



EBI Europe PMC SOAP Web Service

6.8.0

Reference Guide

EMBL-EBI
Wellcome Trust Genome Campus
Hinxton
Cambridge
CB10 1SD
UK
www.ebi.ac.uk

<http://www.ebi.ac.uk/support>

^{11h} May 2022, Doc Version 1.50

EBI Europe PMC SOAP Web Service 6.8.0 Reference Guide

Contents

1. Introduction	6
2. Method Summary	7
3. Data Sources	10
4. Sorting Results	11
5. Method: profilePublications	12
6. Method: searchPublications	14
7. Method: getDatabaseLinks	26
8. Method: getReferences	29
9. Method: getCitations	32
10. Method: getDataLinks	34
11. Method: getLabsLinks	39
12. Method: getFullTextXML	45
13. Method: getBookXML	47
14. Method: getSupplementaryFiles	49
15. Method: listSearchFields	51
Appendix 1: Indexed Fields	54

Document Revision History

Date	Version	Details
11 th May 2022	1.50	Changed webservice to version 6.8.0 Added new REST API endpoint to search publications with a specific source and its id. Included a new SOAP feature to validate empty queryString in search and so as to return a queryException.
29 th Nov 2021	1.49	Changed webservice to version 6.7.0
03 th Aug 2021	1.48	Changed webservice to version 6.6.0 Added nextPageUrl element after nextCursorMark for both core and lite responses. Added support for Accept header. Included pubType information inside the version element of the core response.
26 th Oct 2020	1.47	Changed webservice to version 6.5.0 Added affiliationOrgId for each author affiliation instead of a single affiliationOrgId for each author
14 th Jul 2020	1.46	Changed webservice to version 6.4.0 Remove full text licence statement in search index Added 'fullTextId' in core response, inside the fullTextIdList field
27 th May 2020	1.45	Changed webservice to version 6.3.0 Added datalinks tags Added multiple author affiliations instead of single affiliations Removed text mined terms
25 th Apr 2019	1.44	Changed webservice to version 6.2.0 Added new index fields ORG_ID and FT_CDATE Added fields fullTextReceivedDate and affiliationOrgId
07 th Oct 2018	1.43	Changed webservice to version 6.1.0 Added new index fields ANNOTATION_TYPE and ANNOTATION_PROVIDER Added fields firstIndexDate, versionNumber and versionList
28 th Mar 2018	1.42	Changed webservice to version 6.0.0 Added filed to datalinks method: NameLong, ImageUrl and CollectionUrl
16 th Jan 2018	1.41	Changed webservice to version 5.3.2 Added link count fields for sections and categories of the datalinks method.
11 th Dec 2017	1.40	Changed web service to version 5.3.0. Introduced the getDataLinks method . Added the hasData field to the searchPublications core response Reformatted document revision history and changed to anti chronological order. Reformatted searchPublications method output table .
23 th May 2017	1.39	Changed the description of the sorting options of the Sorting Results chapter.

Date	Version	Details
10 th May 2017	1.38	Changed web service version from 5.1 to 5.2. Added the new parameter “inlineImages” to the getSupplementaryFiles module. Changed the explanation of the parameter “synonym” in the profilePublications and searchPublications method. The response element “firstPublicationDate” of the searchPublications module is now part of the “lite” response as well.
14 th Mar 2017	1.37	Changed web service version from 5.0 to 5.1. Added the new <authMan>, <nihAuthMan> and <manuscriptId> elements to the core output of the searchPublications module. Also removed the lite output marker of the <epmcAuthMan> element.
17 th Feb 2017	1.36	Changed web service version from 4.5.4 to 5.0. Repository migrated from Oracle/Solr to MongoDB. Promoted bookOrReportDetails to the lite response in the searchPublications method.
2 th Feb 2017	1.35	Changed web service version from 4.5.3 to 4.5.4. Added the new supported data base PRIDE in Method Summary, getDatabaseLinks and Info1to4
13 th Oct 2016	1.34	Changed web service version from 4.5.2 to 4.5.3. Removed the request parameter “offset” from the searchPublications method and introduce a new parameter “cursorMark”. Added a new request parameter “sort” for the searchPublications method. Changed Sorting Results regarding to the new parameter “sort”.
1 th July 2016	1.33	Changed web service version from 4.5.1 to 4.5.2. Changed getLabsLinks Output according to the new added imgUrl response data. Adjusting of the searchPublications Method Output according to the pubYear changes.
12 th May 2016	1.32	Changed web service version from 4.5.0 to 4.5.1. Added the new EMBARGOED_MAN search field to the listSearchFields method and Appendix 1. Changed the <DOI> element of the output of the searchPublications module to lower case <doi>.
27 th Jan 2016	1.31	Changed web service version from 4.4 to 4.5. Added the new <hasSuppl>, <license> and <epmcAuthMan> elements to the output of the searchPublications module. Added the new HAS_SUPPL search field to the listSearchFields method and Appendix 1. Added a description with WSDL examples of the new release procedure into the Introduction.
18 th Nov 2015	1.30	Changed web service version from 4.3 to 4.4. Added the new <hasPDF> element to the output of the searchPublications module.
9 th Oct 2015	1.29	Added the reinstated search term ‘HAS_PDF’ to the listSearchFields method and Appendix 1.

Date	Version	Details
6 th Oct 2015	1.28	Changed web service version from 4.2 to 4.3. Changed the scope of the files available for the getSupplementaryFiles method. All files now available for full text content with the exception of images, which are only available for Open Access content.
8 th Sept 2015	1.27	Added the <hasBook> and <bookid> elements to the output of the searchPublications module. Also added the new accession types. Added the new module getBookXML . Added the new Books related search fields HAS_BOOK and BOOK_ID to the listSearchFields method and Appendix 1 . Added 'NBK' code to the table in Section 3 – Data Sources .
27 th Aug 2015	1.26	Added new accession number types to Appendix 1 (Database citations sub-section).
15 th July 2015	1.25	Changed web service version from 4.1 to 4.2. Added the new pageSize parameter to the searchPublications , getCitations , getReferences , getDatabaseLinks and getTextMinedTerms methods.
12 th June 2015	1.24	Updated the statistics in the Introduction . Updated the statistics on the diagram in Section 2 .
24 th Mar 2015	1.23	Added the new EMBARGO_DATE search field to the listSearchFields method and Appendix 1 .
25 th Feb 2015	1.22	Changed web service version from 4.0 to 4.1. Updated the description for the FIRST_PDATE field in Appendix 1 . Included new elements in the 'core' response of the searchPublications module: electronicPublicationDate, firstPublicationDate, printPublicationDate and embargoDate.
4 th Dec 2014	1.21	Changed web service version from 3.0.1 to 4.0. Updated Europe PMC site statistics in Section 1 and content infographic in Section 2. Removed the dataset parameter from the profilePublications method. Removed the dataset parameter from the searchPublications method. Checked and corrected the parameter requirements and default values for the above methods. Note the the default synonym value is set to 'false'. Updated the listSearchFields method, removed the 'fulltext' value from the <dataSets> element. All fields are now combined into the same index. Updated Appendix 1 , moved the search fields EPMC_AUTH_MAN, NIH_AUTH_MAN, AUTH_MAN and AUTH_MAN_ID into their correct sub-section of 'core bibliographic'. Added the three new date fields E_PDATE, FIRST_PDATE and P_PDATE to Appendix 1 . Updated the WSDL URL in the Introduction .
12 th Sept 2014	1.20	Added the new 'LICENSE' search field to the listSearchFields method and Appendix 1 . Added the new manuscript search fields EPMC_AUTH_MAN, NIH_AUTH_MAN, AUTH_MAN and AUTH_MAN_ID to the listSearchFields module and Appendix 1 .

Date	Version	Details
2 nd Sept 2014	1.19	Updated the listSearchFields module and Appendix 1 to remove the search fields: HAS_PDF, HAS_HTML and HAS_FREE_FULLTEXT.
9 th July 2014	1.18	Added the 'affiliation' element to the 'authorList' of the searchPublications method. Added the new 'efo' parameter to the getTextMinedTerms method. Renamed the 'inUKPMC' element of the searchPublications method to 'inEPMC'.
30 th May 2014	1.17	Added the firstName element to the 'core' response of the searchPublications module.
25 th Mar 2014	1.16	Added a new Section (No.4) to describe how results can be sorted.
4 th Feb 2014	1.15	Removed references to CiteSeer records, these are no longer held in Europe PMC.
4 th Dec 2013	1.14	Added a new section of 'section-level' search terms to Appendix 1 . Updated the listSearchFields module output with the new 'section-level' search terms.
11 th Nov 2013	1.13	Added the 'citationType' element to the getCitations module. Added search terms to Appendix 1 : AUTHORID_TYPE, AUTHORID, LABS_PUBS. Added the 'authorId' and 'authorIdList' elements to the searchPublications module.
3 rd Sept 2013	1.12	Corrected an element name in the profilePublications module.
17 th July 2013	1.11	New getLabsLinks module included, Appendix 1 updated with new HAS_LABSLINKS search term, and new 'Database citations' section.
17 th May 2013	1.10	□ Updated statistics on the diagram in Section 2. ▣ Added <investigatorList> elements to searchPublications module, CORE response. ▣ Added ARXPR as a <dbname> to the <dbCrossReferenceList> of the searchPublications module. ▣ Added <hasTMAccessionNumbers> and <tmAccessionTypeList> to searchPublications method. ▣ Updated the getDatabaseLinks module to include the new ARXPR value for the database parameter. Also updated the table to show the definition of the elements returned when the <dbName> is ARXPR. ▣ Updated listSearchFields method with changes, notably the addition of the new fields: INVESTIGATOR, HAS_ARXPR, ARXPR_PUBS and ACCESSION_TYPE. Also updated Appendix 1 with the new search terms.
		Updated getTextMinedTerms with the revised list of values for <dbName>.
12 th Mar 2013	1.9	Updated site statistics in the Introduction .
6 th Mar 2013	1.8	Added the search term HAS_OMIM to Appendix 1 .
4 th Feb 2013	1.7	Updated Data Sources to reflect the EUROPEPMC category. Also updated the profilePublications method output with the EUROPEPMC category.
24 th Jan 2013	1.6	Replaced the search term 'IN_UKPMC' in Appendix 1 and

Date	Version	Details
		listSearchFields method output with 'IN_EPMC'.
11 th Jan 2013	1.5	Added missing search term 'HAS_CHEMBL' to Appendix 1 .
18 th Dec 2012	1.4	Updated to reflect the rebranding of the Web Service from CiteXplore to Europe PMC.
		Corrected getCitations method, 'Offset' parameter is required.
2 nd Nov 2012	1.3	Corrected the searchPublications and profilePublications methods, the synonym parameter is required.
26 th June 2012	1.2	Corrected case of a few element names, repaired broken links and included EBI Support link on cover page.
25 th June 2012	1.1	Corrected getTextMinedTerms 'source' parameter value explanation.
19 th June 2012	1.0	First released version

1. Introduction

This web service gives programmatic access to all of the publications and related information in Europe PMC. The following is included:

- ❑ Over 40 million abstracts including PubMed, Agricola, and international Patents from the European Patent Office. [Section 3](#) provides background information about the data sources.
- ❑ Over 3.3 million full text articles from PubMed Central, of which one million are open access.
- ❑ Database crosslinks to a number of public biomolecular databases, including UniProt, European Nucleotide Archive (EMBL), Protein Data Bank (PDB), InterPro and others listed in this document.
- ❑ Reference lists from full text articles
- ❑ Citing articles (i.e. articles that cited a given article)
- ❑ Terms that have been text mined from full text articles. These include: genes/proteins, diseases, Gene Ontology terms, organisms, database Accession Numbers, and chemicals.

