

Self-Service Syndication with ICE

Building Informative Web Pages and Catalogs Automatically

BY DAN R. GREENING

Newspapers, product retailers, and Web portals face a common problem: How can they provide the most up-to-date content? They can invest in developing their own original content, as does *Web Techniques* magazine, or they can assemble material from several outside sources and rebrand it under their own name. The *San Francisco Chronicle* newspaper assembles its comic page by buying comic strips from King Features and Marvel. Wyle Electronics creates product-information Web pages by assembling data sheets from electronics manufacturers. Excite buys news from Reuters and UPI. Checkout.com buys movie, music, and game information from All Media Guide and offers it along with DVDs, CDs, and games in an online store, creating an "entertainment buying experience."

Providing the best possible information on products in a timely fashion is an important form of customer service. All other factors being equal, customers tend to shop with retailers that give the best presale product information. This is forcing online retailers to become information portals where consumers can go to find out more about products (and incidentally, buy those products).

But the retailers don't have the best information—manufacturers do. They have the greatest incentive to create highly informative product collateral—presale brochures, specifications, manuals, and rebate coupons—to help customers buy

and use a product. If manufacturers syndicate collateral on the Internet and let retailers subscribe to it, customers can obtain enormous detail on a product before and after the sale. With syndication, online retailers should have a great advantage over brick-and-mortar stores, where maintaining detailed product information is very costly.

The Problem

At present, online syndicators usually provide content through proprietary or roll-your-own solutions. Some subscribers assemble and rebrand syndicated content by placing links to a syndicator's Web site. Others perform a nightly download from syndicator FTP sites. And sometimes syndicators create and maintain Web sites under different names (kind of an inverted syndication) to which the subscribers can safely link.

These ad hoc forms of syndication are fraught with technical problems. Usually a subscriber is assembling material from several syndicators. On the subscriber side, adding content from a new syndicator involves engineering and Web-design effort. On the syndicator side, every new subscriber has to be trained to use the syndication system, deployment becomes much more dependent on high-demand Web skills, and in the end customers usually get less information.

Ad hoc approaches typically require human involvement to negotiate the simplest operational issues. Does the

subscriber have to credit the author or syndicator? Can the syndicator push the content at a particular time of day? Can the subscriber edit the material? Is there an incremental charge for each piece of content? Can the subscriber download only some of the content offered by the syndicator?

In short, without a standard protocol, syndication doesn't scale very well on the Web.

ICE Syndication

To address these problems, a consortium of application server and content companies, led by Vignette, created a standard syndication protocol based on XML—Information and Content Exchange (ICE). XML is a simple standard to represent data-hierarchies using familiar HTML-style tags. The ICE protocol standardizes the following functions:

- A potential subscriber requests a catalog of subscriptions offers.
- A syndicator responds with a subscription-offer catalog, each offer detailing the type of content, usage restrictions, and available delivery methods, times, and frequencies.
- A subscriber subscribes to one or more offers, negotiating for specific delivery methods and times.
- A subscriber pulls a package of subscribed content from a syndicator.
- A syndicator pushes a package of subscribed content to a subscriber.
- A syndicator or subscriber cancels or changes a subscription.

Self-Service Syndication with ICE

LISTING ONE

```
<?xml version="1.0"?>
<!DOCTYPE ice-payload SYSTEM "http://www.ice-ag.org/ICE.dtd">
<ice-payload payload-id="12984751223"
  timestamp="13:31:21" ice-version="1.0">
  <ice-header>
  <ice-sender sender-id="5fde2359-439c-1201-56ab-1964dca39034"
    name="Andromedia Aria 3.0" role="subscriber"/>
  </ice-header>
  <ice-request request-id="1998-07-05T02:02:23@xyz.com-1">
    . . .
  ...ICE Directives...
  . . .
  </ice-request>
</ice-payload>
```

