

# Use of Z39.50 for the Delivery of Current Awareness Products

Peter N. Ryall, LEXIS-NEXIS (email: [peterr@lexis-nexis.com](mailto:peterr@lexis-nexis.com))

## **ABSTRACT**

In the context of LEXIS-NEXIS, a Current Awareness Product (CAP) is an information product which provides concise, relevant data to a customer, related to a particular topic (or topics) of interest to that specific customer, or to customers within a broader market or industry group. This paper will discuss how Z39.50 is being used to deliver Current Awareness products to a variety of customers in a wide range of user environments. One essential objective of this delivery facility is to provide the necessary relevant information, proactively, to each customer in their own native environment, with minimal deviation from their standard methods for interacting with their groupware and/or individual workstation applications.

An assortment of Z39.50 services is used in this facility to provide a flexible delivery platform that supports such unconventional features as: information alerts, uploading record usage information for accounting and billing purposes, initiating subscriptions to specific Current Awareness products, establishing information filtering profiles, and providing access to (and delivery of) products consisting of 'webs' of pre-fabricated, related result sets.

## **Z39.50 Conformance Caveats**

The application discussed in this paper was implemented and initially deployed in mid-1994, so that it was not able to take advantage of more recent Z39.50 Implementor's Group (ZIG) thinking and decisions with regards to particular Z39.50 services. In addition, some of the required features of the CAP implementation were driven by the need to accomplish all CAP access and delivery interactions within the same protocol. As Z39.50 was chosen as this single unified protocol, some compromises were required to accomplish this objective. The following is a brief statement of the intentional deviations from conformance to the Z39.50 Version 3 standard, along with a set of predictions as to how and when these deviations will be brought in line with the standard.

All of the non-conformant features have been implemented within private groupings of Origin and Target systems. They are not being recommended for use in public Z39.50 Target or Origin implementations, nor are any of them being proposed as enhancements to the Z39.50 standard at this time.

The first CAP feature that fits into this category is the use of Resource Control to deliver 'alerts' from the Target to the Origin system. The decision to use Resource Control in this non-standard way was strongly driven by the need to deliver all product functionality using the Z39.50 protocol. Although alerts could more naturally be sent using an offline

delivery vehicle such as FAX or email, the use of Resource Control reports to convey alerts allowed us to provide 'real-time' reporting using long-running Z39.50 Associations. In future releases, delivery of these alerts will be migrated to other vehicles (FAX and email) in environments where real-time notification is not required.

The second non-conformant CAP feature is the use of a 'virtual Record 0' within a Result Set to provide access to set-level metadata. Record 0 is not defined by the Z39.50 standard as a valid record for retrieval, so this usage of Record 0 does not fit the model of "each record of the result set contains either a database record, a diagnostic record, or a surrogate diagnostic record". To bring this feature into conformance, a possible short-term (Version 3) solution would be to issue a search using the Result Set as the Operand, and requesting that the metadata be returned in additionalSearchInfo. A more flexible solution will be proposed for Version 4, possibly using Record 0 as defined here, or by defining a set model which includes set-level metadata as an extension to each record in the Result Set, thereby allowing it to be retrieved by simply Presenting set-level meta-elements from any record in the set.

The third non-conformant CAP feature is the usage of the implementationID, implementationName, and implementationVersion Init parameters to allow an Origin to connect to a specific set of back-end search and retrieval services. This is used mostly for testing and backwards compatibility purposes, and is not planned for use in public Z39.50 service implementations (as described in Section 4.0).

The final non-conformant CAP feature is the unconventional usage of Persistent and Transient Result Sets. Each of the resultSetIds conveyed in an Alert represents an identifier for a Transient Result Set (one which only endures for the lifetime of the Association). However, in this implementation of CGTI, the Target performs an automatic (and transparent) service for the Origin. It maintains a Persistent Result Set Task Package corresponding to each of the Transient Result Sets it reports to the Origin in an Alert.

Then, when the Origin issues a Present against the Result Set (using the Transient Result Set ID), the Target automatically asks the back-end retrieval service to fetch the essential elements of the Persistent Set and create a Transient Set. This Transient Set is created with a front-end mapping filter which allows it to be accessed via Present using the Persistent Result Set Package ID. Thus the Target and its associated back-end services present a 'virtual Persistent Result Set' image to the accessing Origin. Note: in subsequent Associa-

tions, the Origin still uses the same Transient Result Set ID to access records in the 'virtual Persistent Result Set'.

In a future release, the usage of Persistent and Transient Result Sets will be brought into line with the Z39.50 standard by requiring the Origin to first issue a Present using the Persistent Result Set Package ID, in order to obtain a Transient Result Set ID to be used within the Association.

It is also important to note here that, while the CAP facility does not require the use of the Z39.50 Search service, the Target which supports CAP functionality also supports search services. Thus while the Target is compliant with the Z39.50 requirement to support Search, the Search service is not used by an Origin in accessing and retrieving information from CAPs.