The web service is Simple Object Access Protocol (SOAP)-based and has a modular architecture based on the above types of information retrieved. The WSDL to access the service can be found here:

<http://www.ebi.ac.uk/europepmc/webservices/soap?wsdl>

From January 2016 a new web service release procedure has been introduced. This allows two versions of the web service to be simultaneously available. This approach to release management will allow users to prepare for a new version, rather than having to immediately respond to a version change. The details of web service releases will be communicated to all known users. A mailing list of users is compiled from those that have supplied an email address in the 'Email' parameters of the various methods available.

Apart from the WSDL given above, two other WSDLs can be used, for example:

- ❑ New test version: <http://www.ebi.ac.uk/europepmc/webservices/test/soap?wsdl>
- ❑ Version specific reference (to current production or new test version as appropriate): <http://www.ebi.ac.uk/europepmc/webservices/ver4.5.0/soap?wsdl>

The table in [Section 2](#) gives an overview of the available methods. Further sections of this document provide more details on each method, explaining the input parameters and data element outputs. The Appendix provides details of the search syntax for the fields indexed, these are applicable to the [profilePublications](#) and [searchPublications](#) methods.

2. Method Summary

The method parameters and output elements are documented in the further sections of this document (use the hyperlinks in the table below).

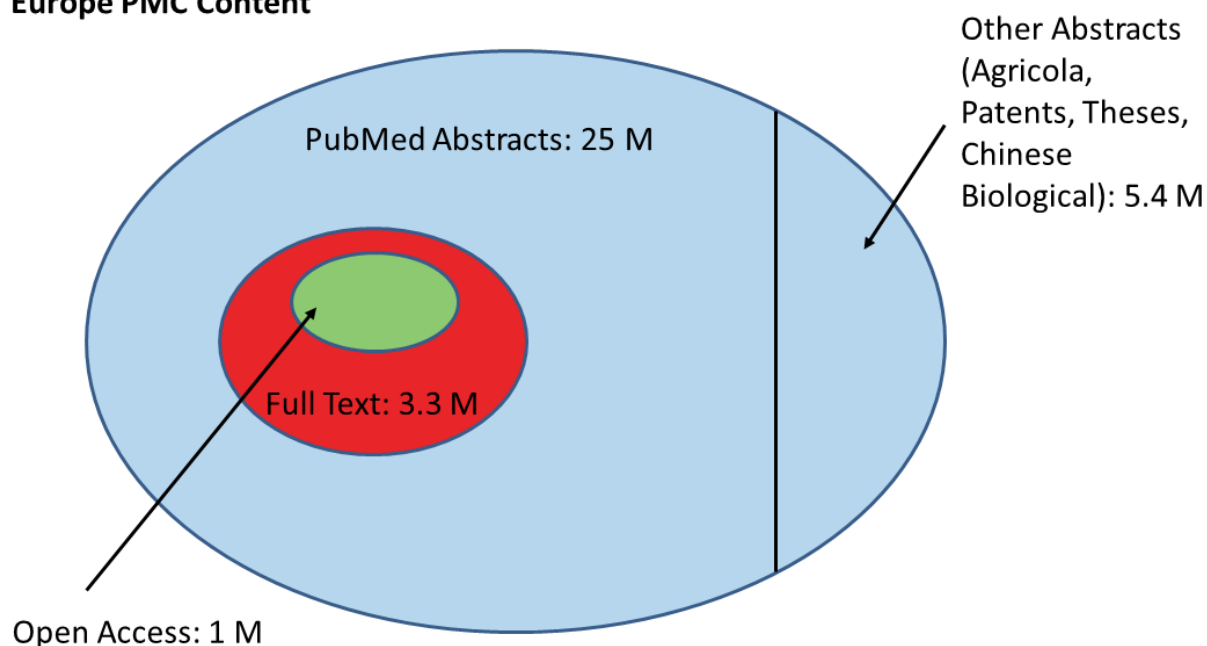
Module Name	Query	Purpose	Dataset
profilePublications	Query String or ID	This gives a profile of the dataset returned, listing records by categories of data source, publication type, and subset. A count of articles in each category is returned.	Citations and full text
searchPublications	Query String or ID	This is the main search method for returning publication metadata (i.e. article metadata). This method has three types of responses: ☐ idlist: only returns publication IDs ☐ lite: limited set of essential metadata ☐ core: all metadata	Citations and full text
getDatabaseLinks	ID and data source	Will retrieve all database cross-references for a given record; a count of total number of database cross-references, a count of each individual database cross-references, and metadata for the records from each. Databases included are: ARXPR, UniProt, PDB, Intact, EMBL, ChEMBL, ChEBI, OMIM, Pride and InterPro.	Citations and full text
getReferences	ID and data source	Gives the total number of references cited in a given article, plus authors, title, journal and publication year as ordered in the reference list.	Citations and full text
getCitations	ID and data source	Retrieves a count and list of articles that cite the query ID.	Citations and full text
getLabsLinks	ID and data source	This module gives access to the External Links provided by 3 rd parties (to extend Europe PMC content). See the Europe PMC External Links Service page for more details. Use this module to return the External Links for a	Citations and full text

		given source (source) and identifier (id) combination, and optionally by the external content provider (providerId).	
getFullTextXML	PMID or PMCID	Retrieves the full text XML of the article. The XML is returned as a SOAP attachment.	For Open Access articles
getBookXML	PMID or NBK number	This method will retrieve the book full text XML.	For Open Access books
getSupplementaryFiles	PMID or PMCID	Retrieves supplementary files for a full text article as a zipped SOAP attachment. Images are only available for Open Access content.	Full text
listSearchFields	N/A	Retrieves all the field names of the Lucene index in both the metadata and the fulltext index files.	N/A

The following table shows which web services methods apply to the datasets available:

All Citations	Full Text	Open Access Full Text
profilePublications	profilePublications	
searchPublications	searchPublications	
getDatabaseLinks	getDatabaseLinks	
getReferences	getReferences	
getCitations	getCitations	
getLabsLinks	getLabsLinks	
		getFullTextXML
		getBookXML
		getSupplementaryFiles

Europe PMC Content



The content scope of Europe PMC covers both abstracts and full text articles, with some full text articles being available as Open Access. All the full text articles in Europe PMC have a corresponding PubMed abstract record, but only about 10-12% of PubMed abstracts have a corresponding Europe PMC full text article available (see figure above).

In terms of the web service, all the content in Europe PMC can be searched. However, in the case of a full text search, in which the complete text of the article is searched, only the metadata and abstract will be sent in the response for most content. This is because most full text articles are under copyright and licencing restrictions that do not permit us to distribute them. However, for the open access subset of full text articles, it is possible to not only retrieve the metadata, but also the XML of the text, along with associated files such as figures and supplemental data.

Each full text article therefore has a unique PMCID and a corresponding PubMed ID (PMID). In these cases, the PMCID and PMID can be used interchangeably to retrieve the same results for the “get” requests. So for example, the list of citing articles returned in response to a PMCID request will be the same for the corresponding PMID. If an article does not have a PMID, for example if it is an Agricola record, with a unique ID in the format IND92070825, there will be no corresponding full text article.

3. Data Sources

Europe PMC contains data from a number of different sources, shown in the following table:

Code	Definition	Details of Data Source
AGR	Agricola	Agricola is a bibliographic database of citations to the agricultural literature created by the US National Agricultural Library and its co-operators. http://agricola.nal.usda.gov
CBA	Chinese Biological Abstracts	CBA: http://www.cba.ac.cn/ and the Shanghai Institutes for Biological Sciences (SIBS at http://www.sibs.ac.cn/) provide EBI with citation data not available in MEDLINE.
CTX	CiteXplore	Manual user-submitted records, added by the EBI.
ETH	EthOs Theses	PhD theses (British Library)
HIR	NHS Evidence	UK Clinical guidelines
MED	PubMed/MEDLINE NLM	PubMed: http://www.ncbi.nlm.nih.gov/pubmed NLM: the National Library of Medicine (NLM) is the world's largest biomedical library. It explores the uses of computer and communication technologies to improve the organization and use of biomedical information. http://www.nlm.nih.gov/ Medline: the National Library of Medicine's database of bibliographic citations and abstracts in the fields of medicine, nursing, dentistry, veterinary medicine, health care systems, and preclinical sciences. http://www.nlm.nih.gov/pubs/factsheets/medline.html
NBK	Europe PMC Book metadata	This source type denotes full text books on the Europe PMC Bookshelf that are not initially provided with a PMID 'MED' source metadata record. Where a full text book is received from the NCBI without a corresponding PMID 'MED' type metadata record, an 'NBK' metadata record will be created with the same 'NBK' number as the book. The metadata will be replaced by a PMID 'MED' type record if later made available. The full text book always retains its 'NBK' number.
PAT	Biological Patents	http://www.epo.org/
PMC	PubMed Central	PMC is a free full-text archive of biomedical and life sciences journal literature at the U.S. National Institutes of Health's National Library of Medicine (NIH/NLM). http://www.ncbi.nlm.nih.gov/pmc/
EUROPEPMC	Europe PMC	This category of data source is assigned for full

EBI Europe PMC SOAP Web Service 6.8.0 Reference Guide

© EMBL-EBI 2022, Document Version 1.50

		text queries to count the records in Europe PMC. This code is not usable as a parameter value.
--	--	---

4. Sorting Results

By default, the Europe PMC SOAP Web Service search results are sorted by relevance, with the most relevant result being presented first in the list.

The [searchPublications](#) method provides a 'sort' parameter for every single-valued field, e.g. P_PDATE_D, AUTH_FIRST, CITED etc. The sort field must be combined with "asc" or "desc" sort order.

The SOAP Web Service still retains the legacy sort method of including the sort parameter directly in the query string. The arguments in the 'queryString' parameter of the [searchPublications](#) method are: 'sort_date:y' and 'sort_cited:y'.

