THOMSON REUTERS OPEN CALAIS

UPGRADE GUIDE

THOMSON REUTERS

Contents

Chapter 1   Introduction

Open Calais is a sophisticated Thomson Reuters web service that attaches intelligent metadata-tags to your unstructured content enabling powerful text analytics. The Open Calais natural language processing engine automatically analyzes and tags your input files. It does this in a way that your application can both easily pinpoint relevant data, and effectively leverage the invaluable intelligence contained within the text.

Open Calais analyzes the semantic content of your input files using a combination of statistical, machine-learning, and custom pattern-based methods. Developed by the Text Metadata Services (TMS) group at Thomson Reuters, Open Calais yields highly accurate and detailed metadata.

Open Calais also maps your metadata-tags to Thomson Reuters unique IDs. This supports disambiguation (and linking) of data across all documents processed by Open Calais, and also offers you the opportunity to further enrich your data with related information from the Thomson Reuters datasets.

The new and improved Open Calais is not fully backward compatible with the legacy Open Calais and will require a few manual changes.

The purpose of this guide is to enable a smooth transition from legacy Open Calais to the new and improved Open Calais. **The legacy Open Calais will be discontinued on AUGUST 31, 2015.** Until that time, both environments will be available for your convenience. For more details throughout the transition period, please follow the Open Calais blog.

Chapter 2   Key Changes & Enhancements

Open Calais has undergone major enhancements throughout the last years, and is now faster, more stable, and provides enhanced through-put and metadata capabilities.

The system is available free of charge for up to 5,000 requests a day (at an assumption of average doc size to be 5KB). For information about processing a higher volume of data, please contact us at questions@opencalais.com.

2.1   Metadata Enhancements

The new Open Calais provides the following metadata enhancements: (See also Output Examples.)

- **er/Company** – A disambiguation tag that maps the extracted company to the corresponding company in the Thomson Reuters Organization Authority, a universe of over 3 million organizations (mainly the largest international business organizations worldwide).

  Open Calais Company metadata is substantially enhanced in quality from the previous opencalais.com version. The main enhancements are in the precision of Disambiguation, and Relevance, with more accurate recognition of high relevance mentions.

  The **openpermid** attribute value is a direct link to the relevant company page on the Open PermID website (https://permid.org), and gives you direct access to high quality, curated Thomson Reuters company data.

  Company disambiguation is now available in English, French, and Spanish.

- **em/e/Person**: This extracted metadata tag now includes firstname and lastname attributes; provides the additional person types: attorney, judge.

- **er/Person**: A disambiguation tag that maps the extracted person to the corresponding person in the Thomson Reuters People Authority, in the Officers and Directors dataset, which is a universe of ~1.6M news worthy people. This capability is now available for English, French, and Spanish.

- **er/Geo**: A disambiguation tag that maps an extracted City, Country, or ProvinceOrState to the corresponding geographical entity in the Thomson Reuters Metadata Authority. The geographical entity is identified by name and a Thomson Reuters PermID. This capability is now available for English, French, and Spanish.

- **er/Organization**: A disambiguation tag that maps an extracted Organization to the corresponding organization in the Thomson Reuters Metadata Authority. The Organization is identified by name and the Thomson Reuters PermID. This capability is now available in English, French and Spanish.

- **er/Currency**: A disambiguation tag that maps an extracted Currency to a currency in the Thomson Reuters Metadata Authority. The Currency is identified by name and the Thomson Reuters PermID. This capability is now available for English, French and Spanish.

- **er/Continent**: A disambiguation tag that maps an extracted Continent to a continent in the Thomson Reuters Metadata Authority. The Continent is identified by name and the Thomson Reuters PermID. This capability is now available for English, French and Spanish.

- **er/Region**: A disambiguation tag that maps an extracted Region to the corresponding region in the Thomson Reuters Metadata Authority. The Region is identified by name and the Thomson Reuters PermID. This capability is now available for English, French and Spanish.

- **forEndUserDisplay**: An attribute available for all the Open Calais metadata tags. This attribute is a recommendation of whether the tag is suitable as a search item for a specific document or whether the metadata is primarily of use for aggregation and analytics on large quantities of documents.

2.2   Premium Metadata

The following additional metadata types are available upon subscription to Thomson Reuters Intelligent Tagging (TRIT): (See also Open Calais Packages, and Output Examples.)

- **TopmostPublicParentCompany** (er/TopmostPublicParentCompany): This tag is generated when an extracted entity of the type Company is mapped to a company in the Thomson Reuters dataset that also has a publicly traded parent company. For example, a mention of IBM Japan would also yield an association of the document with IBM Corp.