## 1.0 Overview of CGTI and Current Awareness Products

Traditionally, LEXIS-NEXIS has provided extensive search and retrieval information services using interfaces and formats defined in-house. After free-text searching across our large information databases, users would then determine which result sets are relevant to their search. LEXIS-NEXIS Current Awareness Products (CAPs) address an audience with broad or focused needs for information about specific industries and markets. CAPs are designed as a value-added product consisting of predefined sets of information organized by subject areas. External user environments access CAPs and other search and retrieval services through our Coarse Grain Transaction Interface (CGTI).

The "coarse grain" nature of CGTI reflects a more loosely-coupled interface which allows access to a broader, more flexible family of services than the traditional LEXIS-NEXIS interface. The CGTI is based on the ANSI/NISO (National Information Standards Organization) Z39.50 search and retrieval standard. The Z39.50 standard includes protocol specifications for search and retrieval of information stored in machine-readable databases.

### 1.1 Definitions and Description

This paper introduces and describes the CGTI, by which a broad range of Current Awareness Products (CAPs) may be accessed via an external system. It presents the set of message-based requests that may be constructed and sent to the Information Service (Z39.50 Target) by an external client (Z39.50 Origin). Each Origin request (and associated Target response) is described in detail, and examples are presented showing how each CAP capability is accessed.

The interface used to access CAPs (as well as other search and retrieval services) is based on Version 3 of the Z39.50 standard, which is specifically designed to meet the needs of client environments desiring access to information search and retrieval services such as those provided by LEXIS-NEXIS.

In this paper, the LEXIS-NEXIS implementation of the Z39.50 search and retrieval standard is referred to as the CGTI. Because CGTI is less tightly-coupled to the internal environment of the search and retrieval system, a more open and consistent interface is provided, allowing both the information provider (i.e., LEXIS-NEXIS) and external system and workstation developers to mature and enhance their systems independently of one another.

### 1.2 CGTI

The CGTI model consists of three distinct components:

- Third-party platform domain, consisting of the Z39.50 Origin and local user applications residing on a specific platform;
- Delivery domain, consisting of the lower layer communications protocol components;
- LEXIS-NEXIS domain, consisting of the Z39.50 Target and back-end information products (CAPs) and services.

The third-party platform domain consists of the Z39.50 Origin, a set of APIs, and local user applications residing on a user-specific platform. A set of environment-independent interfaces (APIs) allows local applications to initiate transfer activity through the CGTI to the Z39.50 Target. This transfer activity takes the form of an exchange of message requests and responses. The developer can write code using Z39.50 protocol data units (PDUs) or use a library of functions defined in specialized toolkits to communicate across the CGTI.

A single toolkit function may call one or more PDUs to execute multiple Z39.50 services. LEXIS-NEXIS provides a toolkit (called the "Origin Adapter Toolkit") for this purpose.

This flexible arrangement allows the end user access to CAPs from a variety of environments such as:

- Information desktop workstations/services such as Folio and Lotus Notes
- Message handling systems such as AT&T Personalink and other value-added distributor systems

- Mass market distributor systems like Prodigy
- Integrated corporate and third-party information systems

The delivery domain provides the logical and physical bridge that links the Target and the Origin. A physical connection must be made between the Origin and the remote Target system. This connection can be established over a variety of physical communication mechanisms, such as: leased lines, TCP/IP, or X.25. A logical connection, or Z39.50 association, must also be established between the Origin and Target. Messages from the Origin and Target are translated into Z39.50 requests/responses before transmission through the delivery domain.

Figure 1-1 illustrates the CGTI model.

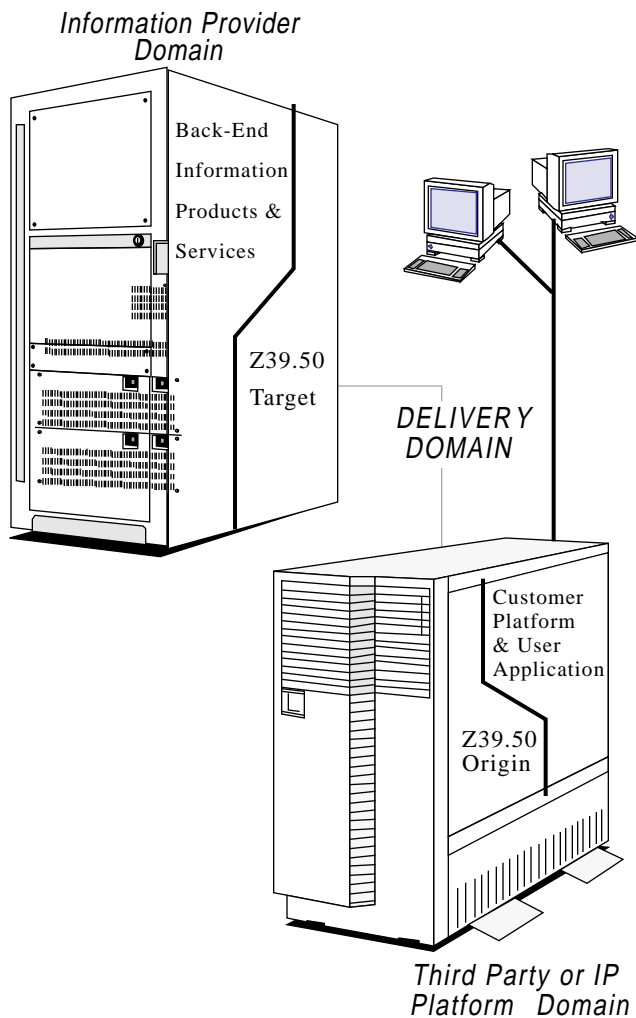


Figure 1-1 CGTI Model

The LEXIS-NEXIS domain consists of the CGTI Target and back-end information products and services. The CGTI Target translates:

- Incoming CGTI requests/responses into function/procedure calls;
- Function call responses/callbacks into the corresponding outgoing CGTI requests/responses

Some situations require that back-end services be called to execute request processing and/or data manipulation.