Examples:

Sort by date of publication, the most recently published article being the first result returned:

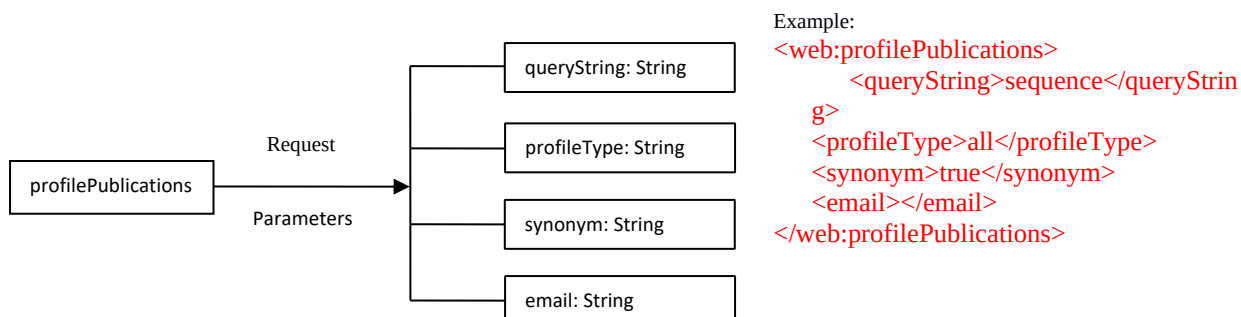
```
<queryString>auth:"Simon Hubbard" sort_date:y</queryString>
```

Sort by the number of citations, the most cited being the first result returned:

```
<queryString>auth:"Simon Hubbard" sort_cited:y</queryString>
```

5. Method: profilePublications

Request/Input Parameters:



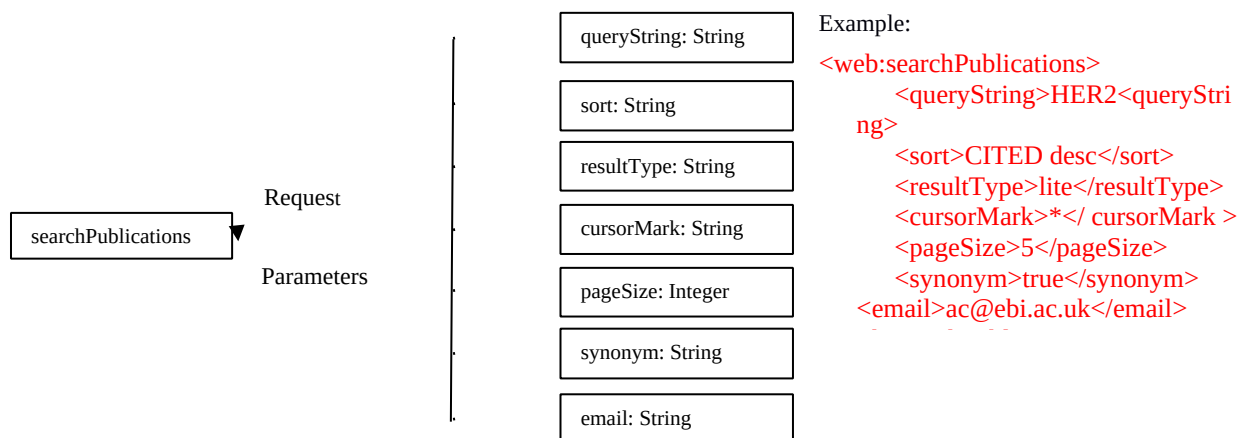
Method: profilePublications	This gives a profile of the dataset returned, listing records by categories of data source, publication type, and subset. A count of articles in each category is returned.		
Parameters	Required	Default	Values
queryString	Y		Any query (see Appendix 1), search string or ID
profileType	Y	all	Groups the data by the following categories: ☐ source – database source of record (See Section 3): ☐ pub_type – type of publication, possible values are listed in the next table of the elements returned. ☐ subset – i.e. BL for all records returned ☐ all – all of the above
synonym	N	false	“True”, “Y”, “YES”, “FALSE”, “N” or “NO” (case insensitive): Entering “TRUE” (and “Y”, “YES”) expands your query using the MeSH Terminology and UniProt synonyms list. For example aspirin, a synonym would be acetylsalicylic acid.
Email	N		Registering your email address allows us to contact send you web service-related news.

profilePublications Output:

Element	Description/Example
source	The count of the data sources as defined in Section 3. For example: <pre> <request> <queryString>sequence</queryString> <profileType>SOURCE</profileType> <synonym>true</synonym> </request> <profileList> <source name="AGR" count="17000"/> <source name="CBA" count="10243"/> <source name="CTX" count="313"/> <source name="ETH" count="1355"/> <source name="HIR" count="7"/> <source name="MED" count="1762123"/> <source name="PAT" count="136619"/> <source name="CIT" count="63"/> <source name="PMC" count="1784"/> </profileList> </pre>
pub_type	The count of the following categories of publication data: <pre> <request> <queryString>sequence</queryString> <profileType>PUB_TYPE</profileType> <synonym>true</synonym> </request> <profileList> <pubType name="all" count="1929507"/> <pubType name="full text" count="750784"/> <pubType name="open access" count="0"/> <pubType name="review" count="128744"/> </profileList> </pre>
Subset	The count for the 'BL' subset

6. Method: searchPublications

Request/Input Parameters:



Method: searchPublications	This is the main search method for returning publication metadata (i.e. article metadata). This method has three types of responses: ☐ idlist: only returns publication IDs ☐ lite: limited set of essential metadata ☐ core: all metadata		
Parameters	Required	Default	Values
queryString	Y		Any query (see Appendix 1), search string or ID
sort	N	score desc	Specify the sort field and sort order. This parameter provides “asc” or “desc” order for every single-valued field: P_PDATE_D, AUTH_FIRST, CITED etc. The sort order must be specified when using this parameter.
resultType	N	lite	idlist, lite or core – see the output section of this method for details
cursorMark	N	*	For the first page request with “*” (asterisk sign). For the following page use the returned value in the element nextCursorMark (for more detail see Pagination of Results).
pageSize	N	25	Range 0 to 1000
synonym	N	false	“True”, “Y”, “YES”, “FALSE”, “N” or “NO” (case insensitive): Entering “TRUE” (and “Y”, “YES”) expands your query using the MeSH Terminology and UniProt synonyms list. For example aspirin, a synonym would be acetylsalicylic acid.
Email	N		Registers the user’s email address with EBI. The user can then be contacted about Web

			Service related news.
--	--	--	-----------------------

searchPublications Method Output:

There are three output formats of the searchPublications method:

- ☐ idlist: returns only the IDs and source, where applicable, of the records
- ☐ lite: returns the key metadata of the record
- ☐ core: returns the full metadata for the record, including abstract, full text links, and MeSH terms

Element	Description	Output Format		
		idlist	lite	core
id	Identifier of the Article	✓	✓	✓
source	The source database of the citation, for example, MED is the code for PubMed. See Section 3 for full list of the 9 citation sources.	✓	✓	✓
pmid	Pubmed identifier of the article if applicable	✓	✓	✓
pmcid	Pubmed Central Identifier if full text is available in Europe PMC	✓	✓	✓
fullTextIdList fullTextId	List of fulltext Ids in the format below: <fullTextIdList> <fullTextId></fullTextId> <fullTextId></fullTextId> </fullTextIdList>	✓	✓	✓
doi	Digital Object Identifier of the article, e.g. 10.1073/pnas.0401194101		✓	✓
title	Title of the article		✓	✓
authorString	Comma separated authors list		✓	✓
journalTitle	Name of the journal which the article belongs to, e.g. PloS One		✓	
issue	Issue of the journal		✓	
journalVolume	Volume of journal		✓	
pubYear	Year of the publication		✓	✓
journalIssn	Journal ISSN number: If ISSN present it shows ISSN other wise it gives ESSN.		✓	
pageInfo	The page range: start page-end page, e.g. 145-178		✓	

Element	Description	Output Format		
		idlist	lite	core
pubType	Category of publication		✓	
brSummary	The summary of book or report details, e.g. Publ: Cancer Research UK 2009 (Pg.tot: 8)		✓	✓
isOpenAccess	Article Open Access status indicated with a Y or an N (only for full text)			✓
inEPMC	Value is either Y (yes) or N (no); this indicates whether the citation is available as full text in Europe PMC.		✓	✓
inPMC	Values: Y or N; this indicates whether the citation is available as full text in PMC		✓	✓
citedByCount	A count that indicates the number of times an article has been cited by other articles in our databases.		✓	✓
hasReferences	Values: Y or N; this indicates whether an article has a reference list or not. The value Y can only be returned for articles that are available as full text in Europe PMC.		✓	✓
hasTextMinedTerms	Values: Y or N; this indicates whether the article has any associated text-mined terms. The value Y can only be returned for articles that are available as full text in Europe PMC.		✓	✓
hasDbCrossReferences	Values: Y or N; this indicates whether an article has any cross references to biomolecular databases.		✓	✓
dbCrossReferenceList	If the value of hasDbCrossReferences is Y, then this element gives a list of the databases cross-referenced. E.g. <pre> <dbCrossReferencesList> <dbName>EMBL</dbName> <dbName>OMIM</dbName> <dbName>UNIPROT</dbName> <dbName>ARXPR</dbName> </dbCrossReferencesList> </pre> See also the ‘database’ parameter of the getDatabaseLinks method for database values and hyperlinks.		✓	✓
hasTMAccessionNumbers	Values: Y or N; this indicates whether an article has any text-mined accession number references to biomolecular databases.		✓	✓

Element	Description	Output Format		
		idlist	lite	core
tmAccessionTypeList accessionType	If the value of hasTMAccessionNumbers is Y, then this element gives a list of the databases cross-referenced. E.g. <pre> <tmAccessionTypeList> <accessionType>arrayexpress</accessionType> <accessionType>bioproject</accessionType> <accessionType>biosamples</accessionType> <accessionType>doi</accessionType> <accessionType>ega</accessionType> <accessionType>emdb</accessionType> <accessionType>ensembl</accessionType> <accessionType>eudract</accessionType> <accessionType>gen</accessionType> <accessionType>go</accessionType> <accessionType>interpro</accessionType> <accessionType>nct</accessionType> <accessionType>omin</accessionType> <accessionType>pdb</accessionType> <accessionType>pfam</accessionType> <accessionType>pxd</accessionType> <accessionType>refseq</accessionType> <accessionType>refsnp</accessionType> <accessionType>sprot</accessionType> <accessionType>treefam</accessionType> </tmAccessionTypeList> </pre> NOTE: The ‘doi’ accession type can be used to link to various data repositories.		✓	✓
hasPDF	A ‘Y’ value indicates that the publisher has provided a PDF version of the article.		✓	✓
hasBook	A ‘Y’ value indicates that full book text is available.		✓	✓

Element	Description	Output Format		
		idlist	lite	core
hasSuppl	A 'Y' value indicates that the article has supplemental data associated (and available) with it.		✓	✓
hasLabsLinks	A 'Y' value indicates that the article is associated with External Links .		✓	✓
hasData	A 'Y' value indicates that the article is associated with data-literature links tagged as 'related_data' or 'supporting_data'.			✓
license	The Creative Commons license (where provided) that is assigned to the article.			✓
authMan	A 'Y' value indicates that the article (author manuscript) is Europe PMC Funded, and was submitted by the author (or representative) via Europe PMC plus .			✓
epmcAuthMan	A 'Y' value indicates that the article (author manuscript) is Europe PMC Funded, and was submitted by the author (or representative) via Europe PMC plus .			✓
nihAuthMan	A 'Y' value indicates that the article (author manuscript) is Europe PMC Funded, and was submitted by the author (or representative) via Europe PMC plus .			✓
manuscriptId	A 'Y' value indicates that the article (author manuscript) is Europe PMC Funded, and was submitted by the author (or representative) via Europe PMC plus .			✓
luceneScore	Gives the rank of the Lucene search engine, the greater the number the higher the ranking.		✓	✓
bookid	If a full book text is available (see 'hasBook' above), then its 'NBK' number will be provided.		✓	✓

Element	Description	Output Format		
		idlist	lite	core
authorList author fullName firstName lastName initials authorId authorAffiliationDetailsList authorAffiliation affiliation affiliationOrgId	List of author details which contains full name, initials and last name: Full Name of author, e.g. Hubbard T First Name of author, e.g. Tim Last Name of author, e.g. Hubbard Initials of author, e.g. T The ID type and value (where available), e.g. type="ORCID">0000-0002-1767-9318 The author's affiliation, note that affiliations for multiple authors are available for MEDLINE records (where provided) from the beginning of 2014. The author organization affiliation identifier is a persistent identifier that uniquely identifies the institutional body listed as an author's affiliated research organization. Examples include ror.org IDs, GRIDs, ISNIs, wikidata IDs.			✓
authorIdList authorId	All the author IDs linked to the article are listed together with the type: e.g. type="ORCID">0000-0002-3908-1122			✓
investigatorList investigator fullName lastName initials	List of investigator details which contains full name, initials and last name: Full Name of author, e.g. Orlandini F Last Name of author, e.g. Orlandini Initials of author, e.g. F			✓