LISTING TWO

```
Request:
...request-header...
<ice-get-catalog/>
...end-request-header...

Response:
...response-header...
<ice-catalog>
<ice-contact description="Contract Approvals"
  name="Infocetus Corp" url="http://www.infocetus.com/">
  Brad Husick
  Syndication Czar
  Infocetus Corporation
  1452 Burns Drive
  Black Rock City, NV 89412
  +1 (775) 555-1212
</ice-contact>

<ice-offer description="Comic:Daily:Bonko Dog and Me"
  rights-holder="Sea Slug Comix"
  showcredit="true" atomic-use="true">
<ice-delivery-policy startdate="2000-01-01"
  stopdate="2000-12-31T23:59:59">
<ice-delivery-rule mode="pull" starttime="00:00:00"
  duration="P14400S" maxcount="1">
<ice-negotiable type="starttime"/>
</ice-delivery-rule>
</ice-delivery-policy>
<ice-business-term lang="en" type="payment">
Net-30 acceptable
</ice-business-term>
<ice-business-term lang="en" type="licensing">
Syndicator makes no claims as to the validity of this material.
Rights holder is solely responsible for this content.
</ice-business-term>
</ice-offer>

<ice-offer description="News:Continuous:Events"
  rights-holder="Smart Art News">
<ice-delivery-policy>
<ice-delivery-rule mode="pull" starttime="00:00:00"
  duration="P86400S" maxcount="12">
</ice-delivery-rule>
</ice-delivery-policy>
<ice-business-term lang="en" type="payment">
Free service with embedded advertising.
</ice-business-term>
<ice-business-term lang="en" type="licensing">
Syndicator makes no claims as to the validity of this material.
Rights holder is solely responsible for this content.
</ice-business-term>
</ice-offer>

<ice-offer-group description="Local:Art">
<ice-offer description="Local:Art:Daily:Burning"
  rights-holder="Larry Harvey">
<ice-delivery-policy> ... </ice-delivery-policy>
... </ice-offer>
<ice-offer description="Local:Art:Daily:Electroluminescent"
  rights-holder="Larry Harvey">
... </ice-offer>
</ice-offer-group>
</ice-catalog>

...response-header...
```

LISTING THREE

```
<ice-package new-state="2" old-state="ICE-ANY" package-id="pkg1124-2"
  subscription-id="S001012">
<ice-item-group item-group-id="2" name="ProdFolders1" >
  <ice-item
    item-id="1333"
    name="100301"
    content-filename="/pf1/10/100301.xml"
    subscription-element="/pf1/10/100301.xml"
  ><?xml version="1.0" standalone="yes" ?>
  <!--DOCTYPE document [
  ]>
  <!--document-->
  <!--ProductFolder-->
  <!--MENU title="General Description" -->
  <!--[CDATA[
  <!--p-->The 100301 is a monolithic triple 5-input OR/NOR gate.
  All inputs have 50 k Ohm pull-down resistors
  and all outputs are buffered.<!--p-->
  ]>
  <!--/MENU-->

  <!--MENU title="Features" -->
  <!--[CDATA[
  <!--ul-->
  <!--li-->23&#amp;#37; power reduction of the 100101<!--li-->
  <!--li-->2000V ESD protection<!--li-->
  <!--li-->Pin/function compatible with 100101<!--li-->
  <!--li-->Voltage compensated operating range = -4.2V to -5.7V<!--li-->
  <!--li-->Standard Microcircuit Drawing (SMD) 5962-915280<!--li-->
  <!--ul-->
  <!--/MENU-->

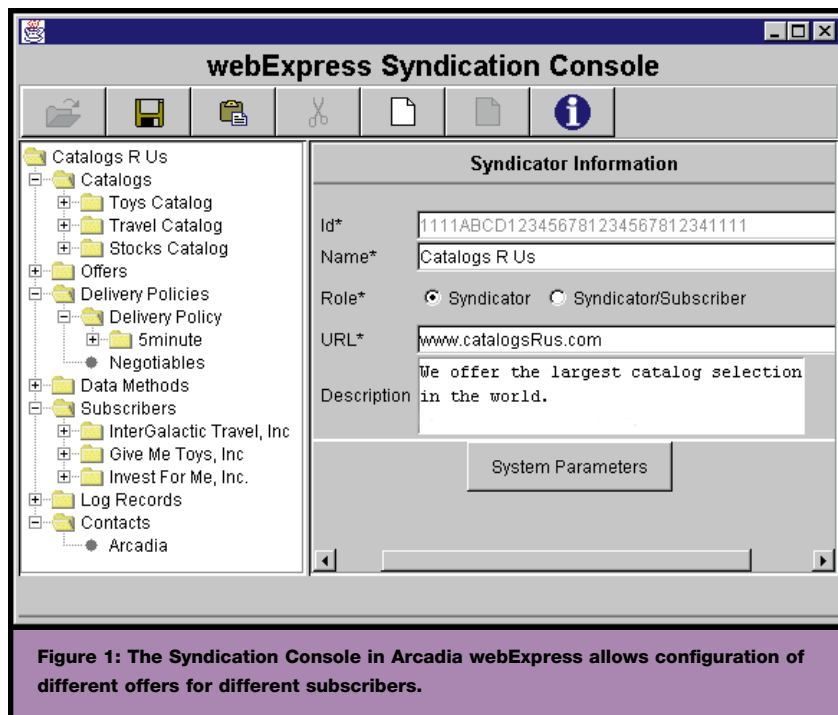
  <!--/ProductFolder-->
  <!--document-->
</ice-item>
  <ice-item
    item-id="1334"
```

