EBSCO recognizes the importance of being able to accurately determine the source of user traffic. When users link directly from origin to destination, it is easy to see what site is sending traffic your way. However, when users link from origin through another site and then on to the destination, it becomes more difficult to capture the details needed to ascertain where the user started.

Consider that there are many flights available from LA to Boston, some of which go direct and others that stop in Denver, Chicago, and Philadelphia. All flights will get you to your destination, but those flights with layovers present a challenge for the person who has to pick you up at the airport, because those flight will show you coming from Chicago and not LA. The same is true for publishers attempting to trace the ultimate origin of their site’s incoming traffic.

Link resolvers such as Full Text Finder (EBSCO), SFX (Ex Libris), and 360 Link (Serial Solutions) often act as intermediaries in connecting discovery services to sites containing the full text, and the fact that link resolvers and discovery services may be available from the same companies but not necessarily used in concert by libraries causes further confusion. For example, if a library subscribes to EBSCO Discovery Service and uses SFX for their linking, publishers may attribute incoming traffic erroneously to Ex Libris.

EBSCO has developed this fact sheet to help familiarize publishers with the topic and the aspects involved in identifying user origin. In addition, EBSCO has initiated a conversation with NISO with the goal of forming industry-wide recommendations for tracking incoming traffic.

## NISO Origin Identification Project

In Fall 2015, Scott Bernier and Oliver Pesch from EBSCO proposed a new work item to NISO. The goal was to “Develop a recommended practice that would allow libraries, publishers and other content providers to accurately track the sites/platforms from which incoming links originate when they pass through a link resolver” (Bernier and Pesch). See the [proposal](http://www.niso.org/apps/group_public/download.php/15858/Tracking%20Link%20Origin%20in%20a%20Networked%20Environment%20Final%20for%20Voting%20Members.pdf) for more information.

On February 17, 2016, NISO announced its approval of the project in a [press release](http://www.niso.org/news/pr/view?item_key=c2ab810f6ee30db6113af9bd6638748d5caf94f7). To summarize,

Members of the National Standards Organization (NISO) have approved a new project: the development of a Recommended Practice for Tracking Link Origins in a Networked Information Environment. As libraries strive to improve the ways in which users access their collections, gaining a definitive understanding of where a user began his or her library search before ultimately arriving at library-licensed content is an important factor for library staff in determining the value of a platform and how to allocate resources. Additionally, publishers may turn to web log analysis to track where users are coming from. (NISO)

For more information on the NISO initiative, please contact [Nettie Lagace](http://www.niso.org/news/pr/contact?item_key=c2ab810f6ee30db6113af9bd6638748d5caf94f7&pr_contact=1) at NISO.

## Key Terms & Definitions

When discussing how to identify origin, it helps to use a common vocabulary.

### Domain

* “A **domain name** is an identification string that defines a realm of administrative autonomy, authority or control within the Internet” (*Wikipedia* <https://en.wikipedia.org/wiki/Domain_name>).
* Examples
	+ EBSCO: <http://eds.b.ebscohost.com/eds/detail/detail?vid=8&sid=cd56ebd1-9ca0-43fb-a239-481b65bff7a8%40sessionmgr112&hid=114&bdata=JnNpdGU9ZWRzLWxpdmU%3d>
	+ IEEE Xplore: [http://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText= IEEE%20Transactions%20on%20Pattern%20Analysis%20and%20Machine%20Intelligence&sid=EBSCO](http://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText=%20IEEE%20Transactions%20on%20Pattern%20Analysis%20and%20Machine%20Intelligence&sid=EBSCO)
	+ Full Text Finder Link Resolver: <http://resolver.ebscohost.com/openurl?EPCustID=ns238282&genre=article&sid=EBSCO&aTitle=12345TITLE&site=ftf-live>
	+ Cranfield University’s SFX Link Resolver: <http://sfx.cranfield.ac.uk/cranfield?genre=article&issn=01628828&title=IEEE%20Transactions%20on%20Pattern%20Analysis%20and%20Machine%20Intelligence&volume=37&issue=6&date=20150601&atitle=Robust%20high%20dynamic%20range%20imaging%20by%20rank%20minimization&spage=1219&sid=EBSCO:edselc&pid>=

### HTTP Referrer

* “The HTTP [referrer] is an HTTP header field that identifies the address of the webpage (i.e. the URI or IRI) that linked to the resource being requested” (*Wikipedia* <https://en.wikipedia.org/wiki/HTTP_referer>).
* When a user links directly from a discovery service to full text on a publisher’s site, the HTTP Referrer will give the discovery service’s domain, which is the user’s point of origin.
	+ Examples, Direct Links:

<http://eds.b.ebscohost.com/eds/detail/detail?vid=8&sid=cd56ebd1-9ca0-43fb-a239-481b65bff7a8%40sessionmgr112&hid=114&bdata=JnNpdGU9ZWRzLWxpdmU%3d>

[http://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText= IEEE%20Transactions%20on%20Pattern%20Analysis%20and%20Machine%20Intelligence&sid=EBSCO](http://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText=%20IEEE%20Transactions%20on%20Pattern%20Analysis%20and%20Machine%20Intelligence&sid=EBSCO)