Element	Description	Output Format		
		idlist	lite	core
journalInfo issue volume journalIssueId dateOfPublication monthOfPublication yearOfPublication printPublicationDate journal title ISOAbbreviation medlineAbbreviation NLMid ISSN ESSN	journalInfo element gives the full details of a journal Issue of the journal Volume of the journal Identifier of the journal issue Date of publication Month of publication Year of publication Print publication date of journal issue, when an article appeared in print format Title of the journal ISO abbreviation of journal name Medline abbreviation of journal name National Library of Medicine's Identifier ISSN number of the journal ESSN number of the journal			✓
abstractText	The abstract of the article			✓
affiliation	The affiliation of the article, e.g. Health Protection Agency Microbiology Services, Birmingham, UK.			✓
language	Language of the article			✓
pubModel	The publication model of the article, e.g. Print-Electronic or media			✓
pubTypeList pubType	Lists all the categories of the publication. A list of the publication categories. E.g. journal article, Research Article or report			✓

Element	Description	Output Format		
		idlist	lite	core
brSummary ☐ bookOrReportDetails edition extraInformation isbn10 isbn13 numberOfPages publisher seriesIssn seriesName comprisingTitle comprisingTitleNonAs cii dayOfPublication monthOfPublication yearOfPublication	brSummary gives the summary of book or report details. E.g. Publ:Cancer Research UK 2009 (Pg.tot: 8) bookOrReportDetails: gives the details of book or report information Edition information Extra information 10 digit International Standard Book Number 13 digit International Standard Book Number Total number of pages Publisher name Name of the series Non-ascii title Day of Publication Month of publication Year of publication		✓	✓
grantsList grant grantId agency acronym orderIn	List of all grants Identifier of the Grant, e.g. CA-06927 Agency of grant, e.g. NCI NIH HHS Acronym of the grant, e.g. CA Order, e.g. 1			✓

Element	Description	Output Format		
		idlist	lite	core
meshHeadingList meshHeading majorTopic_YN descriptorName meshQualifierList abbreviation qualifierName majorTopic_YN	meshHeadingList contains the list of all meshTerms for the article. Descriptor name meshQualifierList is the list of mesh Qualifier information Qualifier abbreviation Qualifier name Major topic Y/N indicator			✓

Element	Description	Output Format		
		idlist	lite	core
patentDetailsInfo countryCode country typeCode typeDescription classifierList classifier classification classificationType hyperlink application applicationNumber applicationDate orderIn priorityList priority priorityNumber priorityDate orderIn familyList family familyNumber orderIn	patentDetailsInfo (given when the record is a patent) details of the patent article Code name of the country, e.g. AU Name of the country, e.g. Australia Patent type code, e.g. A1 Patent Type Description, e.g. Comp. Spec. open to Pub. Insp. Classifier List gives list of all classifiers information Patent classification, e.g. A61K31/4025 Type of classification, e.g. EPO URL of classification Patent application information Number of the application Date of the application Order of the application This gives the list of priorities Number of priority, e.g. US20090224524P Date of priority Order of priority List of patent families Family number Order of family			✓
keywordList keyword	List of Keyword names Name of the keyword, e.g. Holocene epoch			✓

Element	Description	Output Format		
		idlist	lite	core
chemicalList chemical name registryNumber	chemicalList contains the chemicals information Name of the chemical, e.g. Granzymes, Peptides, Hydrogen Registry Number, e.g. EC 3.4.21.			✓
subsetList subset code name	SubSet list contains the subsets information Code of the subset, e.g.IM Name of the subset, e.g. Index Medicus			✓
fullTextUrlList fullTextUrl availability availabilityCode documentStyle site url	This list contains the full text web link information Availability of the article. See table that follows for list of all possible values Code of availability Format of the fulltext document Ex: HTML, PDF , abs(abstract)or DOI Name of website which has fulltext document URL of the document			✓
commentCorrectionList commentCorrection id source reference type note orderIn	This shows the comments and corrections of an article Identifier of the article. Source of the article Reference of the article, e.g. N Engl J Med. 2004 Jul 8;351(2):200 Type of comment, e.g. Comment in, Comment On or Erratum in etc Note. Order of the comment, e.g. 10			✓
extCommentList extComment extCommentSource info1 info2	The comments list is extracted from cross reference database name CRD(Centre for Reviews and Dissemination) Cross reference database name, e.g. CRD Info1 is the identifier of the record Info2 has either DARE or EED value			✓

Element	Description	Output Format		
		idlist	lite	core
dateOfCompletion	Article or book or report completion date			✓
dateOfCreation	Article or book or report creation date			✓
dateOfRevision	Article or book or report revision date			✓
elecronicPublicationDate	Electronic publication date, when an article was first published online			✓
firstPublicationDate	The date of first publication, whichever is first, electronic or print publication. Where a date is not fully available e.g. year only, an algorithm is applied to determine the value		✓	✓
embargoDate	The date from which Europe PMC is permitted to provide access to the full text article.			✓
firstIndexDate	The date when the article has been first indexed in Europe PMC		✓	✓
versionNumber	The preprint version number of article (if exists)		✓	✓
versionList version id source firstPublishDate versionNumber pubTypeList pubType	This list contains the preprint versions information of the article (if exists) Identifier of the article corresponding to this version Source of the article corresponding to this version The date when the article corresponding to this version has been first published Number of the version The publication type of the article which can be "preprint", "preprint-withdrawal" or "preprint-removal"			✓
fullTextReceivedDate	Date when fulltext article is first loaded in Europe PMC			✓
DataLinksTagsList dataLinkstag	available data links tags to be used in retrieving data links using tag parameter			✓
publicationStatus	New field indicating the status of article if it is published or ahead of print.			✓

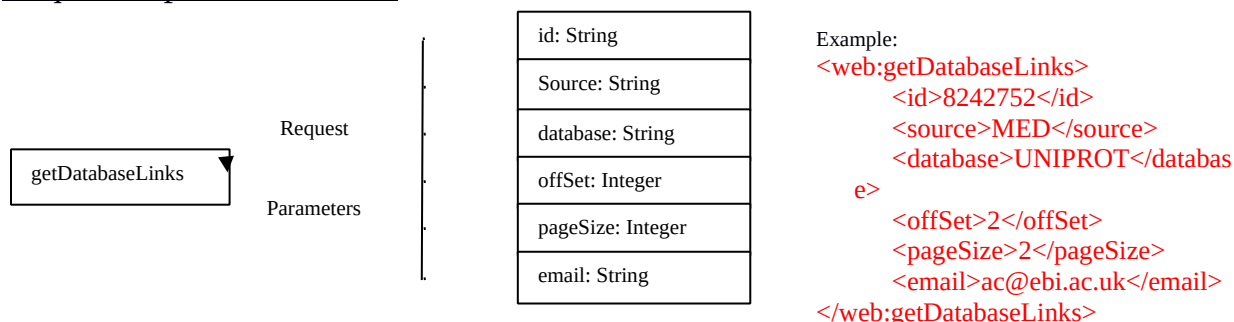
Availability Codes

The [fullTextUrlList](#) section in the above table refers to a list of availability codes. The possible values are shown below:

Availability Code	Description
F	Free
F2	Free after 2 months
F4	Free after 4 months
F6	Free after 6 months
F12	Free after 12 months
F18	Free after 18 months
F24	Free after 24 months
F36	Free after 36 months
OA	Open Access
S	Subscription required
U	Unrestricted

7. Method: getDatabaseLinks

Request/Input Parameters:



Method: getDatabaseLinks	Will retrieve all database cross-references for a given record; a count of total number of database cross-references, a count of each individual database cross-references, and metadata for the records from each.		
Parameters	Required	Default	Values
id	Y		An ID related to the data source as defined in Section 3 .
Source	Y		The data sources code as defined in Section 3 .
Database	N	All values	☐ ARXPR ☐ CHEBI ☐ CHEMBL ☐ EMBL ☐ INTACT ☐ INTERPRO ☐ OMIM ☐ PDB ☐ PRIDE ☐ UNIPROT
offSet	N	0	Offset 0 retrieves the first page, offset 1 the second and so on.
pageSize	N	25	Range 0 to 1000
Email	N		Registers the user's email address with EBI. The user can then be contacted about Web Service related news.

getDatabaseLinks Method Output:

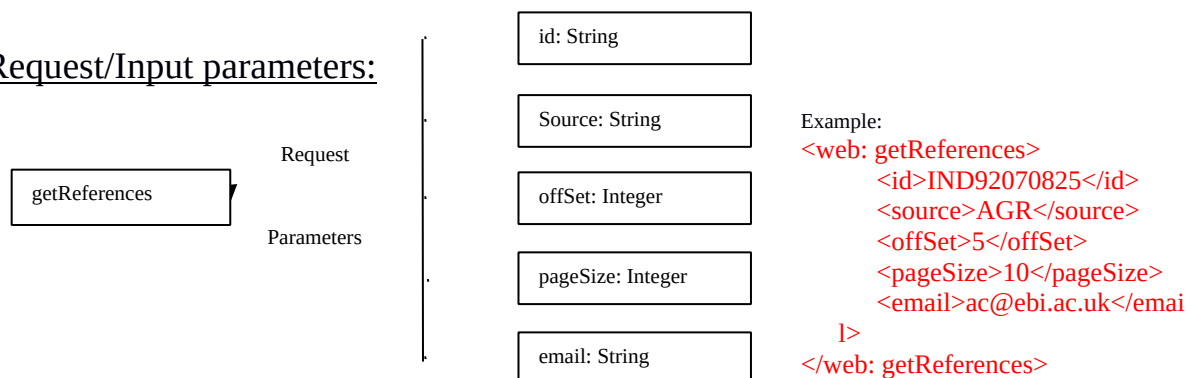
Element	Data Type	Description	Example
dbCountList	ArrayList	It gives the list of all databases counts	
dbName	String	Name of the database	UNIPROT , other possible values are also shown in the table below
count	Integer	Total number of results for a database	8
dbCrossReferenceList		Lists the cross-references	
dbCrossReference	ArrayList	This element gives the cross-reference information for each database returned	
dbName	String	Name of the database	UNIPROT , other possible values are also shown in the table below
dbCount	Integer	Total number of results for a database	1
dbCrossReferenceInfo		Cross-reference information section	
info1	String	Dependent on the dbName value	See table below
info2	String	Dependent on the dbName value	See table below
info3	String	Dependent on the dbName value	See table below
info4	String	Dependent on the dbName value	See table below