  The **openpermid** attribute value is a direct link to the relevant company page on the Open PermID website (https://permid.org), and gives you direct access to high quality, curated Thomson Reuters company data.

- **PharmaceuticalDrug** (em/e/PharmaceuticalDrug): This tag is generated when Open Calais identifies a mention of a pharmaceutical drug.

- **Deal (em/r/Deal, er/Deal)**: These tags are generated when a deal is identified by Open Calais and mapped to a deal in the Thomson Reuters Deal Authority. The Thomson Reuters Deal Authority is a repository of hundreds of thousands of corporate Mergers and Acquisitions.

- **Industry Tags**: Industry tags identify the industries related to the companies mentioned in the text. The industry is identified by name, the Thomson Reuters classification ID (RCS Code), and the Thomson Reuters PermID.

- **Category Tags** (DocCat): Calais identifies the topic or topics that are being discussed in the document. The list of possible topics is defined by the International Press Telecommunications council (IPTC) news taxonomy and is also defined by the Thomson Reuters Coding Schema (TRCS) for subscribers to Thomson Reuters Intelligent Tagging. TRCS is an extensive Thomson Reuters News Taxonomy, covering more than 1000 different News Topics . The Category is identified by name, Thomson Reuters classification ID (RCS Code), and the Thomson Reuters PermID.

- **er/Geo**: A disambiguation tag that maps an extracted City, Country, or ProvinceOrState to the corresponding geographical entity in the Thomson Reuters Metadata Authority. The geographical entity is identified by name, the Thomson Reuters classification ID (RCS Code), and a Thomson Reuters PermID. This capability is now available for English, French, and Spanish.

- **er/Organization**: A disambiguation tag that maps an extracted Organization to the corresponding organization in the Thomson Reuters Metadata Authority. The Organization is identified by name, the Thomson Reuters classification ID (RCS Code), and the Thomson Reuters PermID. This capability is now available in English, French and Spanish.

- **er/Currency**: A disambiguation tag that maps an extracted Currency to a currency in the Thomson Reuters Metadata Authority. The Currency is identified by name, the Thomson Reuters classification ID (RCS Code), and the Thomson Reuters PermID. This capability is now available for English, French and Spanish.

- **er/Continent**: A disambiguation tag that maps an extracted Continent to a continent in the Thomson Reuters Metadata Authority. The Continent is identified by name, the Thomson Reuters classification ID (RCS Code), and the Thomson Reuters PermID. This capability is now available for English, French and Spanish.

- **er/Region**: A disambiguation tag that maps an extracted Region to the corresponding region in the Thomson Reuters Metadata Authority. The Region is identified by name, the Thomson Reuters classification ID (RCS Code), and the Thomson Reuters PermID. This capability is now available for English, French and Spanish.

2.3　Discontinued and Changed Functionality

Functionality that is no longer available:

- GenericRelations: Please contact us for further details about how to use other features to accomplish similar goals.

- Simple Output Format: The simple output format cannot fully support the full richness of Open Calais metadata, therefore we have deprecated this format.

- SOAP interface is no longer supported. Please use our REST API.

- er/Product: Extracted Product entities are no longer mapped to CNET URL.

- Calais LinkedData Repository: This function was deprecated to make room for a new and enhanced form of Linked Data, exposing Thomson Reuters entity masters via https://permid.org. Calais now links to entities available at https://permid.org.

- Please see the section on Request Headers. Note that headers other than the ones listed in this section are not supported by Open Calais.

- The links to company data that appeared in the legacy Open Calais output are no longer supported.

  For the convenience of legacy Open Calais users, a new attribute, **legacyid**, has been temporarily added to the er/company and er/TopmostPublicParentCompany disambiguation tags. The purpose of this attribute is to enable you to map the old company links, which are no longer supported, to the corresponding new links. Please see the section Company Links for further details.

- The RelevanceInfo tag generated for all extracted entities now provides discrete scores instead of continuous scores, (except for the RelevanceInfo tag generated for company entities, which includes both the **relevance** and **relevancecont** attributes). For more details, see Relevance Scores.

2.3.1 Company Links

The old links to company data that were included in the legacy Open Calais output are no longer supported.

For the convenience of legacy Open Calais users, a new attribute, **legacyid**, has been temporarily added to the er/company and er/TopmostPublicParentCompany disambiguation tags.

The purpose of the **legacyid** attribute is to enable you to map the old links, which are no longer supported, to the corresponding new links.

This attribute will be available until August 31, 2015.

The attribute value is the link to the company as it would have been displayed in the legacy Open Calais output.