Related Standards

ICE is one of the most innovative uses of XML, in part because ICE is mainly a protocol with a little bit of data, while most XML standards focus solely on data. ICE doesn't even specify the format of the syndicated content data. In XML terms, it's just bytes inside an `ice-item` entity. The character data inside an `ice-item` can be structured using one of the industry-specific XML data representations—BizTalk, RosettaNet, CommerceOne, WDDX, and so on.

Other standards efforts are often placed in the same box with ICE, because Web publishing is only now becoming automated. The World Wide Web Consortium's Resource Description Framework (RDF) and related standards specify how descriptive terms can be attached to content files, allowing content to be identified and selected according to a filtering or sort criteria. These frameworks can be used together with ICE, letting subscribers select offers using complex criteria.

The syndicate/subscribe model ICE defines is almost the same as what computer scientists call "publish/subscribe." And it turns out that ICE is most similar to binary publish/subscribe protocol standards, such as CORBA and DCOM. But in ICE, messages are delivered through XML, typically delivered over an HTTP connection, as opposed to a lower-level binary protocol. ICE is much easier to read and use, but it is also much more verbose. If you're constrained by network bandwidth, either compress the ICE packets or use something else. Finally, ICE defines many typical syndication operations and constraints



Ad hoc forms of syndication are fraught with technical problems. To address them, a consortium of application server and content companies, led by Vignette, created a standard syndication protocol based on XML—Information and Content Exchange (ICE).

that CORBA and DCOM leave to vertical industry implementations.

ICE provides only minimal security and access-control facilities. Typically the HTTP transport layer handles security, using such technology as SSL. Access control is handled on the syndication server, which performs authentication through the usual Web server password-access system, and provides different subscription offers to different subscribers.

Request-Response

The ICE protocol defines a set of request-response pairs coded in XML. The ICE standard doesn't specify the underlying transfer protocol, but does suggest an implementation using the HTTP POST/response mechanism called "ICE/HTTP". The body of the HTTP POST contains the ice-request, and its associated HTTP response contains the ICE response. As far as I know, all current ICE

ICE Standardization

ICE is truly a demand-driven standard. Vignette recognized the need for a syndication protocol, but quickly realized it had to obtain wide industry adoption to make it work. It sought syndication-software vendors (including competitors), syndicators, and subscribers to participate in an ad hoc standards group. Most were part of an advisory council; the active core became the ICE Authoring Group.