Use of Elements info1 to 4 for the getDatabaseLinks Method

dbName	INFO Element	Description/Example
UNIPROT	info1	UniProt database number
	info2	Protein name
	info3	Organism
	info4	Source of the cross-reference (e.g. "UniProt")
EMBL	info1	EMBL/GenBank/DDBJ database id
	info2	Description of nucleotide sequence record
	info3	Sequence length
PDB	info1	PDB database id
	info2	Experiment type
	info3	Protein structure name
INTERPRO	info1	InterPro database id
	info2	Protein family/domain short name
	info3	Protein family/domain name
OMIM	info1	OMIM database id
	info2	Reference number in OMIM record ("1", "2", etc.)
	info3	Type of record (e.g. "gene", "description, not locus")
	info4	Title
CHEBI	info1	ChEBI database id
	info2	Chemical entity name
	info3	Type of record ("CHEM")
ChEMBL	info1	ChEMBL database id
	info2	Entity name or description
	info3	Type of entity (protein or organism target, compound, assay: "PROT", "ORG", "CHEM", "ASSAY")
INTACT	info1	IntAct database id
	info2	Experiment name
	info3	Interaction detection method ("pull down", "protein array" etc.)
ARXPR	info1	ArrayExpress accession (e.g. E-GEOD-23504)
	info2	ArrayExpress ID (e.g. 269645)
	info3	Bibliography accession (PMID)
PRIDE	info1	PRIDE ID (e.g. PXD004668)
	info2	PRIDE Name (e.g. Interactome of Nematostella GW182 in HEK293 cells)
	info3	Omics Type (e.g. Proteomics)
	Info 4	URL link (e.g. http://www.ebi.ac.uk/pride/archive/projects/PXD004668)

8. Method: getReferences

Request/Input parameters:



Method: getReferences	Gives the total number of references cited in a given article, plus authors, title, journal, publication year as ordered in the reference list.		
Parameters	Required	Default	Values
id	Y		An ID related to the data source as defined in Section 3 .
Source	Y		The data sources code as defined in Section 3 .
offSet	N	0	Offset 0 retrieves the first page, offset 1 the second and so on.
pageSize	N	25	Range 0 to 1000
Email	N		Registers the user's email address with EBI. The user can then be contacted about Web Service related news.

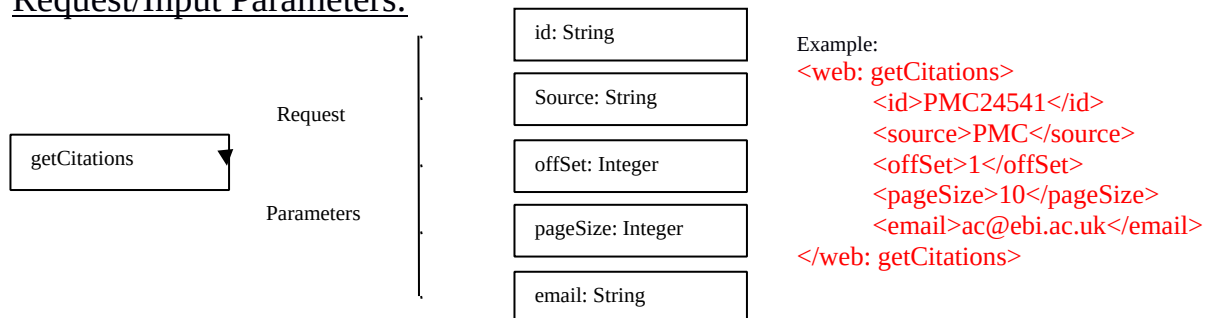
getReferences Method Output:

Element	Data Type	Description	Example
id	String	Identifier of the article	ADL88035909
source	String	Source of the article	AGR (See Section 3 for list all sources)
citationType	String	Type of article	journal article
title	String	Title of the article	Modern biogeographic theory: are there any lessons for nature reserve design?.
authorString	String	Comma separated authors list	Margules CR, Nicholls AO, Pressey RL.
journalAbbreviation	String	ISO abbreviation of the journal name	Biol. Conserv.
Issue	String	issue	1
pubYear	Integer	Publication year of the journal	1988
volume	String	volume	43
ISSN	String	ISSN number	0006-3207
ESSN	String	ESSN	
ISBN	String	ISBN	1842571443
pageInfo	String	Range of page information	63-76
publicationTitle	String	Source publication name (for an unmatched reference)	e.g. "Nucleic Acids Res."
citedOrder	Integer	Order of reference	123
match	String	Indicates whether this reference is a known (matched) publication	N
publisherLoc	String	Location of the publisher	Philadelphia, PA

publisherName	String	Name of the publisher	Hanley & Belfus, Inc
seriesName	String	Name of a series of books or reports	proc sixth annual san rancisco cancer symposium oct 1970 frontiers of rad ther oncology
edition	String	Edition of book or report	4
editors	String	Editor(s) of book or report	Gershwin ME, Vierling JM, Manns MP.
Doi	String	Digital Object Identifier	10.1016/0006-3207(88)90007-9
unstructuredInformation	String	Unstructured metadata (for an unmatched reference)	Boublik, J.H. <i>et al.</i> <i>Nature</i> 301, (1983).
externalLink	String	Link to full text when metadata is incomplete	uri:http://www.doh.gov.uk/gmscontract/neshomeless.pdf:http://www.doh.gov.uk/gmscontract/neshomeless.pdf
comments	String	comments	DHHS Publ. No. (NIH) 91-3242, 5 th Ed.

9. Method: getCitations

Request/Input Parameters:



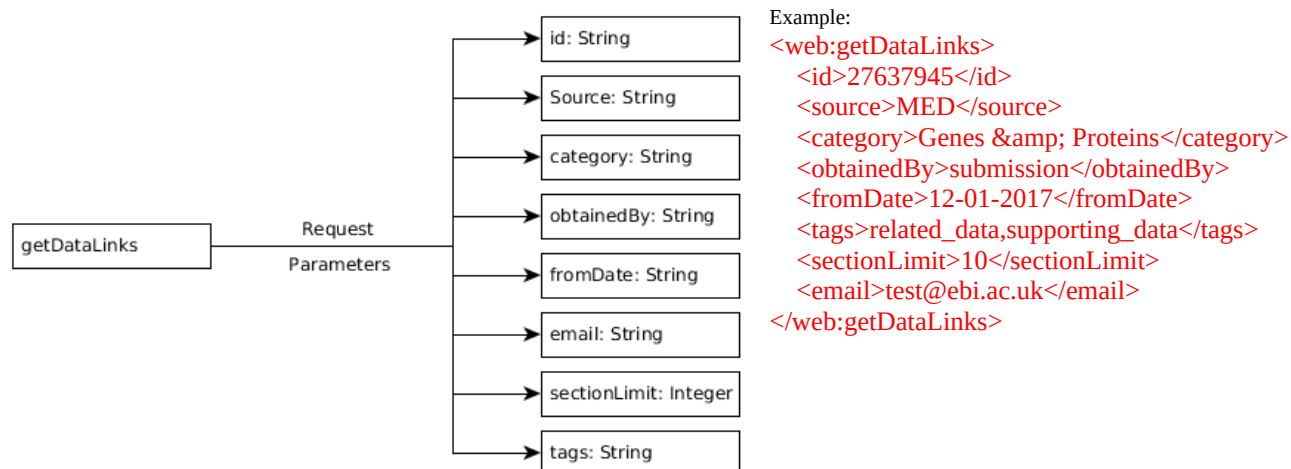
Method: getCitations	Retrieves a count and list of articles that cite the request ID.		
Parameters	Required	Default	Values
id	Y		An ID related to the data source as defined in Section 3 .
Source	Y		The data sources code as defined in Section 3 .
offSet	N	0	Offset 0 retrieves the first page, offset 1 the second and so on.
pageSize	N	25	Range 0 to 1000
Email	N		Registers the user's email address with EBI. The user can then be contacted about Web Service related news.

getCitations Method Output:

Element	Description	Example
id	Identifier of the article	17211498
source	Source of the article	MED (See Section 3 for list all sources)
citationType	Type of publication	Journal Article
title	Title of the article	Functional genomics analysis of low concentration of ethanol in human hepatocellular carcinoma (HepG2) cells. Role of genes involved in transcriptional and translational processes.
authorString	Comma separated authors list	Castaneda F, Rosin-Steiner S, Jung K
journalAbbreviation	ISO abbreviation of the journal name	Int J Med Sci
pubYear	Publication year of the journal	2007
volume	Journal volume	4
issue	Journal issue	1
pageInfo	Start to end page of the article (note, for some journals, this is of the form 310-25, which means 310-325)	28-35
citedByCount	The number of times the publication has been cited by other publications in the Europe PMC database	2
text	This is the fragment of text from the fulltext article immediately preceding the citation	To evaluate the gene expression profile between group 1 (ethanol-treated) and group 2 (control) a hierarchical cluster analysis was performed

10. Method getDataLinks

Request/Input Parameters:



Method: getDataLinks	This method gives combined access to Data Links independent of the way they arrived at Europe PMC. Links available here will either have been obtained by text mining, external links or our Database Crosslinks process. Links are returned in Scholix format.		
Parameters	Required	Default	Values/Examples
id	Y		An ID related to the data source as defined in Section 3 .
Source	Y		The data sources code as defined in Section 3 .
category	N		The link category/provider. Possible values listed in Table X
obtainedBy	N		The process through which these links were obtained. (one of: tm_accession, tm_term, ext_links, submission)
fromDate	N		A lower date cutoff for data link retrieval. Only datalinks created at this date or later are returned.
tags	N		Tags the Europe PMC website uses to request datalinks
sectionLimit	N		The maximum number of links per section in the response.
email	N		Registers the user's email address with EBI. The user can then be contacted about Web Service related news.

Possible Category Names:

Category Name	Database Name
Genes & Proteins	SwissProt/UniProt
Protein Structures	PDBe
Nucleotide Sequences	ENA
Protein Families	InterPro
Protein Interactions	IntAct
Diseases	OMIM
Chemicals	CHEBI/CHEMBL
Functional Genomics Experiments	ArrayExpress

Proteomics Data
 Gene Ontology (GO) Terms
 Data Citations
 SNPs
 Pfam
 Clinical Trials

 Electron Microscopy Data Bank
 BioSamples
 ProteomeXchange
 Treefam
 Quick GO
 European Genome-Phenome Archive
 Open Access at Bielefeld University
 PhenoMiner
 WormBase
 Publons
 Access to Understanding
 Wellcome Trust
 Arthritis Research UK
 Neuroscience Information Framework
 iPTMnet
 FlyBase
 OpenAIRE
 Dryad Digital Repository
 EMBL Press Releases
 Kudos
 PANGAEA
 Open Access at Lund University
 DEPOD
 GenomeRNAi
 HAL Open Archive
 EBI Train Online
 Centre for Reviews and Dissemination (UK)
 Wikipedia
 Linköping University Digital Archive
 EuroFIR Document Repository
 NHGRI-EBI GWAS Catalog
 BioStudies
 EBI Metagenomics
 National Centre for Text Mining (NaCTeM)
 Bibliomics and Text Mining Group (BiTeM)
 DisGeNET
 IntAct
 Open Targets Platform
 GOA Project

PRIDE
 Gene Ontology
 DOIs
 NCBI dbSNP
 Pfam
 ClinicalTrials.gov/EU Clinical Trials
 Register
 Electron Microscopy Data Bank
 BioSamples
 ProteomeXchange
 Treefam
 Quick GO
 European Genome-Phenome Archive

Altmetric
Marie Curie Press Releases
Worldwide Cancer Research
Ximbio
Related Immune Epitope Information - Immune
Epitope Database and Analysis Resource