The CGTI "Origin Adapter Toolkit" provides the necessary functions to allow a local application to access a CAP without the complications of building an Origin supporting the CGTI Z39.50 profile (private to LEXIS-NEXIS).

### 1.3 Current Awareness Products Conceptual Model

The basic conceptual model of Current Awareness Products is as follows:

- LEXIS-NEXIS, through its advanced authoring facilities (combining human Subject Experts and automated processes), creates value-added relevancy-based information products which are designed for use in specific industries and markets, and for access by users who have either broad or focused needs for relevant information related to specific subject areas;

**Note:** These products are referred to as Current Awareness Products to distinguish them from standard search and retrieval services, which provide free-text searching across large subsets of the information warehouse but do not guarantee the return of highly-focused, relevant answer sets.

- LEXIS-NEXIS makes these products available via the CGTI, using a minimal subset of the Z39.50 standard requests to provide the following capabilities:
  - access to, and retrieval of, one or more directories of CAPs that are potentially available to end users attached to the client (Z39.50 Origin) system;

**Note:** These directories may be presented in an 'active' way to the end user - e.g., as forms which allow the user to subscribe to an individual CAP; once subscribed, the user could receive new information (documents/articles) automatically whenever the CAP is updated.

The CGTI server will notify the CGTI client when new material is available; the degree of automation in obtaining results is determined by the client system implementation, and can vary from real-time alerts at the user's workstation, to notifications sent through a store/forward mechanism (e.g., E-Mail system), to requiring the user to check a shared folder or database where new results are placed as they are received.

In the LEXIS-NEXIS CAP implementation, one option is for the server to send notifications (alerts) to the client via a non-standard mechanism which uses the Z39.50 Resource Control request to convey the alert data. The specifics of these alerts and how they are encoded is described later in this paper.

- the ability for the CGTI Origin to establish a persistent (long-running) Alert Specification for a particular CAP. This Specification instructs the CAP Provider to notify the Origin system whenever the specified CAP is updated with new information;
- once an Alert Specification is established, and the Origin system is notified of new information in a topic/subject category, the Origin may then retrieve one or more of the documents in the updated category;
- the ability to access the directory (table of contents) of a CAP on an episodic basis; this involves retrieving the desired directory from the Target, presenting it to the user, and then allowing the user to select and navigate through the hierarchy of the directory;
- also on an episodic basis, the ability to request retrieval of one or more of the available forms of the documents/articles in a specific topic/subject category of the CAP directory (examples of possible forms are 'cite-list', 'preview', 'full-text', and 'full-text with graphic');
- the ability to access the descriptive and hierarchy (parent/child) information related to a specific topic/subject category.
- A future capability to be supported for access to CAPs is the capability to search through a particular product directory (or the master di-

rectory of all products accessible to an Origin system). The Origin will be allowed to issue a Z39.50 Search request that searches across a product (or all products) for a specified topic/subject of interest, or a set of terms. The Origin will be able to specify whether it wants the search to be:

- restricted to the Topic names and descriptions of the various categories;
- restricted to the abstract/preview portion of the documents in each topic category, or;
- free to search across the entire full-text of the documents in each category.

## 1.4 CGTI User Model

There are two classes of users who gain access to LEXIS-NEXIS services via the CGTI: *direct* and *indirect*. This classification of users is important to the model of CGTI CAP services outlined in this paper, so it is described in some detail below.

A *direct* (end) user is an individual user who accesses the system via a direct CGTI connection, either through a workstation which contains a Z39.50 Origin, or via a multi-user (server) system which executes the Origin within the server, but 'passes through' any information from the user dealing with ID, password, and/or token authentication, so that the user is still directly attached to the system. The attributes of a *direct* user are:

- data about the end user is contained in the customer information database; it is entered by an administrator at the time the user signs up for the service, and it is modified whenever changes are required;
- in order to access LEXIS-NEXIS services, a user must first go through an individual sign-on process and provide identifying information to the Authentication Service, which will return a ticket to the user's system, granting permission to access authorized services and products;
- records are logged for billable events initiated by the user, specifically identifying the user as the billable party;
- typically, invoices for charges incurred by the user for usage of services will be delivered directly to the user (or the user's firm with itemization by user);
- when users experience difficulty connecting to the system or using a particular service, they will contact Customer Services (i.e, the CAP

Provider supplies the first line of Customer support).

