear lexicographic definition, and accordingly construes the phrase "transformation component" to mean: A **transformation component is the same as a transformation object, which is a software binary file acting as an individual unit that possesses built-in autonomy and encapsulates the functionalities of a transformation.**

5. "staging data in a first said plurality of transformation components"

The phrase "staging data in a first said plurality of transformation components" appears in claim one of the '990 patent. Claim one reads (with the disputed phrase in bold):

1. A computer implemented method for transforming data in a data warehousing application, comprising the steps of:

specifying at least one source containing data;

constructing a plurality of transformation components for manipulating data according to pre-determined sets of rules;

coupling the transformation components to form one or more pipelines;

specifying a target for storing data generated by one or more of the pipelines;

staging data in a first of said plurality of transformation components; and

streaming data in a second of said plurality of transformation components, wherein said staging and said streaming of data are performed automatically by software without human intervention.

Informatica proposes the definition "In a first transformation component, storing some amount of incoming data for processing." BODI proposes the construction "Staging data is the storing of incoming data fields as the transformation component processes the data fields. The degree of requisite staging by each transformation component is automatically determined and implemented, without any human intervention. Depending on the nature of the transformation, each transformation component will automatically select the optimal amount of staging. The staging can range continuously from zero staging (also known as streaming) to full staging."

Informatica contends that the term "staging" when used in the context of data is generally understood in the field to mean storing data prior to processing. Further, Informatica maintains that the patent sets out that staging can range continuously from zero, which is referred to as streaming, to full staging. Therefore, staging involves storing some amount of incoming data. BODI derives its proposed construction from the Summary of the Invention section of the specification. The complete text surrounding the language BODI urges states (with the BODI's proposal in bold):

The autonomy of each transformation software component is manifested in two dimensions: The first dimension of a component autonomy relates to how each transformation component implements staging (storing) **the incoming data fields as it processes these data fields. The degree of requisite staging by each transformation component is automatically determined and implemented, without any human intervention. Depending on the nature of the transformation, each transformation component will automatically select the optimal amount of staging. The staging range continuously from zero staging (also known as streaming) to full staging.** A transformation with zero staging is called streaming transformation. The second dimension of component autonomy relates to how each transformation component automatically selects its own mode of manipulating data input/output. Depending on the nature of the transformation, each transformation component, together with the server, automatically selects a push mode, a pull mode, or a push and pull mode that optimizes throughput.

Informatica posits that BODI's construction is incorrect for two primary reasons. First, the portion of the specification upon which BODI relies is discussing only a preferred embodiment. Second, the proposed construction reads into this term specific features that are included separately in the claim language, namely that the requisite degree of staging of data is performed automatically by software without human intervention. The Court agrees that BODI's construction results in redundancy in the claim language and renders portions of the remaining claim language superfluous. Claim one explicitly teaches that "staging and said streaming of data are performed automatically by software without human intervention." It is unlikely that the patentee contemplated that the phrase "staging data in a first said plurality of transformation components," appearing earlier in the same claim, also contained that feature. Indeed the additional text would be unnecessary if persons of skill in the art understood that staging data in a first said plurality of transformation components inherently serve such a function.

The Court adopts Informatica's proposed construction and construes the phrase to mean: **In a first transformation component, storing some amount of incoming data for processing.**

The '374 Patent

There are five terms at issue in the '374 patent. All five terms appear, *intra alia*, in claim one, which reads (with the disputed terms in bold):

1. A database system for storing database system for storing data, comprising:

a **global repository** containing metadata;

a **shared folder** residing in the **global repository,** wherein the **shared folder** contains metadata which is to be shared;

a first **data mart linked** to the **global repository;**

a **reference** residing within the **first data** mart, wherein the **reference** specifies a **link** to of [sic] the metadata in the **shared folder** of the **global repository** and a user accesses the metadata through the first **data mart** by using the **reference** to read the metadata residing in the **shared folder.**

6. "data mart"

The term "data mart" appears in several claims of the '374 patent (claims 1-5, 7-12). Informatica proposes

the construction "Software that extracts, transforms and loads data into one or more databases, where the database(s) contain a subset of corporate data useful for decision support for an aspect of a business." BODI proposes the construction "An analytical database containing a subset of corporate information useful for decision support for an aspect of a business."