An example of an old company link in the legacy Open Calais er/company tag (RDF output):

```
<rdf:Description rdf:about="http://d.opencalais.com/er/company/ralg-tr1r/39d9a86f-6241-3688-b05b-4595eeaba2d1">
```

The company link and the legacyid attribute as they appear in the new Open Calais er/company tag (RDF output):

```
<rdf:Description rdf:about="http://d.opencalais.com/er/company/ralg-oa/5000608903">
        <rdf:type rdf:resource="http://s.opencalais.com/1/type/er/Company"/>
        <c:docId rdf:resource="http://d.opencalais.com/dochash-1/7586b818-40af-3d55-ac16-bf3520cddfa6"/>
        <c:permid>4295905573</c:permid>
        <c:score>0.9928677</c:score>
        <c:legacyid rdf:resource="http://d.opencalais.com/er/company/ralg-tr1r/39d9a86f-6241-3688-b05b-4595eeaba2d1"/>
        <!--Apple-->
        <c:subject rdf:resource="http://d.opencalais.com/comphash-1/705cd5cf-93e1-323c-8d4e-1ea3200d37e4"/>
        <c:name>Apple Inc</c:name>
        <c:commonname>Apple</c:commonname>
        <c:ticker>AAPL</c:ticker>
        <c:primaryric>AAPL.OQ</c:primaryric>
        <c:openpermid rdf:resource="https://permid.org/1-4295905573"/>
</rdf:Description>
```

The company link and the legacyid attribute as they appear in the new Open Calais er/company tag (N3 output):

```
<http://d.opencalais.com/er/company/ralg-oa/5000608903>
        a               <http://s.opencalais.com/1/type/er/Company> ;
        c:commonname    "Apple" ;
        c:docId         <http://d.opencalais.com/dochash-1/268b4e49-098d-389d-9d76-816cb722a630> ;
        c:legacyid      < http://d.opencalais.com/er/company/ralg-tr1r/39d9a86f-6241-3688-b05b-4595eeaba2d1> ;
        c:name          "Apple Inc" ;
        c:openpermid    <https://permid.org/1-4295905573> ;
        c:permid        "4295905573" ;
        c:primaryric    "AAPL.OQ" ;
        c:score         "0.9928677" ;
        c:subject       <http://d.opencalais.com/comphash-1/705cd5cf-93e1-323c-8d4e-1ea3200d37e4> ;
        c:ticker        "AAPL" .
```

Note that this attribute is not relevant to the JSON output format.

2.3.2   Relevance Scores

Continuous relevance scores, which were supported by legacy Open Calais have been slightly modified; in the new Open Calais, relevance scores are expressed as discrete values (except for the company entity which is given both discrete and continuous relevance scores).

All extracted entities have an associated *RelevanceInfo* tag that indicates how centric the entity is to the containing document. The tag includes the *relevance* attribute, which defines a discrete relevance score.

A *RelevanceInfo* tag for an extracted person:

```
<rdf:Description rdf:about="http://d.opencalais.com/dochash-1/f4707556-c36e-39af-b0e6-0103f889be3e/Relevance/9">
        <rdf:type rdf:resource="http://s.opencalais.com/1/type/sys/RelevanceInfo"/>
        <c:docId rdf:resource="http://d.opencalais.com/dochash-1/f4707556-c36e-39af-b0e6-0103f889be3e"/>
        <c:subject rdf:resource="http://d.opencalais.com/pershash-1/e4808181-2cd0-3670-b992-7467229ba691"/>
        <c:relevance>0.8</c:relevance>
    </rdf:Description>
```

Currently, the following *relevance* attribute values (discrete relevance scores) are supported:

- 1.0 – This value is reserved for the company identified as the reporting company in a document with a predefined format such as an SEC report.
- 0.8 – Entities defined as having high relevance will receive this score.
- 0.5 – Entities defined as having medium relevance will receive this score.
- 0.2 – Entities defined as having low relevance will receive this relevance score.
- 0.0 – This value is reserved for entities identified as irrelevant to the story. For example, mentions of companies as rating agencies, reporting agencies, stock exchanges, and social applications receive a relevance score of 0.

The *RelevanceInfo* tag for the company entity also includes the *relevancecont* attribute (a continuous relevance score).

For more details, see the Thomson Reuters Open Calais API User Guide.

2.4   Peripheral Tools

- Calais Marmoset - http://www.opencalais.com/documentation/calais-marmoset is no longer supported.

- Calais Tagaroo - http://tagaroo.opencalais.com/ will undergo modifications to support the new API and will be published once we complete the transition.