Any syndication protocol would fail if it didn't address the needs of end users, so the ICE Authoring Group decided that syndicators and subscribers should comprise a minimum of 50 percent of the participants. No new vendors could join unless there were enough new end users to provide a balance. The result is a standard shaped and driven by market need.

Some of the first ICE applications were surprising. Despite the participation of large news syndicators in the standardization effort, the first apps came from manufacturers who used ICE to publish parts catalogs. These companies continue to push the boundaries of the ICE standard heavily. One vendor, Arcadia Technologies, is now building ICE tools and middleware specifically for parts catalogs.

Andromedia adopted ICE for a unique purpose: syndicating and subscribing to Web visitor events. Andromedia's network monitors sniff packets, plug in to Web servers or application servers, and then syndicate the results through the ICE protocol. Web-traffic analyzers and personalization servers that adapt to visitor behavior are ICE subscribers, and on occasion they are ICE syndicators as well. Andromedia standardized on ICE to let partner companies easily integrate new dynamic content servers, event monitors, and analyzers with its products.

Although the ICE standardization group remained ad hoc until the standard was nearly complete, it needed a formal organization to promote the standard and approve changes. It selected an organization centered around publishers, the Graphics Communication Association, rather than a Web- or computer-centric group. The advantage is that users exert greater control than engineers; however, a possible disadvantage is that the standard may not undergo rigorous technical evaluation. —DG

Self-Service Syndication with ICE

ICE is one of the most innovative uses of XML, in part because ICE is mainly a protocol with a little bit of data, while most XML standards focus solely on data.

1/2
Island
4-1/2' x 6-1/2'

Ad
Color?
P.U. info
page No.
RSN ?

Keyline does not print

implementations use ICE/HTTP. This article assumes ICE/HTTP is the transport.

Every `ice-request` is contained within an `ice-payload`, which identifies the ICE version and the sender, and provides a `request-id`. Listing One shows a sample request payload. Most of the header is devoted to debugging information.

Upon receiving a request, the respondent creates a response payload, which contains whatever the subscriber asked for, if available. The response payload contains many of the same header tags as the request payload.

Catalog

Before syndication can occur, the syndicator must configure a syndication server, specifying which offers are available to which subscribers at what time. Figure 1 shows a syndication console that lets a syndicator add and delete catalogs, offers, delivery policies, and subscribers.

For a subscriber to know what content it can subscribe to, it needs to obtain a catalog of offers from the syndicator. This can be handled in two ways: In the old-fashioned way, the subscriber telephones the syndicator, asks "What's available?" and the syndicator provides a list. In the modern ICE way, the subscriber sends an `ice-get-catalog` request to the syndicator, which responds with a collection of offers and offer-groups.

Listing Two shows an example `ice-get-catalog` request, followed by a response. The ICE catalog first provides contact information for a person who can provide more details on the catalog. Then it provides a set of product offers.

Each offer includes the name of the content and the copyright owner. It can also include several typical usage constraints: `atomic-use` indicates that all items in the subscription must be offered to the user, otherwise individual items may be deleted from the presentation. `Editable` indicates the subscriber can modify the content. `Ip-status` values indicate the intellectual-property rights status of the subscription: `public-domain`, `free-with-ack`, `see-license`, `severe-restrictions`, or `confidential`. `showcredit` indicates that the subscriber must display the copyright owner with the content. `Usage-required` indicates that information regarding viewers of the content must be provided to the syndicator. Other constraints can be encoded in a mutual-