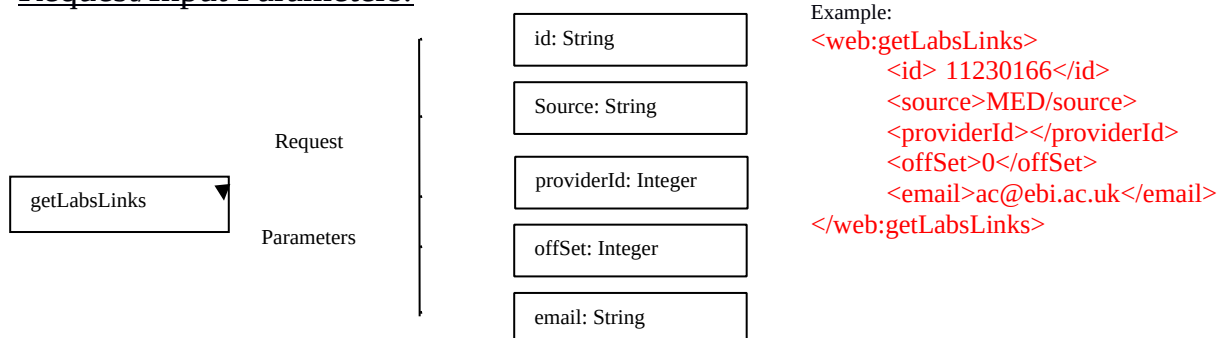
getDataLinks Method Output:

Element	Description
dataLinkList	List of data link objects
Category Name NameLong Tags CategoryLinkCount Section	Wrapper for a category of links. The name of the category. This typically is the name of the linked database. A longer, more descriptive version of the category. List of Tags to identify and request links of the category. Number of total links in this category including the ones cut off through the section limit Section of links objects
Tag	Element of the tags list. Identifies groups of links. Possible tags are 1. related_data 2. supporting_data A datalink does not need to have a tag.
Section ObtainedBy Tags SectionLinkCount CollectionURL Linklist Link	Pubmed Central Identifier if full text is available in Europe PMC Describes how the links were obtained. One of: submission, ext_links, tm_accession, tm_term. List of tags to identify and request links of the section. Number of total links in this section including the ones cut off through the section limit Link to a collection of external data records. List of link packages Individual scholix link package
Link ObtainedBy PublicationDate LinkProvider RelationshipType	Link package object Describes method of obtaining the link. One of: submission, ext_links, tm_accession, tm_term. Point of time the link package was made public (format: DD-MM-YYYY) Object describing the provider of the link package. Descriptor of the relationship between the Source and Target object

Source Target Frequency	The source object in the relationship. In Europe PMC this is always the publication. The target object in the relationship. Here this is always the external target of the relationship. The number of times a text
LinkProvider Name	Object describing the provider of the link package. The name of the link provider. This will be 'Europe PMC' as we are providing these links.
RelationshipType Name	Descriptor of the relationship between the Source and Target object. Name of the relationship. One of: IsSupplementTo, IsSupplementedBy, References, IsReferencedBy, IsRelatedTo. For more detail refer to http://www.scholix.org/schema .
Source/Target Type Identifier Title Publisher ImageURL	Linked elements of a link package Link element complex type element Identifiers Title of the linked object (e.g. "Antifungal protein 2") Publisher description complex type Link to an image supporting ext_links type datalinks
Type Name SubType	Linked element type Name of the linked element type. Either "dataset" or "literature". Subtype of the linked element.
Identifier ID IDScheme IDURL	Identifier complex type. The identifier (e.g. 10.1016/j.jamcollsurg.2013.04.039 for a doi) Schema of identifier (e.g. "doi") A resolveable URL based on the identifier (e.g. https://doi.org/10.1016/j.jamcollsurg.2013.04.039)
Publisher Name	Publisher description complex type Name of the object publisher (e.g. "Europe PMC" or "UniProt")

11. Method: getLabsLinks

Request/Input Parameters:



Method: getLabsLinks	This module gives access to the External Links provided by 3 rd parties (to extend Europe PMC content). See the Europe PMC External Links Service page for more details. Use this module to return the External Links for a given source (source) and identifier (id) combination, or by the external content provider (providerId).		
Parameters	Required	Default	Values
id	Y		An ID related to the data source as defined in Section 3 .
Source	Y		The data sources code as defined in Section 3 .
providerId	N		The identifier – as assigned to a provider.
offSet	Y	0	Results are returned in batches of 25 records. Offset 0 retrieves the first 25, offset 1 the second 25 and so on.
Email	N		Registers the user's email address with EBI. The user can then be contacted about Web Service related news.

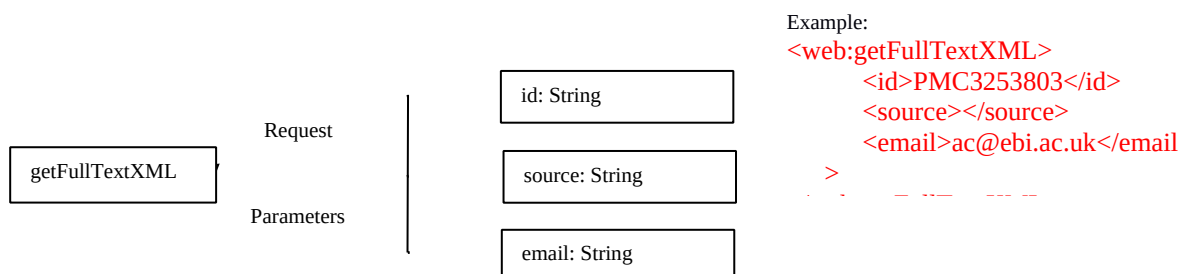
getLabsLinks Output:

Element	Data Type	Description	Example
linksCountList		The section element listing the link providers together with the count of their links	
provider		The element heading for each provider	
providerName	string	The name of the links provider	EBI Train Online
linksCount	integer	The number of links available from each provider	1
providers		The section element listing the link URLs made available from the providers	
provider		The element heading for each provider	
id	integer	The identifier of the provider	4321
name	string	The name of the links provider	EBI Train Online
description	string	The description of the type of links provided	Train Online Courses from the EMBL-European Bioinformatics Institute
frontTab	string	Unused element for potential future implementation	Y or N
link		The URL link section element	
title	string	The optional title associated with the URL link	Proteomics: an introduction to the EBI resources
url	string	The actual URL of the link provided	http://www.ebi.ac.uk/training/online/course/proteomics-introduction-ebi-resources
imgUrl	string	An optional image URL.	https://api.altmetric.com/v1/donut/3421689_64.png

12. Method: getFullTextXML

Reference can also be made to: <http://dtd.nlm.nih.gov/archiving/>. This describes the format of the XML output.

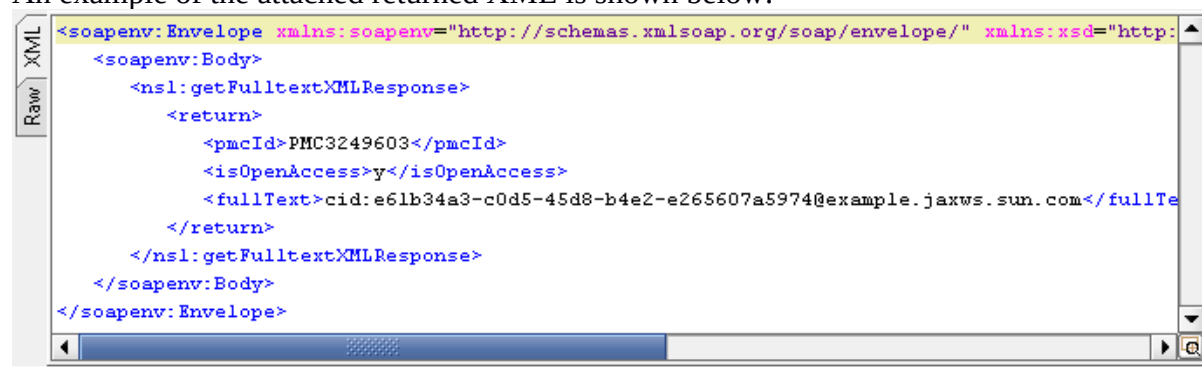
Request/Input Parameters:



Method: getFullTextXML	Retrieves the full text XML of the article. The XML is returned as a SOAP attachment. The full text XML is available only for the full-text OA subset of the Europe PMC database.		
Parameters	Required	Default	Values
id	Y		A PubMed (PMID) or PubMed Central ID (PMCID)
source	N		Only articles with a PMC data source are available. See Section 3 for data source definitions.
Email	N		Registers the user's email address with EBI. The user can then be contacted about Web Service related news.

getFullTextXML Method Output:

An example of the attached returned XML is shown below:

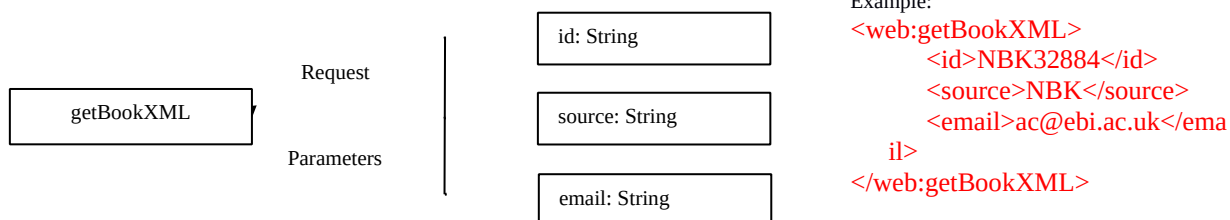


```
<?xml version='1.0' encoding='UTF-8'>
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/" xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  <soapenv:Body>
    <ns1:getFulltextXMLResponse>
      <return>
        <pmcId>PMC3249603</pmcId>
        <isOpenAccess>y</isOpenAccess>
        <fullText>cid:e61b34a3-c0d5-45d8-b4e2-e265607a5974@example.jaxws.sun.com</fullText>
      </return>
    </ns1:getFulltextXMLResponse>
  </soapenv:Body>
</soapenv:Envelope>
```

Name	Content type	Size	Part	Type	Content
e61b34a3-c0d5-45d8-b4e2-e265607a5974@example.jaxws.sun.com	application/xml	76524	e61b34a...	MIME	<e61b3...