An *indirect* user, on the other hand, is one who accesses the system via some intermediate (or agent) system, which may or may not be owned, serviced, and/or managed by the Provider. The intermediate system is (logically) the *direct* user of the system, as it contains the knowledge of the user ID(s) and password(s) needed to access Provider services, and it is responsible for managing the CGTI interactions with the Target system. The customer's system must provide the access points to its end users, including local authentication and authorization, communications connectivity, customer support, customer sign-up and subscription services, individual user billing and invoicing services (generally via a charge back system of some type) and anything else which requires identification and tracking of individual end users and their activities. The unique attributes of an *indirect* versus a *direct* user are:

- the Information Provider (IP) maintains no persistent information about *indirect* end users - they are known only within the administrative domain of the customer firm; the IP also has no involvement in signing up or subscribing *indirect* users to specific services;
- a system within the customer's firm is responsible for 'signing on' to the Provider's service; the end user interacts directly only with his/her local system to login and enter a password (if necessary);
- events visible to the IP will be logged (for billable events initiated by any user attached to the customer system), but no data is logged specifically identifying the end user;
- typically, invoices for charges incurred by users for usage of Provider services will be delivered to the customer firm or the third party agent, who is responsible for any charge back billing or invoicing to individual end users within the firm, or under the administrative control of the agent;
- when users experience difficulty connecting to the system or using a particular service, they will contact their local administrator or help desk (i.e., the IP provides only the second line of Customer support).

For those CGTI services which require that the authenticity of the Origin system's identity needs to be verified, the IP will return a ticket to the Origin, which must be attached to the Origin's request in

order to access the service (for both *direct* and *indirect* users).

Multiple User ID's may be permitted from a single Origin system; depending on the classification of the end user according to the descriptions above:

- for *indirect* users, the Origin system manages User IDs to grant different spans of authority to different users and/or local applications. The IP will follow a standard Authorization process for these IDs. As a part of this process, the Origin system will be given tickets for each of these IDs, which will permit access to different groups of Provider services and products.
- for *direct* users, the end user's workstation manages User IDs to allow multiple users to use the same workstation.

## 2.0 How CGTI Uses the Z39.50 Standard

The CGTI is one LEXIS-NEXIS implementation of the Z39.50 search and retrieval standard. At present, CAPs are delivered using only the following Z39.50 facilities of CGTI:

- Initialization
- Extended Services
- Resource Control
- Retrieval
- Termination

CGTI also supports the Z39.50 Search facility, but as it is not required for delivery of Current Awareness products, it is not discussed in this document. The following sections provide brief descriptions of the five facilities listed above, and tell how the CGTI utilizes these facilities to provide access to Current Awareness Products.

### 2.1 Initialization Facility

The CGTI Target supports the Z39.50 Initialization facility exactly as defined in Version 3 of the standard. The Origin sends an Init request to the CGTI Target, including setup information such as operations that should be supported, user authentication, and message size.

An "accept" result from the Init response indicates that the association is established and the Origin can proceed to access CAPs for which it is authorized. If the association is unsuccessful, the Origin can attempt another initialization.

## 2.2 Extended Services Facility

The Extended Services facility implemented in the CGTI system allows the Origin to:

- Set up a delivery notification mechanism (notifying the user of additions and updates to a CAP) by issuing an Alert Specification Extended Service;
- Send confirmation of billable information delivery to the user via a Final Delivery Notice Extended Service;
- Send a Usage Accounting Report Extended Service to the Target indicating CAP usage within the confines of the external delivery system provider;
- Send a Subscription Accounting Report Extended Service to the Target containing end user subscription requests.

## 2.3 Resource Control Facility

As mentioned earlier, Resource Control is used in a non-standard way by the CGTI, and as such it is only discussed here to provide completeness to the overall CAP delivery system description. In a future release, there is a plan to migrate this feature over so that it uses more standard Z39.50 facilities (e.g., Search and Present against a standard database containing ‘product update’ records, as well as discontinuing the use of Resource Control for delivery of alerts, in favor of email and FAX alert delivery).

The Resource Control request is issued by the CGTI Target to notify the Origin of the availability of new information topics or additional documents for a CAP. In response, the Origin sends a Resource Control response which indicates to the Target that the Origin is ready to receive another alert. The Origin can send a request to retrieve the information after the Resource Control response is sent to the Target.

In order to start this CAP notification process, the Origin must first establish an ‘Alert Profile’ by sending an ‘Information Alert Specification’ Extended Service request to the Target (see Section 5.1 “Creating and Sending an Alert Specification”).

## 2.4 Retrieval Facility

The CGTI Target supports the Present and Segment services. As with any standard Target, these services define how result records appear when retrieved from the appropriate database.

Information retrieved from a CAP is contained in a result set maintained by the Target. The result set is a data structure with a pointer indicating where a record is located within the appropriate database; therefore, records are referenced by their position within the result set.

The Present service allows the Origin to request the retrieval of records from a specified result set. The Origin issues a Present request specifying a range of records that should be retrieved. The request message can also specify subsets of the records that define the “view” of what the user sees. These “views” include Cite, Preview, and Full.

Full database records may consist of document text or metadata. See Section 6.1 for an explanation of the various classes of metadata.

The CGTI Target supports Level 1 Segmentation, which allows large documents (records) to be broken down into manageable “fragments” for transfer to the Origin.