* When a user links from a discovery service to the full text on a publisher’s site using a link resolver, the HTTP Referrer will give the link resolver’s domain, which is *not* the point of origin.
	+ Examples, Link Resolvers: <http://resolver.ebscohost.com/openurl?EPCustID=ns238282&genre=article&sid=EBSCO&aTitle=12345TITLE&site=ftf-live>

<http://sfx.cranfield.ac.uk/cranfield?genre=article&issn=01628828&title=IEEE%20Transactions%20on%20Pattern%20Analysis%20and%20Machine%20Intelligence&volume=37&issue=6&date=20150601&atitle=Robust%20high%20dynamic%20range%20imaging%20by%20rank%20minimization&spage=1219&sid=EBSCO:edselc&pid>=

### Query String

* “On the World Wide Web, a **query string** is the part of a uniform resource locator (URL) containing data that does not fit conveniently into a hierarchical path structure” (*Wikipedia* <https://en.wikipedia.org/wiki/Query_string>).
* When a user links from a discovery service to the full text on a publisher’s site using a link resolver, the query string will contain the discovery service’s identifier, which indicates the point of origin.

### SID, RFR.ID

* Parameters added to URL query strings to help identify the source of the traffic.
* Can be a static value (i.e., SID=EBSCO) or a dynamic one (i.e., SID={SID}) depending on whether the link is a direct link or going through a link resolver that should pass the origin on.
* Example of SID in direct URL: [http://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText= IEEE%20Transactions%20on%20Pattern%20Analysis%20and%20Machine%20Intelligence&sid=EBSCO](http://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText=%20IEEE%20Transactions%20on%20Pattern%20Analysis%20and%20Machine%20Intelligence&sid=EBSCO)
* Examples of SID in link resolver URL where the origin is the EBSCO platform:

<http://resolver.ebscohost.com/openurl?EPCustID=ns238282&genre=article&sid=EBSCO&aTitle=12345TITLE&site=ftf-live>

<http://sfx.cranfield.ac.uk/cranfield?genre=article&issn=1095323X&title=IEEE%20Aerospace%20Conference%20Proceedings&volume=4&issue=&date=20000101&atitle=In-flight%20performance%20of%20the%20NSTAR%20ion%20propulsion%20system%20on%20the%20deep%20space%20one%20mission&spage=123&sid=EBSCO:edselc&pid=%3Cauthors%3EPolk,%20J.E.;Kakuda,%20R.Y.;Anderson,%20J.R.;Brophy,%20J.R.;Rawlin,%20V.K.;Sovey,%20J.;Hamley,%20J.%3C/authors%3E%3Cui%3Eedselc.2-52.0-0034432352%3C/ui%3E%3Cdate%3E20000101%3C/date%3E%3Cdb%3Eedselc%3C/db%3E>

## Recommendations for Analyzing Incoming Traffic

Each site will have different methods for capturing incoming traffic information and reporting on it, but here are some guidelines for identifying traffic originating from EBSCO.

1. Look for the SID on the **Query String**.
	1. If found, use this value in the origin analysis
	2. If not found, proceed to step 2.
2. Look to see if the **HTTP Referrer** has a domain that matches an EBSCO domain.\*
	1. If so, use “EBSCO” as the origin for origin analysis
	2. If not, proceed to step 3.
3. Check to see if the **HTTP Referrer** is an OpenURL Resolver. Check to see if the HTTP Referrer has “SID=” as part of the URL
	1. If so, use the value in the origin analysis
	2. If not, proceed to step 4
4. Use the domain of the **HTTP Referrer** for the origin analysis

\*Note that step 2 is required because EDS and EBSCOhost URLs include an SID parameter that represents the sessionID and would conflict with the check done in step 3.

If you would like to partner with EBSCO to do more in-depth analysis, please contact your EBSCO Publisher Relations representative.

## EBSCO Domains

EBSCO products use a variety of domain names in our URLs as shown below. For the most up-to-date listing, see the EBSCO Support site <http://support.ebscohost.com/knowledge_base/detail.php?id=966>.

Web domains

* ebscohost.com
* \*.ebscohost.com
* \*.\*.ebscohost.com
* web.a.ebscohost.com
* web.b.ebscohost.com
* web.c.ebscohost.com
* web.d.ebscohost.com
* web.e.ebscohost.com
* \*.ebsco-content.com
* global.ebsco-content.com
* ebscovideos.ebscohost.com
* epnet.com
* \*.ebsco.com

Database or service-specific domains

* EBSCOhost
* [www.ebscohost.com](http://www.ebscohost.com)
* web.ebscohost.com
* search.ebscohost.com
* weblinks1.ebscohost.com - weblinks3.ebscohost.com
* imageserver.ebscohost.com
* imagesrvr.ebscohost.com
* content.ebscohost.com
* exports.ebscohost.com
* plinks.ebscohost.com
* therapy.ebscohost.com

EBSCO Discovery Service, EBSCOhost Integrated Search

* eds.a.ebscohost.com
* eds.b.ebscohost.com
* eds.c.ebscohost.com
* eds.d.ebscohost.com
* eds.e.ebscohost.com