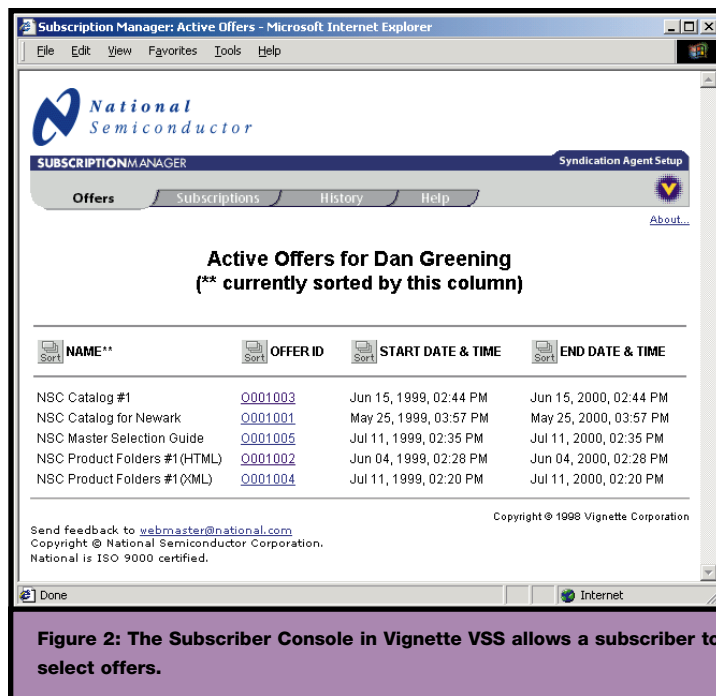


Figure 2: The Subscriber Console in Vignette VSS allows a subscriber to select offers.

ly agreed upon format and offered in a `constraints-url`.

Offers can be built in to a display hierarchy for convenient navigation using `ice-offer-group` tags. A subscriber can't subscribe to an `ice-offer-group`; the name of the offer-group is only a mnemonic. Subscriptions are made using the name in the `ice-offer` tag, even if they are embedded inside an `ice-offer-group`.

There are four offers shown in Listing Two: `Comic:Daily:Bonko Dog and Me`, `News:Continuous:Events`, `Local:Art:Daily:Burning and Local:Art:Daily:Electroluminescent`. The last two are organized under the `Local:Art` hierarchy.

Delivery Policy

With every offer is a set of delivery policies, including the delivery-mode, availability dates, and more detailed availability information.

Content can be delivered in pull or push mode. If the subscriber specifies pull delivery, the subscriber always makes the requests and the syndicator always responds—content is delivered only when the subscriber requests it. Pull delivery makes programming a subscriber quite simple: You get what you ask for when you want it.

Push delivery, on the other hand, requires the subscriber to run a Web server to handle pushed deliveries, which come in the form of an HTTP request from the syndicator. The HTTP request could contain a large payload with one or more articles. The subscriber usually confirms with an OK.

There are some fairly complex rules for specifying when and how often a subscriber obtains new content from a subscription. In Listing Two, `Bonko Dog and Me` can be pulled at most once per day (`maxcount="1"`) between midnight (`starttime="00:00:00"`) and 4:00am (`duration="P1440S"` specifies 14400 seconds from `starttime`).

Subscribing

To subscribe to an offer, the subscriber simply sends an `ice-offer` back to the syndicator in an `ice-request`. The offer is usually taken verbatim from an `ice-catalog`.

ICE Resources

ICE specification

www.gca.org/ice

The ICE specification provides a DTD and reference material for developing ICE-compliant applications.

Andromedia

www.andromedia.com

Andromedia uses ICE to syndicate visitor events from its customer behavior monitors to its high-traffic Web behavior analysis and personalization software. Partner vendors and customers can interoperate with Andromedia products by syndicating or subscribing to different types of customer behavior.

Arcadia Technologies

www.arcadiatech.com

Arcadia provides syndication software called `webExpress` primarily for business-to-business parts and product catalog syndication. The company intends to integrate its software with major application servers to provide seamless syndication regardless of the content-management platform.

iSyndicate

www.isyndicate.com

iSyndicate is a prominent online content syndication service. Individual artisans submit their work to iSyndicate, then iSyndicate handles royalties, distribution, and advertising. iSyndicate offers over 3000 sources of information.