- Submission Tool - http://www.opencalais.com/documentation/calais-submission-tool, the sample application for submitting documents to the old API will no longer be supported. We will publish a new sample application for submitting documents to the new API.

- Calais SemanticProxy - http://www.semanticproxy.com/ is no longer supported.

2.5 Error Codes

Note that the error codes in the new and updated Open Calais differ from the error codes in legacy Open Calais. Please refer to the Open Calais API User Guide for further details.

Chapter 3   Output Examples (Legacy vs. New Open Calais)

This section highlights the differences between legacy and new Open Calais metadata output.

Metadata Tag Type	Legacy Open Calais Output	New Open Calais Output
er/Company	```<rdf:Description
rdf:about="http://d.opencalais.c
om/er/company/ralg-
tr1r/23d07771-c50b-315b-8050-
3cdaf47ac0d0">
<rdf:type
rdf:resource="http://s.opencalai
s.com/1/type/er/Company"/>
<c:docId
rdf:resource="http://d.opencalai
s.com/dochash-1/bdc76fb2-e65f-
320c-9447-fae386bcda44"/>
<!--Apple Inc-->
<c:subject
rdf:resource="http://d.opencalai
s.com/comphash-1/d412df50-1061-
3276-a052-5bbd85244633"/>
<c:score>1.0</c:score>
<c:name>APPLE INC.</c:name>
<c:shortname>Apple</c:shortname>
<c:ticker>AAPL</c:ticker>
<c:symbol>AAPL.OQ</c:symbol>
</rdf:Description>```	```<rdf:Description
rdf:about
http://d.opencalais.com/er/company/r
alg-oa/4295905573">
<rdf:type
rdf:resource="http://s.opencalais.co
m/1/type/er/Company"/>
<c:docId
rdf:resource="http://d.opencalais.co
m/dochash-1/bdc76fb2-e65f-320c-9447-
fae386bcda44"/>
<c:score>1.0</c:score>
<c:legacyid rdf:resource="
http://d.opencalais.com/er/company/r
alg-tr1r/23d07771-c50b-315b-8050-
3cdaf47ac0d0"/>
<!--Apple Inc-->
<c:subject
rdf:resource="http://d.opencalais.co
m/comphash-1/d412df50-1061-3276-
a052-5bbd85244633"/>
<c:name>Apple Inc</c:name>
<c:commonname>Apple</c:commonname>
<c:ticker>AAPL</c:ticker>
<c:primaryric>AAPL.OQ</c:primaryric>
<c:permid>4295905573</c:permid>
<c:openpermid
rdf:resource="https://permid.org/1-
4295905573"/>
</rdf:Description>```

Metadata Tag Type	Legacy Open Calais Output	New Open Calais Output
em/e/Person	```<rdf:Description
rdf:about="http://d.opencalais.c
om/pershash-1/054d1351-39a1-
3146-887a-d453c24cebba">
<rdf:type
rdf:resource="http://s.opencalai
s.com/1/type/em/e/Person"/>
<c:name>Bill Gates</c:name>
<c:persontype>economic</c:persontype
>
<c:nationality>N/A</c:nationality>
<c:commonname>Bill
Gates</c:commonname>
</rdf:Description>```	```<rdf:Description
rdf:about="http://d.opencalais.com/p
ershash-1/054d1351-39a1-3146-887a-
d453c24cebba">
<rdf:type
rdf:resource="http://s.opencalais.co
m/1/type/em/e/Person"/>
<c:name>Bill Gates</c:name>
<c:persontype>economic</c:persontype>
<c:nationality>N/A</c:nationality>
<c:confidencelevel>0.987</c:confidencel
evel>
<c:firstname>Bill</c:firstname>
<c:lastname>Gates</c:lastname>
<c:commonname>Bill Gates</c:commonname>
</rdf:Description>```
er/Person	N/A	```<rdf:Description
rdf:about="http://d.opencalais.com/e
r/person/ralg-pa1/34413157709">
<rdf:type
rdf:resource="http://s.opencalais.co
m/1/type/er/Person"/>
<c:docId
rdf:resource="http://d.opencalais.co
m/dochash-1/b1d5bde9-3d9c-32c7-96e4-
204880ccceac"/>
<c:name>William Henry Gates</c:name>
<c:personid>1975266</c:personid>
<c:paid>34413157709</c:paid>
<c:officerid>28066</c:officerid>
<c:commonname>Bill Gates</c:commonname>
<c:score>0.8797866</c:score>
<!--Bill Gates-->
<c:subject
rdf:resource="http://d.opencalais.co
m/pershash-1/054d1351-39a1-3146-
887a-d453c24cebba"/>
</rdf:Description>```