## 2.5 Termination Facility

In the CGTI system, the Close service operates according to the Version 3 standard, allowing either the Origin or Target to terminate a Z39.50 association. Reasons for termination include system problems, security violations, protocol errors, lack of activity, and completion of a ‘user session’. The Close request terminates a single Z39.50 association between the Origin and Target. The recipient of the Close request responds with a Close response confirming the termination.

## 3.0 How to Access Current Awareness Products via Z39.50

To retrieve information from one or more Current Awareness Products, a Z39.50 Origin initiates a sequence of Z39.50 request/response exchanges with the Target. The following list defines the functions which must be implemented when developing a CGTI Origin:

- Establish an Association
- Identify new information
- Retrieve information
- Provide notification, usage, and subscription information
- Provide diagnostic and error message information
- Terminate an association

In order to start a Z39.50 Association, the Origin will first need to establish a physical connection with the Target system, and then send a Z39.50 Init request to establish an (application level) Association. Once an Association has been established, the Origin will typically set up a delivery notification mechanism for new CAP information. Given the fact that CAPs are regularly updated, a client will most likely want the Target to create an alert which notifies the Origin/client when new information is available.

Once the Target sends notification of new CAP information, the Origin can retrieve that information from the CAP Provider. Documents (records) can be retrieved at different hierarchical levels (topics/subtopics) and can be delivered in different “views” (e.g., cite lists, previews, full document text), by specifying different Element Set Names in the Z39.50 Present request.

After setting up retrieval options, the Origin should specify how it wants the Information Provider to track CAP usage and activity for billing and accounting purposes. In many cases, it will be desirable to allow the Origin to create and send notices/reports that provide notification, subscription, and usage information to the Target for processing.

#### **4.0 Establishing a CGTI Association**

Before attempting to access Current Awareness Products, an Origin must first establish a Z39.50 connection with a CGTI Target, just as it would with any other Target. This process consists of these steps:

- Making a physical connection between the Target and Origin environments using TCP/IP or X.25 (using either a leased line or a dial-up connection)
- Establishing a Z39.50 association between the Target and Origin

In order to conform to the CGTI model, the client must set the proper Init Option flags to indicate that it supports Present, Resource Control, Extended Services, and Close.

In order to distinguish among LEXIS-NEXIS Targets which support different functionality, the three Init parameters implementationId, implementationName, and implementationVersion may be used to specify a particular LEXIS-NEXIS Target implementation with which the Origin desires to interact. This is only used for testing out new implementations of Target functionality, such as a beta release of CAP capabilities.

For backwards compatibility, new values may be used to allow new CAP features to be provided to Origin systems wishing to take advantage of the new features. For Origin systems not wanting to implement or take advantage of new features, the Target continues to support the previous Target implementation, which the Origin indicates by specifying null (default) values in these three parameters.

#### **5.0 Identifying New Information**

Once the Origin has established an association with the Target, the next step is to set up an automated notification facility that alerts users of changes to the CAP(s). The effectiveness of a CAP is degraded if the product information is inaccurate and/or out-of-date. A CGTI product domain can update CAPs on a daily, weekly, or monthly basis.

CAPs may change due to the availability of new source information, the roll-out of new CAPs, and scheduled promotions. Since a key component of a product’s value is determined by the CGTI server’s ability to deliver that product in a timely manner, CGTI provides an automated notification capability via the Z39.50 interface.

This process of identifying new information consists of these steps:

- Creating and sending an Alert Specification to the Target using Extended Services;
- Receiving one or more alerts from the Target using Z39.50 Resource Reports.

The following sections describe the alert notification cycle and other related functions.

#### **5.1 Creating and Sending an Information Alert Specification**

The Present service allows product information records/documents to be retrieved via a synchronous request from the Origin. Product Alerts (usually delivered via Resource Control reports in the current release) provide an asynchronous notification mechanism to an Origin system.

An Origin system must send an Alert Specification (‘profile’) to initially activate delivery of Product Alerts from the CAP Provider system to the Origin system. The Alert Specification instructs an agent within the CAP Provider domain to search for new information in a specified subject area (or attributed to a particular CAP) and to notify the Origin whenever any relevant information is found.

The Origin specifies how alerts are to be managed by the Target by the setting of the Action parameter in the Alert Specification Extended Service. When this parameter is set to 'queued', the Target is instructed to send all queued alerts related to the specified product and topic immediately upon completion of the Extended Service exchange.

The functional flow of the Product Alert capability is as follows:

- First, the Origin sends an Extended Services request to establish an Alert Specification at the CGTI Target system,
  - as a part of the request, the Origin specifies a set of parameters, to be used by the Target in delivering notifications of new, relevant information:
    - The Alert Delivery Vehicle, which is generally set to 'Z39.50 Resource Control' for the first release of CGTI CAP services;
    - the Product Name, which specifies the name of the CAP to be tracked;
    - an Action, which contains either 'queued' (meaning that all the alerts on the 'new information' queue at the Target are delivered to the Origin, but once the queue is cleaned out, additional alerts will not be sent until another Alert Specification is sent), 'realtime' (meaning that alerts can be delivered to the Origin at any time after the Alert Specification is created, and will continue to be sent until the Alert Specification is deleted), or 'refresh' (meaning re-send or 'refresh' the entire CAP, including Topics and documents previously delivered to the Origin);
    - a flag (alertCombinations Desired) specifying whether the Origin is prepared to handle multiple topics and/or multiple topics for multiple products in a single alert, or whether it will only handle one topic per alert (a topic is a single category of information in a CAP which covers a single subject area - e.g., baseball scores under the Sports CAP);
    - if alertCombinationsDesired specifies "multipleTopicsPer Alert", the maxTopicsPerAlert parameter indicates the maximum number of Topics

which are allowed to be packaged into a single Resource Control 'Alert'.