National Semiconductor

ice.national.com

National Semiconductor is one of the first companies to use ICE for catalog syndication. It uses Vignette VSS. National rolled out its ICE syndicator in under two weeks. There are currently 12 distributors in National's syndication network. National claims it can enroll a new ICE subscriber in less than 30 minutes.

ShiftKey

www.shiftkey.com

Shiftkey provides Pure Java syndication software (including an ICE syndicator), and subscriber and Web-based subscription and site management.

Vignette

www.vignette.com

Vignette initiated the ICE standardization effort, and provides an ICE-compliant syndication system, Vignette Syndication Server. VSS allows Vignette StoryServer sites to syndicate content. VSS can manage subscriber groups, offering different subscription catalogs to different subscribers, and can monitor subscriber usage.

Xenosys

www.livebiz.com

Xenosys provides Java frameworks and Enterprise JavaBeans for financial and business transactions on the Internet. It provides an ICE-complaint Java API, called JICE, allowing customers to easily create applications that use the ICE protocol. —DG

Self-Service Syndication with ICE

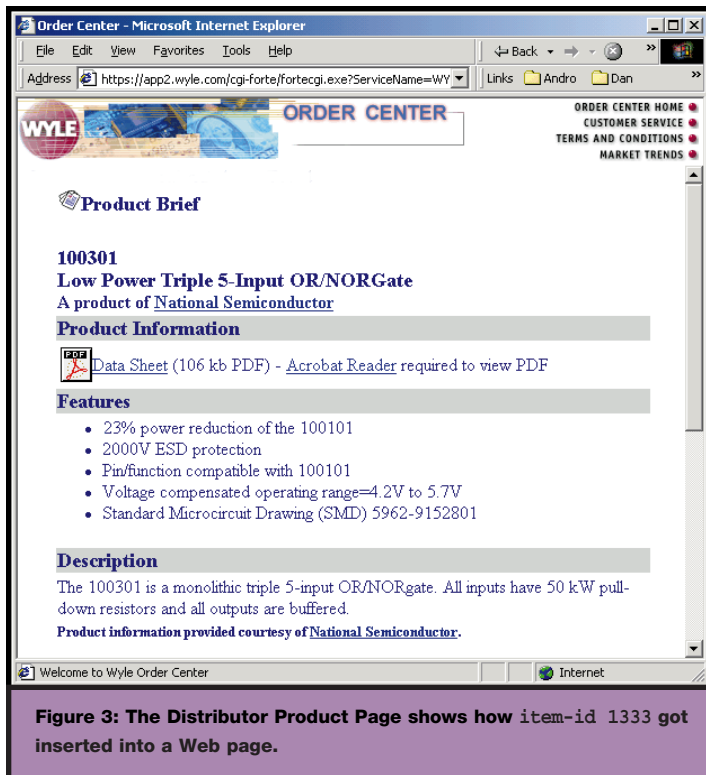


Figure 3: The Distributor Product Page shows how item-id 1333 got inserted into a Web page.

1/4
Page
3-1/4' x 4-5/8'

Ad
Color?
P.U. info
page No.
RSN ?

Keyline does not print

The subscriber can modify any field marked negotiable in an ice-catalog. If the syndicator responds with OK, then the negotiated offer is accepted. If the syndicator responds with Sorry, it was rejected with no further information. If the syndicator responds with a different ice-offer, the subscriber can consider it a counteroffer, and submit it back to the syndicator as an ice-request, fairly confident it will be accepted.

A catalog of ice-offers can be presented "the old-fashioned way," by sending offer information not through an ice-catalog, but rather via the Web, email, fax, or voice. People who process subscriptions then click on offers of their choosing or cut-and-paste to subscribe.

Figure 2 shows an example Web-page catalog from the National Semiconductor subscription site (ice.national.com). In this case, it is a "catalog of catalogs": It contains various National Semiconductor catalogs to which my site can subscribe. The NSC Product Folders #1(XML) subscription, for example, presents product information in XML on every National Semiconductor product. Clicking on an offer establishes a subscription.