Metadata Tag Type	Legacy Open Calais Output	New Open Calais Output

er/Geo
 	
This example is specific to er/Geo/ProvinceOrState, but the RDF output will be the same for er/Geo/City and er/Geo/Country as well.
	``xml
	<rdf:Description
	        rdf:about="http://d.opencalais.c
	        om/er/geo/provinceorstate/ralg-
	        geo1/62914270-0a1c-1bcc-880e-
	        6dbe9fb82eb9">
	    <rdf:type
	        rdf:resource="http://s.opencalai
	        s.com/1/type/er/Geo/ProvinceOrSt
	        ate"/>
	    <c:docId
	        rdf:resource="http://d.opencalai
	        s.com/dochash-1/bdc76fb2-e65f-
	        320c-9447-fae386bcda44"/>
	    <!--Texas-->
	    <c:subject
	        rdf:resource="http://d.opencalai
	        s.com/genericHasher-1/b6607136-
	        0945-32d7-b4fa-a99019467fd9"/>
	    <c:name>Texas, United States</c:name>
	    <c:shortname>Texas</c:shortname>
	    <c:containedbycountry>United
	        States</c:containedbycountry>
	    <c:latitude>31.4484328889</c:latitud
	        e>
	    <c:longitude>-
	        97.7816569778</c:longitude>
	</rdf:Description>
	``
		``xml
		<rdf:Description
		        rdf:about="http://d.opencalais.com/e
		        r/geo/provinceorstate/ralg-
		        geo1/62914270-0a1c-1bcc-880e-
		        6dbe9fb82eb9">
		    <rdf:type
		        rdf:resource="http://s.opencalais.co
		        m/1/type/er/Geo/ProvinceOrState"/>
		    <c:docId
		        rdf:resource="http://d.opencalais.co
		        m/dochash-1/bdc76fb2-e65f-320c-9447-
		        fae386bcda44"/>
		    <c:name>Texas, United States</c:name>
		    <c:shortname>Texas</c:shortname>
		    <c:latitude>31.4484328889</c:latitude>
		    <c:longitude>-
		        97.7816569778</c:longitude>
		    <c:containedbycountry>United
		        States</c:containedbycountry>
		    <c:rcscode>G:22</c:rcscode>
		    <c:permid>100023</permid>
		    <!--Texas-->
		    <c:subject
		        rdf:resource="http://d.opencalais.co
		        m/genericHasher-1/b6607136-0945-
		        32d7-b4fa-a99019467fd9"/>
		</rdf:Description>
		``

er/Organization	N/A	```xml
		<rdf:Description
		        rdf:about="http://d.opencalais.com/e
		        r/organization/comphash-1/2c96fe9b-
		        88c8-3059-8aed-4e8824835b5c">
		    <rdf:type
		        rdf:resource="http://s.opencalais.co
		        m/1/type/er/Organization"/>
		    <c:docId
		        rdf:resource="http://d.opencalais.co
		        m/dochash-1/672e4d50-39db-3fbb-ab51-
		        53c6c10e9cb2"/>
		    <c:name>United Nations</c:name>
		    <!--United Nations-->
		    <c:subject
		        rdf:resource="http://d.opencalais.co
		        m/genericHasher-1/fa36ccaa-3402-
		        3be8-993a-00a6a8fbfb2e"/>
		    <c:rcscode>M:8N</c:rcscode>
		    <c:permid>8589934571</c:permid>
		</rdf:Description>
		``

Metadata Tag Type	Legacy Open Calais Output	New Open Calais Output
er/Currency	N/A	```<rdf:Description rdf:about="http://d.opencalais.com/er/currency/comphash-1/4f50c5c9-ae72-353c-86a2-020cabca89f5"> <rdf:type rdf:resource="http://s.opencalais.com/1/type/er/Currency"/> <c:docId rdf:resource="http://d.opencalais.com/dochash-1/7399183e-c4c7-31cf-b528-06de23b1acb9"/> <c:rcscode>C:6</c:rcscode> <c:name>US Dollar</c:name> <!--USD--> <c:subject rdf:resource="http://d.opencalais.com/genericHasher-1/736a403d-157b-3e80-86a7-acc404607cb2"/> </rdf:Description>```
er/Continent	N/A	```<rdf:Description rdf:about="http://d.opencalais.com/er/continent/comphash-1/04683db9-5680-3561-bc 70-87e8d62924a0"> <rdf:type rdf:resource="http://s.opencalais.com/1/type/er/Continent"/> <c:docId rdf:resource="http://d.opencalais.com/dochash-1/369104f6-1d29-3d5f-a6e1-63235ab3bd73"/> <c:name>Europe</c:name> <!-- Europe --> <c:subject rdf:resource="http://d.opencalais.com/genericHasher-1/09bebb87-1ef5-373d-bd01-263e77470c65"/> <c:rcscode>G:A</c:rcscode> </rdf:Description>```
er/Region	N/A	```<rdf:Description rdf:about="http://d.opencalais.com/er/region/comphash-1/b5edcd0b-ff9f-396d-87ae-ebc5ddb17242"> <rdf:type rdf:resource="http://s.opencalais.com/1/type/er/Region"/> <c:docId rdf:resource="http://d.opencalais.com/dochash-1/7399183e-c4c7-31cf-b528-06de23b1acb9"/> <c:rcscode>G:Q</c:rcscode> <c:name>Middle East</c:name> <c:permid>103394</c:permid> <!--Middle East--> <c:subject rdf:resource="http://d.opencalais.com/genericHasher-1/55ccb46b-68ce-3750-a326-6bdd742672fc"/> </rdf:Description```