The crux of the parties' dispute is whether a data mart in the context of the '374 patent is limited to a passive database or incorporates both the end product target database and the associated metadata and applications used to build and maintain a data mart database, i.e. a data mart environment. Informatica posits that the term data mart and the term data mart application FN2 should be construed similarly. The Court finds that there is no support in the patent for this proposition and declines to adopt Inormatica's proposal. In addition, while the Court is mindful that the patent as a whole teaches a "method and system for accessing and sharing metadata amongst a number of data marts through the use of object referencing," neither the specification nor the claims describe data marts as having software functionality.

FN2. The parties previously reached an agreement regarding the construction of data mart application. They agree that a data mart application is software that extracts, transforms and loads data into one or more databases, where the database(s) contain a subset of corporate data useful for decision support of an aspect of a business.

The specification states that a "data warehouse is comprised of an analytical database containing information useful for decision support." "Data marts are similar to data warehouses, except that data marts usually contain only a subset of corporate data which is directed towards a singles aspect of that business (e.g., a separate finance data mart, sales data mart, human resources data mart, etc.) ." BODI's proposed construction combines these two concepts and moreover, comports with the ordinary meaning as understood by one skilled in the art.

The Court adopts BODI's construction and construes the phrase "data mart" to mean: **An analytical database containing a subset of corporate information useful for decision support for an aspect of a business.**

7. "global repository"

The term "global repository" appears in claims one through five of the '374 patent. Informatica proposes the construction "An abstraction for a database that stores metadata that is also accessible from one or more local database repositories." FN3 BODI proposes the construction "A repository is an abstraction for a database where metadata is stored. There are three types of repositories: independent, linked, and global. A global repository is distinguished from independent and linked repositories in that it is the only repository which stores metadata that can be shared with other linked repositories." The parties agree that the general meaning of the term "repository" is "an abstraction for a database," as set forth in the specification.

FN3. Informatica indicated in its opening brief that it does not object to the use of BODI's synonymous "linked repositories" language in lieu of "local database repositories."

The parties proffer similar constructions, but for BODI's proposed language regarding the three types of repositories, independent, linked, and global. The Court finds the term global repository can be construed

without reference to and discussion of the other types of repositories. The Court construes the phrase "global repository" to mean: **An abstraction for a database that stores metadata that can be shared with other linked repositories.**

8. "shared folder"

The term "shared folder" appears in claims one and two of the '374 patent. Informatica proposes the construction: "A folder whose contents may be accessed through or by a local repository." FN4 BODI proposes the construction "A folder represents the abstraction for grouping related objects and metadata in a repository. Folders can be of type regular or shared. The content of a shared folder can be accessed by other folders in the same repository, and by other linked repositories if the shared folder is in a global data mart repository through the use of a 'reference.' "

FN4. Again, Informatica indicated in its opening brief that it does not object to the use of BODI's "linked" modifier in lieu of its proposed modifier "local."

Informatica argues that the term should be given its plain meaning. BODI counters that the term has no ordinary meaning in the art, and thus points the Court to the specification. Informatica's proffered construction is too broad and provides insufficient detail to be useful going forward. The specification and drawings contemplate that the term "folder" represents an abstraction for grouping related objects and metadata in a repository or data mart. There is no indication that this definition is limited to a preferred embodiment. At a minimum, the final construction should contain this statement.

BODI's proffered construction, on the other hand, is too narrow. The proposed construction reads in detail that is unnecessary to the definition of a "shared" folder, *i.e.* that folders can be either regular or shared. Further, the construction includes a specific feature that is addressed separately in the claim language, namely that the content of a shared folder in a global data mart repository can be accessed by other linked repositories. This feature in the claim language would be unnecessary if persons of skill in the art inherently understood that the contents of a shared folder located in a global data mart repository could be accessed by other linked repositories. Moreover, the final construction can be structured to convey similar meaning to the term without reading in a limitation from the preferred embodiment.