13. Method: getBookXML

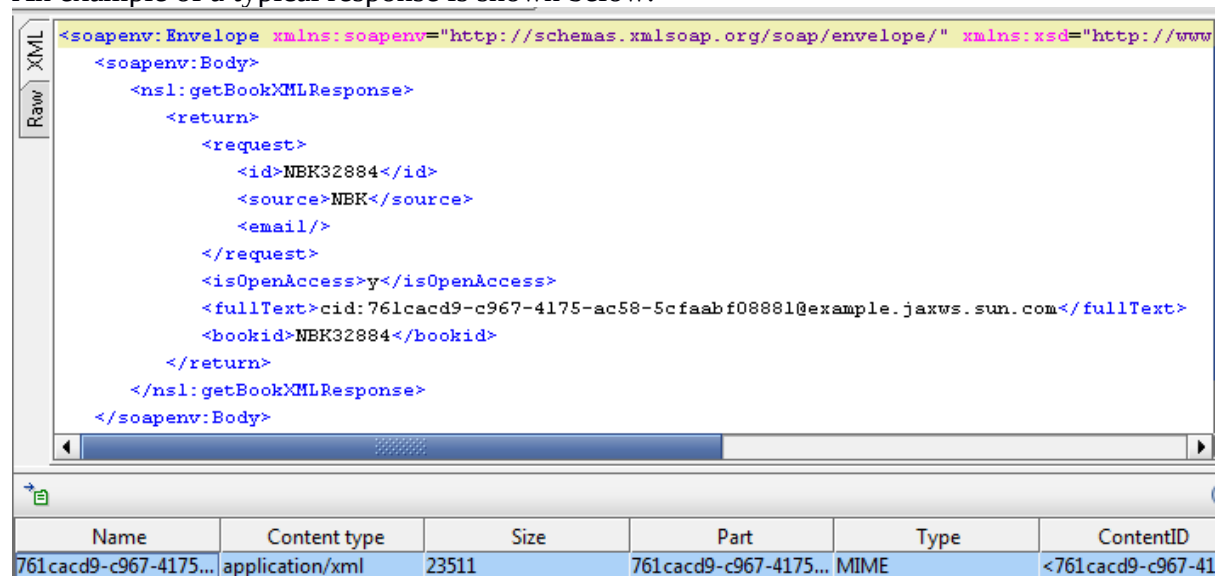
Request/Input Parameters:



Method: getBookXML	Retrieves the full text XML for an Open Access Book as a SOAP attachment.		
Parameters	Required	Default	Values
id	Y		A PubMed (PMID) or Europe PMC 'NBK' number.
Source	Y		Use 'MED' with a PMID, or use 'NBK' with a Europe PMC 'NBK' number. See Section 3 for more details of Europe PMC data sources.
Email	N		Registers the user's email address with EBI. The user can then be contacted about Web Service related news.

getBookXML Output:

An example of a typical response is shown below:



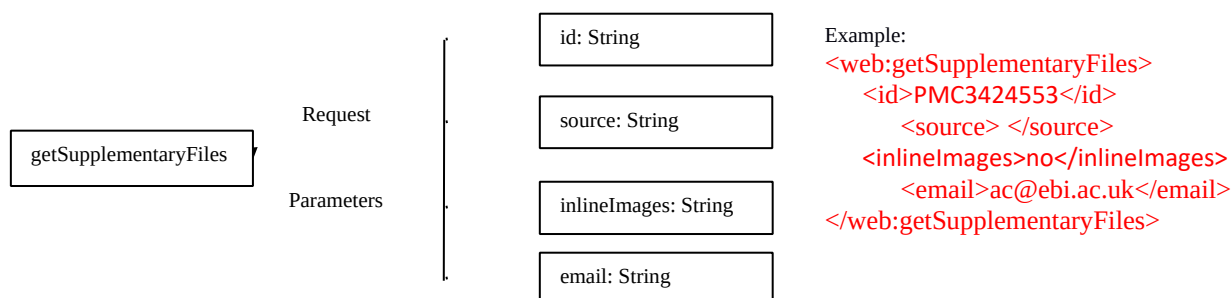
```
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/" xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  <soapenv:Body>
    <ns1:getBookXMLResponse>
      <return>
        <request>
          <id>NBK32884</id>
          <source>NBK</source>
          <email/>
        </request>
        <isOpenAccess>y</isOpenAccess>
        <fullText>cid:761cacd9-c967-4175-ac58-5cfaabf08881@example.jaxws.sun.com</fullText>
        <bookid>NBK32884</bookid>
      </return>
    </ns1:getBookXMLResponse>
  </soapenv:Body>
```

Name	Content type	Size	Part	Type	ContentID
761cacd9-c967-4175...	application/xml	23511	761cacd9-c967-4175...	MIME	<761cacd9-c967-41

The XML can then be retrieved from the attached file.

14. Method: getSupplementaryFiles

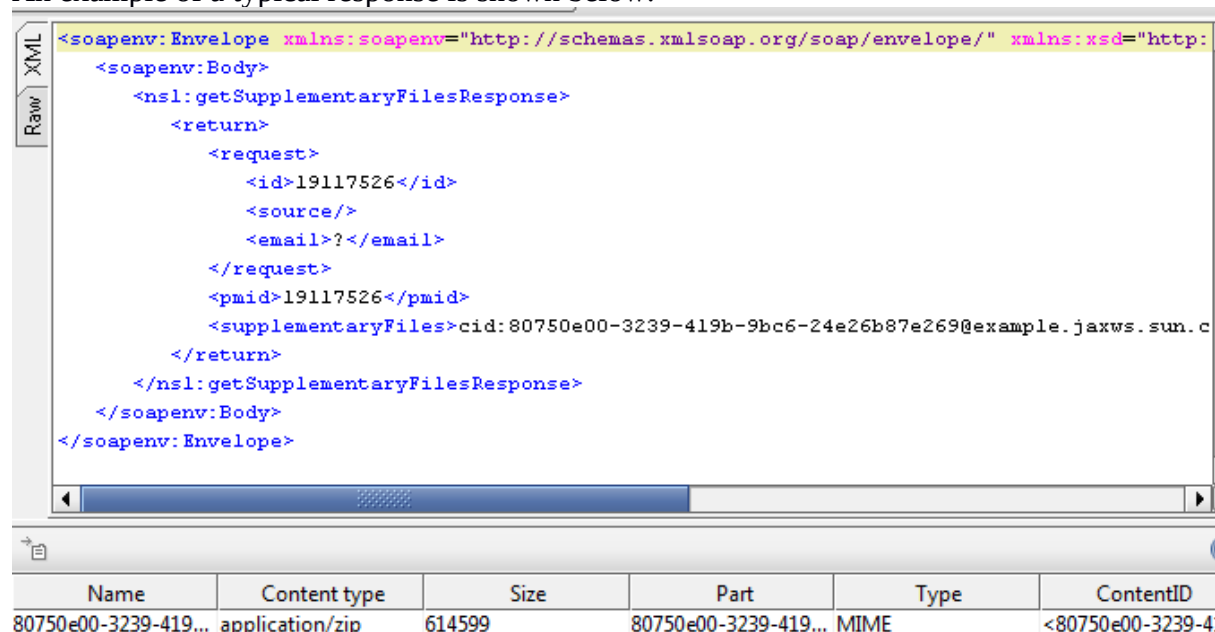
Request/Input Parameters:



Method: getSupplementaryFiles	Retrieves supplementary files for a full text article as a zipped SOAP attachment. Images are only available for Open Access content.		
Parameters	Required	Default	Values
id	Y		A PubMed (PMID) or PubMed Central ID (PMCID)
source	N	N/A	This parameter is not currently used
inlineImages	N	Y	By setting this parameter to “n”, “no” or “false” the inline images get excluded from the download.
email	N		Registers the user’s email address with EBI. The user can then be contacted about Web Service related news.

getSupplementaryFiles Output:

An example of a typical response is shown below:



The screenshot displays a SOAP response in a web browser. The top section shows the raw XML, and the bottom section shows a table of the supplementary files.

```
<?xml version='1.0' encoding='UTF-8'>
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/" xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  <soapenv:Body>
    <ns1:getSupplementaryFilesResponse>
      <return>
        <request>
          <id>19117526</id>
          <source/>
          <email>?</email>
        </request>
        <pmid>19117526</pmid>
        <supplementaryFiles>cid:80750e00-3239-419b-9bc6-24e26b87e269@example.jaxws.sun.c
      </return>
    </ns1:getSupplementaryFilesResponse>
  </soapenv:Body>
</soapenv:Envelope>
```

Name	Content type	Size	Part	Type	ContentID
80750e00-3239-419...	application/zip	614599	80750e00-3239-419...	MIME	<80750e00-3239-4...

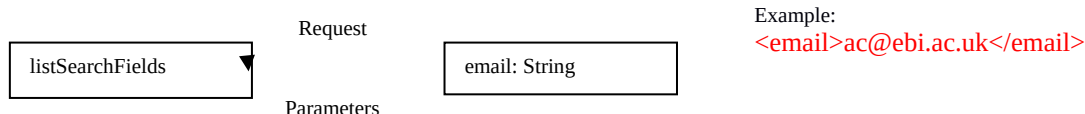
The screen shot below shows the contents of the attached zip file when opened:

Name	Type	Compressed size	Password ...	Size	Ratio
.DS_Store	DS_STORE File	1 KB	No	7 KB	98%
1471-213X-8-120-S1.jpeg	JPEG image	171 KB	No	286 KB	41%
1471-213X-8-120-S2.jpeg	JPEG image	359 KB	No	462 KB	23%
1471-213X-8-120-S3.mov	QuickTime Movie	3,499 KB	No	3,899 KB	11%
1471-213X-8-120-S4.mov	QuickTime Movie	3,602 KB	No	3,995 KB	10%
PMID19117526_citations.xml	XML Document	2 KB	No	4 KB	73%
PMID19117526_coremetadata.xml	XML Document	3 KB	No	17 KB	83%
PMID19117526_databaselinks.xml	XML Document	1 KB	No	6 KB	90%
PMID19117526_fulltext.xml	XML Document	24 KB	No	91 KB	75%
PMID19117526_references.xml	XML Document	4 KB	No	21 KB	82%
PMID19117526_textminedterms.xml	XML Document	1 KB	No	10 KB	92%

15. Method: listSearchFields

Also see the Appendix for a list of the field names.

Request/Input Parameters:



Method: listSearchFields	Retrieves all the field names of the Lucene index in the unified index file.		
Parameters	Required	Default	Values
email	N		Registers the user's email address with EBI. The user can then be contacted about Web Service related news.

Output:

Element	Data Type	Description	Example
term	String	The indexed field name	ABSTRACT, ACCESSION_ID, AFF
datasets	String	The name of index	metadata

Field Name	In Unified Index
ABBR	Y
ABSTRACT	Y
ACCESSION_ID	Y
ACCESSION_TYPE	Y
ACK_FUND	Y
AFF	Y
APPENDIX	Y
ARXPR_PUBS	Y
AUTH	Y
AUTHORID	Y
AUTHORID_TYPE	Y
AUTH_CON	Y
AUTH_EXACT	Y
AUTH_FIRST	Y
AUTH_LAST	Y
AURG_LIST	Y
AUTH_MAN	Y
AUTH_MAN_ID	Y
BOOK_ID	Y
CHEBITERM	Y

CHEBI_PUBS	Y
CHEM	Y
CHEMBL_PUBS	Y
CITED	Y
CITES	Y
COM_INT	Y
CONCL	Y
CRD_LINKS	Y
CREATION_DATE	Y
CREATION_DATE_NUM	Y
DISCUSS	Y
DISEASE	Y
DOI	Y
ED	Y
EFO	Y
EFO_ID	Y
EMBARGO_DATE	Y
EMBARGOED_MAN	Y
EMBL_PUBS	Y
EPMC_AUTH_MAN	Y
EUROFIR_KW	Y
EXT_ID	Y
E_PDATE	Y
FIG	Y
FIRST_PDATE	Y
FULLTEXT_SITE	Y
GENE_PROTEIN	Y
GOTERM	Y
GRANT_AGENCY	Y
GRANT_ID	Y
HAS_ABSTRACT	Y
HAS_ARXPR	Y
HAS_BOOK	Y
HAS_CHEBI	Y
HAS_CHEMBL	Y
HAS_CRD	Y
HAS_DATA	Y
HAS_DOI	Y
HAS_EMBL	Y
HAS_FULLTEXT	Y
HAS_FULLTEXTDATA	Y
HAS_INTACT	Y
HAS_INTERPRO	Y
HAB_LABSLINKS	Y
HAS_OMIM	Y
HAS_PDB	Y
HAS_PDF	Y

HAS_REFLIST	Y
HAS_SUPPL	Y
HAS_UNIPROT	Y
HAS_XREFS	Y
INTRO	Y
INVESTIGATOR	Y
IN_PMC	Y
IN_EPMC	Y
ISBN	Y
ISSN	Y
ISSUE	Y
IS_SCANNED	Y
JOURNAL	Y
JRNL_ISS_ID	Y
KEYWORD	Y
KW	Y
LABS_PUBS	Y
LANG	Y
METHODS	Y
NIH_AUTH_MAN	Y
OMIM_PUBS	Y
OPEN_ACCESS	Y
OTHER	Y
ORGANISM	Y
PAGE_INFO	Y
PDB_PUBS	Y
PDF	Y
PMCID	Y
PUBLIC_ACCESS	Y
PUBLISHER	Y
PUB_TYPE	Y
PUB_YEAR	Y
P_PDATE	Y
REF	Y
REFFED_BY	Y
RESULTS	Y
SB	Y
SORT_DATE	Y
SPAGE	Y
SRC	Y
SUPPL	Y
TABLE	Y
TITLE	Y
TS_LASTUPDATE	Y
UNIPROT_PUBS	Y
UPDATE_DATE_NUM	Y
VOLUME	Y

Appendix 1: Indexed Fields

The citations and full text articles can be searched using the indexed fields listed below. See the 'queryString' parameter for both the [profilePublications](#) and [searchPublications](#) methods. Click on the links below to see example search results. Note the use of double quotes for phrases.