Metadata Tag Type	Legacy Open Calais Output	New Open Calais Output
DocCat (RCS topic*) (A DocCat tag assigns a topic from either the TRCS or IPTC taxonomy. In this example, the topic is from TRCS.)	N/A	```<rdf:Description` ` rdf:about="http://d.opencalais.com/d` ` ochash-1/bdc76fb2-e65f-320c-9447-` ` fae386bcda44/cat/5">` ` <rdf:type` ` rdf:resource="http://s.opencalais.co` ` m/1/type/cat/DocCat"/>` ` <c:docId` ` rdf:resource="http://d.opencalais.co` ` m/dochash-1/bdc76fb2-e65f-320c-9447-` ` fae386bcda44"/>` ` <c:forenduserdisplay>false</c:forenduse` ` rdisplay>` ` <c:rcscode>F:22</c:rcscode>` ` <c:name>Judicial Process / Court Cases` ` / Court Decisions</c:name>` ` <c:score>0.884</c:score>` `</rdf:Description>```
Industry Tag*	N/A	```<rdf:Description` ` rdf:about="http://d.opencalais.com/d` ` ochash-1/bdc76fb2-e65f-320c-9447-` ` fae386bcda44/Industry/1">` ` <rdf:type` ` rdf:resource="http://s.opencalais.co` ` m/1/type/tag/Industry"/>` ` <c:docId` ` rdf:resource="http://d.opencalais.co` ` m/dochash-1/bdc76fb2-e65f-320c-9447-` ` fae386bcda44"/>` ` <c:name>Department Stores -` ` NEC</c:name>` ` <c:rcscode>B:1464</c:rcscode>` ` <c:permid>4294951537</c:permid>` ` <c:relevance>0.200</c:relevance>` `</rdf:Description>```

Metadata Tag Type	Legacy Open Calais Output	New Open Calais Output

er/TopmostPublicParentCompany*	N/A	`<rdf:Description rdf:about="http://d.opencalais.com/er/topmostpublicparentcompany/ralg-oa-tm/4295912712">`
		`<rdf:type rdf:resource="http://s.opencalais.com/1/type/er/TopmostPublicParentCompany"/>`
		`<c:docId rdf:resource="http://d.opencalais.com/dochash-1/6a4c07a9-07e8-3e32-9e63-99e4998a5371"/>`
		`<c:permid>4295912712</c:permid>`
		`<!--Broadway Federal Bank-->`
		`<c:subject rdf:resource="http://d.opencalais.com/comphash-1/9d648832-7a10-3047-a7de-e9d4a5fa4b6e"/>`
		`<c:legacyid rdf:resource="http://d.opencalais.com/er/company/ralg-tr1r/a2ecacae-d62b-3eaa-83d4-cb7d091d512c"/>`
		`<c:name>Broadway Financial Corp</c:name>`
		`<c:commonname>Broadway Finl</c:commonname>`
		`<c:ticker>BYFC</c:ticker>`
		`<c:primaryric>BYFC.OQ</c:primaryric>`
		`<c:openpermid rdf:resource="https://permid.org/1-4295912712"/>`
		`</rdf:Description>`