- the Target will next save the Alert Specification (if valid) and reply with an Extended Services response specifying that the request is valid and has been processed (i.e, the Alert Specification has been created);
- once the Specification is created, the Target then proceeds to send alerts to the requesting Origin system (based on the setting of the action parameter) whenever new documents are received for the specified product(s), using Z39.50 Resource Control 'Alert' reports.

## 5.2 Sending Information (Product) Alerts

Once an Alert Specification is created, the Target is responsible for sending alerts to the Origin whenever new topics are added to a CAP, or a CAP is updated with new records/documents.

Resource Control requests, each containing one or more Alerts, are issued by the Target to notify the Origin of the availability of new CAP topics, deleted topics, or additional documents added to an existing CAP topic. In response, the Origin sends a Resource Control response telling the Target whether or not it wants to continue receiving Alerts. The Origin can then send a Present request to retrieve the new records, after sending the Resource Control response to the Target.

As stated earlier in this paper, this use of Resource Control for sending alerts is not conformant with the use of Resource Control as defined within the Z39.50 standard. These Resource Control 'alerts' do not correlate with any specific request (as required by the Z39.50 state tables) nor do they relate to a Resource Report pertaining to the entire association. In future releases of the CAP delivery system, notification mechanisms (such as FAX and Email) will be used for delivery of alerts, thus deprecating Resource Control as an alert delivery mechanism.

The Target will issue a request with a referenceId, a resourceReport ('Alert', see definition in 5.3 below), and a responseRequired flag set to 'ON'.

Within each alert is an alertAction parameter; it can be set to either "new", "update", or "remove". If it is set to "new", this is a new Topic (i.e, it has not been retrieved by this Origin previously). If alertAction is set to "update", this is an existing Topic which has been updated with new documents. "Remove" indicates that the designated TopicID has



been removed from the Product; if the Topic is a hierarchy node, it indicates that all the 'child' nodes have been removed from the Product as well.

ResultSetID, of course, indicates the ID of the Result Set which corresponds to the latest version of the Topic 'set'. The Origin uses this as the Result Set ID when it builds the Present request to retrieve records from the Result Set (see the Note on 'virtual Persistent Result Sets' below). NumItems specifies the total number of records in the 'Topic' result set.

AlertsQueue is a parameter which indicates how many remaining alerts are waiting to be sent relative to this Product Alert Specification. If alertCombinationsDesired is set to "multipleTopicsPerAlert" or "multipleProductsPerAlert" in the Alert Spec, then the Topics structure will carry a sequence of potentially multiple pairs of topicPath, resultSetId parameters. The Origin will send a response by returning the referenceId received during the request, in addition to a continueFlag set to "ON".

**Note:** Each of the resultSetIds conveyed in an Alert represents an identifier for a Transient Result Set. However, in this implementation of CGTI, the Target performs an automatic service for the Origin. It maintains a Persistent Result Set Task Package for each of the Transient Result Sets it reports to the Origin in an Alert. Then, when the Origin issues a Present against the Result Set (using the Transient Result Set ID), the Target calls the back-end retrieval service to create a Transient Set from the Persistent Set. This Transient Set may then be accessed via Present using a result set name that is actually associated with the Persistent Result Set Task Package. Thus the Target and its associated back-end services present a 'virtual Persistent Result Set' image to the accessing Origin.

The SetType parameter indicates whether this set (which is the subject of this alert) is a hierarchy 'node' or a leaf 'node' in the Product structure (see "Understanding CAP Hierarchy" in Section 6.1 below). In simple terms, a leaf node set always contains only records/documents (or pointers to records), whereas a hierarchy node set may contain records, but also contains pointers to sets subordinate to itself in the Product hierarchy.

### 5.3 LN-RR-1 Resource Report Definition

The following is the LN-RR-1 Resource Report definition, which is a privately-registered Resource Report (RR) type to be used by LEXIS-NEXIS for asynchronous notification of events (in this case, Information Alerts).

```
RR (1.2.840.10003.7.1000.14.3) DEFINITIONS ::= BEGIN
```

```
LNResourceReport ::= IMPLICIT SEQUENCE {
    resourceReportId IMPLICIT OBJECT
        IDENTIFIER,
    -- specifies an OID to identify this Resource
    -- Report type. The following is the structure
    -- of the CAP 'Information Alert', which
    -- is used to notify the Origin of newly-received
    -- relevant CAP information.
    operation [0] IMPLICIT VisibleString,
    --set to "alert" for this RR class
    alertTopicsQueue [1] IMPLICIT INTEGER,
    numTopicsThisAlert [2] IMPLICIT INTEGER,
    topicNodeRecord [3] IMPLICIT SEQUENCE
        OF SEQUENCE {
            topics [1] IMPLICIT SEQUENCE OF
                SEQUENCE {
                    topicId [0] IMPLICIT VisibleString,
                    resultSetId [1] IMPLICIT VisibleString,
                    numItems [2] IMPLICIT INTEGER
                    alertAction [3] IMPLICIT INTEGER {
                        new (1),
                        update (2),
                        remove (3) },
                    setType [4] IMPLICIT INTEGER {
                        hierarchy (1),
                        leaf (2) }
                } OPTIONAL } },
}
-- End of LN-RR-1 Resource Report definition
END
```