The only remaining issue is whether the definition of "shared folder" must include the limitation that a reference is used to access the metadata in the shared folder, A review of the patent reveals that the invention was directed, in part, at a "method and system for accessing and sharing metadata amongst a number of data marts through the use of object referencing." The Summary of the Invention teaches that "[g]iven proper authorization, the metadata is accessed through the use of object referencing." Finally, the detailed description states that the "actual mechanism which allows for the sharing of metadata is through the use of 'object references .' " Thus, the specification teaches that the invention, as a whole (and not merely as a preferred embodiment), requires that the content stored in the shared folders is accessed through the use of a reference.

The Court construes the phrase "shared folder" to mean: **A folder represents the abstraction for grouping related objects and metadata in a repository. The contents of a shared folder may be accessed by other folders in the same repository and/or by other linked repositories through the use of a reference.**

9. "reference"

The term "reference" is found in four claims of the '374 patent (claims $1, 4, 7, 9$). Informatica proposes the construction "A denotation of shareable metadata." BODI proposes the construction "Metadata is accessed through the use of an object reference, *i.e.*, a shortcut. Basically, a reference points to the object stored in the global repository. Rather than storing an object directly in each of the data marts desiring to access that particular object, a reference to that object is stored instead in those data marts. By using multiple references, the same object can be accessed, used, and shared by multiple users across many different data marts." Once again, Informatica's proposed construction is too vague and BODI's proposed construction is too narrow. The Court declines to adopt either construction.

The patent discloses in several places throughout the specification that a reference is an abstraction for pointing to an object, which may be stored in various places. This definition is consistent with the use of the term in claims one, four, seven, and nine. The claims instruct that references residing in various data marts "specify a link to metadata in the shared folder," or "point[ ] to the metadata residing within the global repository" or "specify a link to where the metadata is stored in [a particular] database." Thus, in the context of the '374 patent, a reference is an abstraction for either denoting or pointing to the contents of a shared folder, which may reside in either the same repository or in the global data mart repository.

The Court construes the term "reference" to mean: **A reference represents an abstraction for pointing to the contents of a shared folder within the same repository or in the global data mart repository.**

10. "link"

The final term is "link," which appears in claims one and seven of the '374 patent." Informatica proposes that the term be construed to mean "connection." BODI proposes the construction "A link is a connection between a reference in a data mart and specific metadata in a shared folder in a global repository. The linking enables those data marts the ability to gain access to the objects stored in the global repository." Informatica argues that the term link has a plain and ordinary meaning, *i.e.*, a connection, that would be understood by one skilled in the art. BODI agrees that generally speaking a link is a connection, but in the context of the '374 patent, connection is too generic and overly simplistic.

BODI's proposed construction, requiring a link to connect a reference in a data mart and specific metadata in a global repository, fails to acknowledge the use of the term link in a broader sense in claim seven and throughout the specification. Claim seven discloses a method for sharing metadata between a plurality of databases, comprising, *intra alia*, a link between the second database and the first database. The Abstract and the Summary of the Invention discuss that "[a]ny number of data marts can be linked to the global repository." Figure 2 depicts a global data mart repository "connected to a number of linked data marts" (as represented by the arrows pointing from the global data mart repository to the data marts). The description of Figure 2 explains that "[b]y registering a repository with a [global data mart repository], the necessary link between the two repositories is established, and thereafter, the linked repository is able to share metadata with other repositories linked to the same [global data mart repository]." Finally, the detailed description states that the present invention provides for a sharing feature, which is "accomplished by linking certain designated data marts" with the global data mart repository.

When engaging in claim construction, the Court is bound by the presumption that "the same terms appearing in different portions of the claims should be given the same meaning unless it is clear from the specification and prosecution history that the terms have different meanings at different portions of the claims." Fin

Control Sys. Pty, Ltd. v. OAM, Inc., 265 F.3d 1311, 1318 (Fed.Cir.2001). BODI's narrow construction of the term link is unworkable in light of the obvious broader use of the term throughout the patent.

There is no need to go beyond the plain and ordinary meaning here. The Court adopts Informatica's proposed construction and construes the term "link" to mean: **Connection.**

CONCLUSION

Based upon the analysis set forth above, the Court adopts the foregoing constructions of the disputed terms. The parties are ordered to submit a further joint case management report pursuant to Patent Standing Order para. 13 within 21 days of the filing of this Order.

IT IS SO ORDERED.

N.D.Cal.,2005.
Informatica Corp. v. Business Objects Data Integration, Inc.