Metadata Tag Type	Legacy Open Calais Output	New Open Calais Output

em/r/Deal*	N/A	```<rdf:Description
		        rdf:about="http://d.opencalais.com/c
		        omphash-1/70273bf6-b72c-3d26-8867-
		        ae369c35b106">
		    <rdf:type
		        rdf:resource="http://s.opencalais.co
		        m/1/type/em/r/Deal"/>
		    <c:forenduserdisplay>true</c:forenduser
		        display>
		    <c:confidencelevel>0.999893</c:confiden
		        celevel>
		    <c:target
		        rdf:resource="http://d.opencalais.co
		        m/comphash-1/649b219c-5892-30db-
		        83e6-71a0dd62f51f"/>
		    <c:acquirer
		        rdf:resource="http://d.opencalais.co
		        m/comphash-1/1a07a87e-1f56-31bc-
		        ae94-ca8d09b64760"/>
		</rdf:Description>```

er/Deal*	N/A	```<rdf:Description
		        rdf:about="http://d.opencalais.com/e
		        r/deal/154082536283">
		    <rdf:type
		        rdf:resource="http://s.opencalais.co
		        m/1/type/er/Deal"/>
		    <c:docId
		        rdf:resource="http://d.opencalais.co
		        m/dochash-1/4faf0326-3093-31b8-8758-
		        551dfb1974fa"/>
		    <c:subject
		        rdf:resource="http://d.opencalais.co
		        m/comphash-1/70273bf6-b72c-3d26-
		        8867-ae369c35b106"/>
		    <c:permid>154082536283<c:/permid>
		    <c:dealsdcnum>2717301020</dealsdcnum>
		</rdf:Description>```

em/e/PharmaceuticalDrug*	N/A	```<rdf:Description
		        rdf:about="http://d.opencalais.com/d
		        rughash-1/efda702c-5c5b-36eb-9233-
		        d5badc2eb92f">
		    <rdf:type
		        rdf:resource="http://s.opencalais.co
		        m/1/type/em/e/PharmaceuticalDrug"/>
		    <c:forenduserdisplay>true</c:forenduser
		        display>
		    <c:name>Avandia</c:name>
		    <c:confidencelevel>0.999</c:confidencel
		        evel>
		</rdf:Description>```

Metadata Tag Type	Legacy Open Calais Output	New Open Calais Output
em/r/GenericRelations	```<rdf:Description` ` rdf:about="http://d.opencalais.c` ` om/genericHasher-1/9a0bf02d-` ` 985b-309e-a583-329c2648d148">` ` <rdf:type` ` rdf:resource="http://s.opencalai` ` s.com/1/type/em/r/GenericRelatio` ` ns"/>` ` <!--Apple Inc-->` ` <c:relationsubject` ` rdf:resource="http://d.opencalai` ` s.com/comphash-1/d412df50-1061-` ` 3276-a052-5bbd85244633"/>` ` <c:verb>claim</c:verb>` `</rdf:Description>```	N/A (This feature is deprecated.)

*These metadata items are available to subscribers to Thomson Reuters Intelligent Tagging (TRIT). For further information please contact our customer support team at questions@opencalais.com.

Chapter 4   Upgrading to the New Open Calais

4.1   The Workflow

We recommend you make the transition from Legacy Open Calais to Open Calais in two steps:

1. Adapt your application to the new API without adding new functionality.

2. Add new functionality gradually at your own pace once you are adapted to the new API.

You can preview the functionality of the entire inventory of Open Calais semantic metadata tags (including the tags available to Thomson Reuters Intelligent Tagging users) with the Open Calais Demo tool, at http://www.opencalais.com/ Demo page.

4.2 Forming the API Call

Here we present the basic information about forming the API call. For more detailed information, please refer to the Open Calais API User Guide.

4.2.1 Resource URLs

HTTP REST

POST https://api.thomsonreuters.com/permid/calais

Note: **We no longer support multiple end points - a REST end point is the only available end point.**

4.2.2 Request Authentication

In order to post a request, you are required to register and obtain a license key at www.opencalais.com.

Note: **Your previous key will not work for you in the new environment – you will need to register for a new one.**

If you are already registered for Open PermID services (https://permid.org)– the same key and credentials can be used for Open Calais. If you haven't yet seen our Open PermID initiative at https://permid.org/, please check it out!

4.2.3   Request Headers

The request is an HTTPS POST with the following query parameters:

Parameter	Description
Content-Type (Mandatory)	Indicates the input content type (mime type). Values: • text/xml • text/html • text/raw • application/pdf – available to Premium (Thomson Reuters Intelligent Tagging) users text/htmlraw format is no longer supported
x-ag-access-token (Mandatory)	The value of this header is your access key. (You receive an access key upon registration.) Note that the (legacy) x-calais-licenseID header in no longer supported, and that the old licenseID value is invalid.
outputFormat	Defines the output response format (mime type). Mandatory for any output other than RDF. Values: • xml/rdf (default value) • text/n3 • application/json text/simple and text/microformat formats are no longer supported. Note that the new Open Calais does not support the Accept header.
x-calais-language	Specifies the language of the input text. You can use this header to override the automatic language detection functionality. Values: • English • French • Spanish

Note: Headers other than the ones listed in the table above are not supported by Open Calais.