## 6.0 Retrieving Information

Retrieving CAP information is the primary objective of an Origin system which is accessing Current Awareness Products. Each Origin has a defined set of CAPs that it is authorized to access. This is verified through the user ID of the Origin system (see discussion of CGTI User Model in Section 1.4). Each set of user CAPs is known as a root product set. Each root product set consists of:

- Descriptive information about the root product (result) set
- Product (result) sets accessible to the end user and the topics associated with each product set

You can retrieve information from these product sets with or without alert notification. The procedure for retrieving information is similar; however, you must understand the hierarchical nature of a CAP before attempting to retrieve CAP information.

### 6.1 Understanding CAP Hierarchy

A CAP is designed as a hierarchical product with a tree-like structure. The top level of the tree is a topic or information category associated with the product. Subsequent levels of the tree consist of subtopics associated with the topic or subtopic at the next higher level. Each topic that contains one or more subtopics is considered a hierarchy node. The lowest level of a hierarchical system consists of leaf nodes. There are no subtopic levels associated with leaf nodes.

A CAP can also be a flat product with no underlying tree structure (no hierarchy nodes).

From a Z39.50 perspective, each version (e.g., daily update) of a CAP topic corresponds to a unique result set and has a corresponding result set ID. Each version of a subtopic at a hierarchical level also corresponds to a result set. An Origin can request different ‘views’ or subsets of the records in a CAP result set, depending on what ‘views’ it wants to see. Records in a result set can contain the following types of information:

- Set metadata information
- Documents relating to a topic
- If available, the parent of the topic and any other child subtopics

As stated previously, each CAP Topic/subtopic set contains set-level metadata and document data. To access the set-level metadata, the Origin issues a Present request against record (document) 0 of a valid CAP result set.

**Note:** The use of Record 0 for access to and retrieval of set metadata is not conformant with the Result Set model specified in the Z39.50 standard. However, it is being used as a convenient mechanism for access to this metadata. It is foreseen that the set metadata will be accessed via more conventional (and standard) Z39.50 facilities in future releases of CGTI. For instance, one possible solution would be to allow the set-level metadata to be retrieved from any valid record in a set (1-N), simply by specifying the correct set meta-element names/tags.

At a hierarchy node, the Origin can request the following elements of metadata describing a CAP topic or subtopic:

- Set class indicating whether the CAP is flat or hierarchical
- Topic or subject name of the CAP
- Result set name or ID
- Description of the result set or topic
- Existing child subtopics
- the root ancestor of this topic (the Product Name/ID) [future]
- an (optional) list of related/associated topics/subjects [future]

A hierarchy node (result) set contains only “directory” information such as the data structure with a pointer indicating where a record is located in the CAP Provider database. Records are referenced by their relative position within the result set.

At a leaf node, you can request different views of a document such as a:

- Cite list
- Preview
- Full Document text

A leaf node result set is made up of the individual records (documents) in the set. Each record contains the full document content and descriptive information such as author, title, document publication date, and so on. The format of the document can be in ASCII text, SGML, or in any other type supported by the CGTI system.

Figure 6-1 illustrates the CAP hierarchy concept.

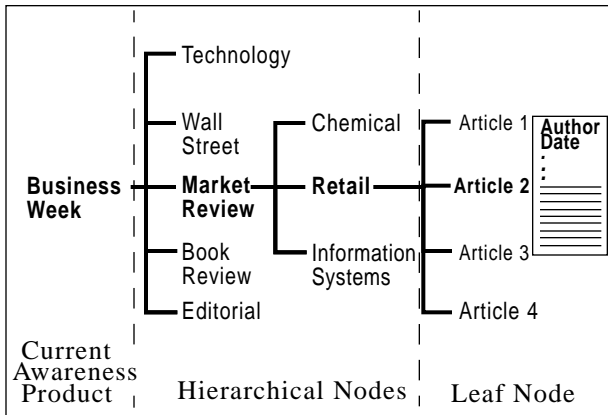


Figure 6-1 CAP Hierarchy

## 6.2 Retrieving Data Information

The CGTI Target allows an Origin to initiate data retrieval using the Present and Segment services. When the Origin issues a Present request, it uses designated Present elements to define the view of the retrieved information. A typical flow looks like this:

- Suppose you want to know what topics are associated with a product. The Origin sends a Present request specifying the result set ID of the product;
- The Target generates internal retrieval responses and returns an aggregate Present response, i.e. zero or more Segment responses followed by a Present response. A single document (within a CAP topic category) is sent to the Origin in each response.
- You can choose to view any level of the product hierarchy, or documents associated with a topic, by issuing additional Present requests.

**Note:** Level 1 Segmentation is used by CGTI services to break down large result sets into segments (records) which fit into the specified maximum message and record sizes.