Packages

To fulfill subscriptions, the syndicator sends directives in an ice-package to update a subscriber from an old-state to a new-state. ice-item, ice-item-ref, and ice-item-group entities specify additions, and ice-item-remove entities specify deletions.

If a package includes an activation field, the subscriber must not perform the operations before the specified time. For news items, this is typically a release time. A business might want earnings news released after the stock market closes. A politician might want a speech transcript released after a press conference. News editors tend to respect these time constraints, in part to ensure future access to "hot news." ICE can automate the delivery process.

Packages have several other attributes drawn from typical syndication requirements. Some of the fields are the same as ice-offer fields. These fields modify the package only, while the ice-offer fields refer to the entire subscription. A number of parameters that appear on the offer can also appear on the package entity—expiration, Atomic-use, Confirmation, Editable, Fullupdate, showcredit, and Exclusion. For example, the *Miss Manners* column might have an exclusionary clause that the article can't be used unless the author's picture is displayed.

Package Sequence

ICE forces packages to be processed in the order specified by the syndicator. The state of a subscriber can be defined by a single value—the package sequence identifier (PSI). Each package sent has the "old PSI" (the required subscriber state prior to receipt of the package) and the "new PSI" (the state of the subscriber following receipt of the package). Prior to subscribing, the subscriber is in the "empty state" indicated by "ICE-INITIAL". If it doesn't matter what the previous state was, a package can be sent with "ICE-ANY" as the PSI.

Using sequence IDs reduces or eliminates the state information stored on the syndicator side. Subscription management becomes the purview of each subscriber. The syndicator has to

remember a subscriber's state only if it supports push delivery.

PSI strings are opaque to the subscriber, except when they need to be compared for equality. This gives the syndicator enough flexibility to use an implementation-specific state encoding. For example, the implementation might use integers, time stamps, or a proprietary database key as the PSI.

National Semiconductor offers thousands of complex parts through distributors and retailers. By syndicating product information through ICE, National's distributors can provide the most up-to-date product information to customers. Syndication is very handy for such vendors.

Listing Three shows an example subscription item from the National Semiconductor XML catalog. Ice-item 1333 appears in the beginning, an XML description for part 100301. Figure 3 shows the same item integrated into a distributor's Web page.

It's tempting to say this article describes the tip of the iceberg, but ICE is a fairly simple standard. Most of the ICE features

omitted in this article relate to error handling and offer negotiation. For more details on ICE or vendors providing ICE applications, refer to the boxes titled "ICE Standardization" and "ICE Resources."


Conclusion

The ICE standard makes it easier for syndicators to deliver information in a controlled way to subscribers. In traditional publishing, writers, artists, composers, and producers "outsource" contract and delivery issues to syndicators, allowing the artisans to focus on their craft. Syndicators then achieve economies of scale by performing the same function for multiple artisans.

Some Web companies follow this traditional model. iSyndicate, for example, assembles content from individual Web pages, packages it, and allows portals to subscribe. If you're an author, it's easy to publish your material through iSyndicate. Using ICE, portals will be able to easily subscribe to syndicated content from iSyndicate.

As usual, the Internet changes tradition-

al definitions, because it makes automated negotiation possible and speeds information transfer. Syndication is no exception. On the Internet, a syndicator can be anyone with a large collection of uniformly structured data made available to multiple subscribers. This means parts manufacturers, icon collections, free software aggregators, stock-photo libraries, and even Web-traffic analyzers can be syndicators.

Now that the Web has matured, and people seek useful information among too much noise, obtaining the best-quality information for your visitors becomes more challenging. Typically, this means subscribing to, assembling, and controlling multiple sources of news and data. The ICE protocol makes it possible for site developers to do this using a single system, allowing everyone to spend more time on creative tasks. 

Dan holds a Ph.D. in computer science from UCLA. He is currently chief technology officer at Andromedia. He can be reached at greening@andromedia.com.

1/2 Horizontal

7' x 4-5/8'

Ad
Color?
P.U. info
page No.
RSN ?

Keyline does not print