4.3   The API Response

Upon successful processing, a metadata-enriched document is returned in the format you specified via the outputFormat header.

Upon error, a JSON is returned with full details of the error.

Parameter	Description
result	This element contains the tagged document. Its format is determined by the request parameter **outputFormat**.
status	A numeric return code for the request. The value corresponds to the HTTP return codes. For example, "200" indicates that the request was successfully processed. "400" is the code for a bad request.
errorCode	Present when an error was encountered when processing the request (i.e., status!= 200). It is a short name for the error type.
errorDescription	Present when an error was encountered when processing the request (i.e., status!= 200). It is a containing a descriptive cause of the error encountered.

4.4   A Sample Input File

```
HTTPS REST
POST https://api.thomsonreuters.com/permid/calais
```

4.4.1   Parameters

```
Content-Type: text/xml
outputFormat:xml/rdf
X-AG-ACCESS-TOKEN: YOUR_LICENSE_KEY_HERE
```

4.4.2   Body

```
<Document>
    <Title>GoPro launches $800 million offering, CEO to sell some shares</Title>
    <Body>
        Wearable camera maker GoPro Inc's chief executive, Nicholas Woodman, plans to sell a portion of his
        stake as part of an $800 million offering of the company's shares. The offering of Class A common shares
        is expected to start in the next couple of weeks and close by November, a company spokesman told
        Reuters. GoPro's shares fell as much as 5.1 percent before easing back a little to trade down 3.8
        percent at $76.04 on Monday. The lock-up period on the stock, which listed in June, expires on Dec. 22,
        allowing employees and early investors to sell shares of the company. Typically, on the day a lock-up
        expires, prices tend to fall as a large number of shares become available for trading. The company said
        the offering was expected to soften the blow of the lock-up expiration on the share price.  (Additional
        reporting by Arathy Nair; Editing by Saumyadeb Chakrabarty)
    </Body>
</Document>
```

For a detailed explanation of Open Calais output, see the Open Calais API User Guide.

4.5　　Upgrade Check List

Step	Description
Register for the new Open Calais, retrieve a valid authentication token	@ *http://www.opencalais.com/*
Review the Open Calais API User Guide	@ *http://www.opencalais.com/ API page*
Review enhancements and changes made to list of concepts, attributes, and features.	Review the Upgrade Guide @ *http://www.opencalais.com/ Upgrade page* for highlights and the Open Calais API User Guide for details. Determine whether changes made to metadata concepts are relevant to you and modify/enhance your client accordingly.
REST	Calais will no longer support SOAP; if relevant modify the client to support REST.
Test text submission with the demo tool to review the output RDF.	@ *http://www.opencalais.com/ Demo page*
Check the changed error message codes and make any necessary modifications.	Read the section about error messages in the Open Calais API User Guide @ *http://www.opencalais.com/ API page*
Test the API	@ *http://www.opencalais.com/*
Contact us with any questions or feedback	@ *http://www.opencalais.com/*

Chapter 5   Open Calais Packages

The following table outlines the features of the different Open Calais packages: Free, Basic, and Thomson Reuters Intelligent Tagging (TRIT):

Feature	Free	Basic	TRIT
Submissions Per Day*	Up to 5000	Up to 1 Million**	Up to 1 Million**
Maximum File Size	Up to 100KB	Up to 100KB	Up to 500KB**
Rate Limit (documents per second)	0.5 - 1**	2 – 20**	2 – 32**
Extensive Set of Semantic Entity Tags	√	√	√
Fact and Event Detections	√	√	√
IPTC Categories (DocCat tag)	√	√	√
TRCS Categories (DocCat tag)			√
Financial Deals (em/r/Deal tag, er/Deal tag)			√
Pharmaceutical Drugs (em/e/PharmaceuticalDrug tag)			√
Company Hierarchy (er/TopmostPublicParentCompany tag)			√
TRBC Industry Mapping (Industry tag)			√
PDF Support			√
Pre-Load			√
Third Party NLP Onboarding Mechanism			√
24 x 7 Professional Customer Support		√	√

*A submission is a request of up to the maximum file size (including the body and xml tags, but excluding the HTTP headers).

**According to your license agreement.

For further information, please contact us at questions@opencalais.com.

Existing Premium Customers, please contact us at questions@opencalais.com. We will be happy to assign a technical account manager to assist you with finding the best methods and tools for incorporating Open Calais metadata into your infrastructure and application.

THOMSON REUTERS