## 6.3 Retrieving New or Updated Information

Suppose the Origin receives an alert triggered by the availability of new CAP topics or the addition of new documents to a CAP. The flow for retrieving alert information is similar to retrieving data information. You still need to send the Present request; however, you need only the result set ID for the new topic alert or new document alert (see the Note in

Section 5.3 on the use of 'virtual Persistent Result Sets' within the current CAP implementation). Like the data information retrieval process, you can select different views of the new or updated information.

The Origin initiates the Z39.50 Present service by sending a Present request to the Target. Typically, the Origin returns to the client application when the Present response is received in full.

The preferredRecordSyntax parameter in the Present request is not currently used by the Target. The current release of CGTI supports only a 'default' record syntax (using a LEXIS-NEXIS private OID) which does not provide encapsulation of the record contents; however, in future releases, both SUTRS and GRS-1 will be supported.

The presentElements parameter is used to specify the desired subset/view of the records expected in the present response. For Release 1, this will simply consist of an elementSetName, which will be structured to contain the desired document/record 'view', in addition to the desired document format.

In future releases, the eSpec-1 structure will be used in place of elementSetName to describe more complex composition specifications by which to retrieve the records.

The following are the document 'views' supported in the current release of CGTI:

- CITE - retrieves "headline" information about a topic
- PREVIEW - retrieves an abstract of a document
- FULL - retrieves entire document contents
- SUBINFO - retrieves subtopic information
- TOPICINFO - retrieves set metadata information

The Origin can also specify the text format of the retrieved documents using these values as part of the Element Set Name:

- FASCII - formatted ASCII text
- UASCII - unformatted ASCII text
- GSGML - text tagged with SGML tags using the 'generalized CAP DTD'

So, using the legal values above, an example of an ElementSetName is 'CITE;FASCII', which would be used to request retrieval of headline information in formatted ASCII.

## 7.0 Delivery Notification, Usage, and Subscription Information

Delivery and document usage activities are tracked by LEXIS-NEXIS to provide usage, billing, and subscription information. These activities provide revenue based on price schedules for documents, document usage, delivery services, and subscription services. Various notices and reports are generated by the Origin to provide tracking information. The Extended Services facility allows the Origin to send this information to the Target and to generate a Final Delivery Notice, Usage Accounting Report, and Subscription Accounting Report.

### 7.1 Providing Delivery Notification

The Final Delivery Notice (FDN) is an Extended Services task that allows the Origin to send confirmation to the Target that a document has been delivered. This confirmation contains the following types of information:

- The delivery status
- The date and time the document was delivered
- The retail price, in cents, charged to the end user
- The suggested wholesale price

The FDN is used in situations where the Origin initiates a Present request to retrieve a document and passes the document directly to one or more end users.

The process of creating and sending an FDN to the CGTI Target is similar to the alert specification process. The Origin builds an FDN Extended Services request, which creates an FDN parameter package at the CGTI server that captures the report. After processing, the Target returns an Extended Services response indicating whether or not the notice was delivered.

### 7.2 Providing Document Usage Information

Mostly on behalf of *Indirect* end users (see Section 1.4), external delivery systems within the Origin domain are required to capture data about CAP document usage. They then send this data to the Target in the form of a Usage Accounting Report (UAR). The UAR contains these types of information:

- The number of copies of a document that were delivered to the Origin
- The date and time the document was delivered

- The retail price, in cents, charged to the end user
- The suggested wholesale price

The UAR is used in situations where the Origin initiates a Present request to retrieve a document and stores the results (document) locally. The Origin then delivers these documents to end users directly from local storage, rather than from the Target. The Origin periodically sends UARs to the Target indicating what and how many documents have been delivered.

Similarly to the FDN, the Origin builds a UAR Extended Services request, which creates a UAR parameter package at the CGTI server that captures the report. After processing, the Target returns an Extended Services response indicating whether or not the report was delivered.

### 7.3 Providing Subscription Information

End users have the ability to request subscriptions to specific CAPs. Mostly on behalf of *Indirect* end users (see Section 1.4), external delivery systems within the Origin domain capture subscription request data and send the data to the Target in the form of a Subscription Activity Report (SAR). The SAR tracks subscription information such as:

- The number of requested subscriptions
- The length of the subscription in months
- The retail price, in cents, charged to the end user
- The suggested wholesale price

Similarly to the FDN and the UAR, the Origin builds an SAR Extended Services request, which creates an SAR parameter package at the CGTI server that captures the report. After processing, the Target returns an Extended Services response indicating whether or not the report was delivered.

## 8.0 Terminating a CGTI Association

The preceding sections have described the CGTI Z39.50 profile which allows an Origin system to establish an association, identify new information, retrieve information, and create administrative reports. As with the Initialization of a Z39.50 association, termination of the association conforms to the Z39.50 Version 3 specification. Reasons for termination range from a security violation to internal system errors at the Origin or Target.

The termination process consists of these steps:

- Terminating the Z39.50 association between the Target and Origin using the Close request
- Terminating the TCP/IP (or X.25) connection between the Target and Origin environments

A Close request can be issued from either the Origin or Target. The reason for termination is specified within the Close request. After processing the request, the Origin or Target returns a Close response.