Field	Description	Example
Core bibliographic		
EXT_ID:	Search for a publication by external ID: i.e. the ID assigned to a publication at repository level. Together with the publication's source, they form a unique id of the publication.	EXT_ID:10826746
PMCID:	Search for a publication by its PubMed Central ID, where applicable (i.e. available as full text)	PMCID:PMC1287967
TITLE:	Search for a term or terms in publication titles	TITLE:aspirin, TITLE:"protein knowledgebase"
ABSTRACT:	Search for a term or terms in publication abstracts	ABSTRACT:malaria, ABSTRACT:"chicken pox"
PUB_YEAR:	Search by year of publication in YYYY format; note syntax for range searching.	PUB_YEAR:2000, PUB_YEAR:[2000 TO 2001]
E_PDATE:	Electronic publication date, when an article was first published online.	E_PDATE:2013-12-15 E_PDATE:20070930 E_PDATE:[2000-12-18 TO 2014-12-30] E_PDATE:[20040101 TO 20140101]
FIRST_PDATE:	The date of first publication, whichever is first, electronic or print publication. Where a date is not fully available e.g. year only, an algorithm is applied to determine the value.	FIRST_PDATE:1995-02-01 FIRST_PDATE:20000101 FIRST_PDATE:[2000-10- 14 TO 2010-11-15] FIRST_PDATE:[20040101 TO 20140101]
P_PDATE:	Print publication date of journal issue, when an article appeared in print format.	P_PDATE:1982-10-01 P_PDATE:20140101 P_PDATE:[2000-12-18 TO

		2014-12-30] P_PDATE:[20031114 TO 20141115]
EMBARGO_DATE:	The date from which Europe PMC is permitted to provide access to the full text article. This date is held for articles received from March 2015.	EMBARGO_DATE:[2015-03-01 TO 2015-03-31] EMBARGO_DATE:2015-03-15
EMBARGOED_MAN:	Limit search results according to accessibility of the full text article regarding the EMBARGO_DATE. Assuming you request EMBARGOED_MAN:Y at the date 20160101 than this will be translated into EMBARGO_DATE:[20160101 TO *] and EMBARGOED_MAN:N to EMBARGO_DATE:[* TO 20160101].	EMBARGOED_MAN:Y, EMBARGOED_MAN:N
AUTH:	Search for a surname and (optionally) initial(s) in publication author lists	AUTH:einstein, AUTH:"Smith AB"
INVESTIGATOR:	Search for a publication by specifying an investigator	INVESTIGATOR:"Orlandini F"
AUTHORID_TYPE:	List publications that are associated with an ORCID	AUTHORID_TYPE:ORCID
AUTHORID:	List all the publications associated with a specified ORCID	AUTHORID:"0000-0002-1767-9318"
AFF:	Search for a term or terms in the author affiliation field	AFF:ebi, AFF:"university of cambridge"
JOURNAL:	Journal title – searchable either in full or abbreviated form	JOURNAL:"biology letters", JOURNAL:"biol lett"
ISSN:	Search for a journal by its ISSN; see the NCBI's list of journals and ISSNs in PubMed	ISSN:0028-0836
VOLUME:	Search for journal volumes, most useful in combination with fields: JOURNAL/PUB_YEAR/VOLUME/ISSUE/SPAGE	JOURNAL:"biology letters" VOLUME:10
ISSUE:	Search for journal issues, most useful in combination with fields: JOURNAL/PUB_YEAR/VOLUME/ISSUE/SPAGE	JOURNAL:"biology letters" VOLUME:10 ISSUE:2

SPAGE:	Search for articles which begin on page N of their journal, best used in combination with JOURNAL or other fields	SPAGE:25
LICENSE:	Search for content according to the assigned Creative Commons license (where provided). The values assigned by publishers have been grouped as follows: <u>CC BY</u> "cc", "cc by", "cc-by" <u>CC BY-SA</u> "cc", "cc by sa", "cc-by sa", "cc by-sa", "cc-by-sa" <u>CC BY-ND</u> "cc", "cc by nd", "cc-by nd", "cc by-nd", "cc-by-nd" <u>CC BY-NC</u> "cc", "cc by nc", "cc-by nc", "cc by-nc", "cc-by-nc" <u>CC BY-NC-ND</u> "cc", "cc by nc nd", "cc-by nc nd", "cc by-nc nd", "cc by nc-nd", "cc-by-nc nd", "cc by-nc-nd", "cc-by nc-nd", "cc-by-nc-nd" <u>CC BY-NC-SA</u> "cc", "cc by nc sa", "cc-by nc sa", "cc by-nc sa", "cc by nc-sa", "cc-by-nc sa", "cc by-nc-sa", "cc-by nc-sa", "cc-by-nc-sa"	LICENSE:"cc by sa" OR LICENSE:"cc-by sa" OR LICENSE:"cc by-sa" OR LICENSE:"cc-by-sa" ...returns the same count as: LICENSE:"CC BY-SA" All articles that have a Creative Commons license assigned can be found as follows: LICENSE:cc
EPMC_AUTH_MAN:	Identify manuscripts that have been submitted via the Europe PMC plus (formerly UKPMC+) Manuscript Submission System	EPMC_AUTH_MAN:y
NIH_AUTH_MAN:	Identify manuscripts that have been submitted the NIH Manuscript Submission System	NIH_AUTH_MAN:y
AUTH_MAN:	Search for an article that has been submitted via a Manuscript Submission System	AUTH_MAN:y
AUTH_MAN_I	Find an article by specifying an manuscript submission ID	AUTH_MAN_ID:EMS59581

D:		
----	--	--

Article metadata		
DISEASE:	Search for mined diseases	DISEASE:dysthymias
GENE_PROTEIN:	Search for records that have GENE_PROTEINS mined	GENE_PROTEIN:gng11
GOTERM:	Search for records that have GOTERM mined	GOTERM:apoptosis
IS_SCANNED:	Search for scanned/OCR content	IS_SCANNED:y
LANG:	Limit your search to publications written in language X	LANG:fre , LANG:eng
GRANT_AGENCY:	Limit your search by the funding agency which supported the research	GRANT_AGENCY:wellcome , GRANT_AGENCY:"medical research council"
GRANT_ID:	Limit your search by ID of the grant which funded the research	GRANT_ID:100229 , GRANT_ID:71672
KW:	Limit your search by keyword, including MeSH and other publisher-supplied terms	KW:galactosylceramides , KW:"recombinant proteins"
CHEM:	Limit your search by MeSH substance	CHEM:propantheline , CHEM:"protein kinases"
HAS_ABSTRACT:	Limit search results according to presence or absence of abstract	HAS_ABSTRACT:y , HAS_ABSTRACT:n
ORGANISM:	Search for mined organisms	ORGANISM:terebratulide
PUB_TYPE:	Limit your search by publication type	PUB_TYPE:review , PUB_TYPE:"retraction of publication"
Full text availability		
BODY:	Search for terms within the body of a fulltext article	BODY:PCR
DOI:	Search for publication by Digital Object Identifier (DOI)	DOI:10.1007/bf00197367 , DOI:10.1097/aln.0b013e3181b87edb
HAS_DOI:	Limit search results to those publications which do or do not have a DOI	HAS_DOI:y , HAS_DOI:n

IN_PMC:	Limit search results according to availability (or not) of fulltext article in PubMed Central	IN_PMC:y
IN_EPMC:	Limit search results according to availability (or not) of fulltext article in Europe PMC	IN_EPMC:y
HAS_PDF:	Limit search results according to availability (or not) of a PDF version of the fulltext article	HAS_PDF:y
Collection metadata		
SRC:	Search for articles from a particular repository; see Section 3 for the available data sources	SRC:ctx , SRC:hir
HAS_XREFS:	Limit search results to articles with cross-references to other databases; see Section 6 where the databases are listed for further details.	HAS_XREFS:y
HAS_CRD:	Limit search results according to presence or absence of links to related content. CRD is a database that provides comments about an article.	HAS_CRD:y
HAS_TM:	Limit search results to text-mined fulltext articles only (or not)	HAS_TM:y
HAS_REFLIST:	Limit search results to only those publications with a reference list (or not)	HAS_REFLIST:y
CREATION_DATE:	Search for publications by date of entry into the Europe PMC database, in YYYY-MM-DD format; note syntax for searching date range	CREATION_DATE:2010-11-11 , CREATION_DATE:[2010-11-11 TO 2010-12-11]
UPDATE_DATE:	Search for publications by date of update in the Europe PMC database in YYYY-MM-DD format; note syntax for searching date range	UPDATE_DATE:2011-11-11 , UPDATE_DATE:[2011-11-11 TO 2011-12-11]
OPEN_ACCESS:	Limit search results to articles that are Open Access	OPEN_ACCESS:y

HAS_LABSLINKS:	Lists articles that have links provided by 3 rd parties (using the Europe PMC External Links Service)	HAS_LABSLINKS:y
HAS_DATA:	Lists articles that have data-literature links with the tags 'related_data' or 'supporting_data'.	HAS_DATA:y
HAS_SUPPL:	List articles that have supplemental data associated with them e.g. spreadsheets, video files, etc.	HAS_SUPPL:y
LABS_PUBS:	List articles that have external links by provider ID	LABS_PUBS:1001
Content set filter		
SB:	Limit search results to EuroFIR subsets	SB:eurofir
Books		
ISBN:	Search for book by ISBN	ISBN:9780815340720
ED:	Search for book by editor	ED:jensen HAS_BOOK:y
PUBLISHER:	Search for book by publisher	PUBLISHER:"OUP Oxford" HAS_BOOK:y
HAS_BOOK:	List the full text books on the Europe PMC Bookshelf	HAS_BOOK:y
BOOK_ID:	Find a full text book on the Europe PMC Bookshelf by specifying its 'NBK' number	BOOK_ID:NBK27326
Database cross references		
ARXPR_PUBS:	Show publications with links to given ArrayExpress ID	ARXPR_PUBS:E-GEOD-22481
UNIPROT_PUBS:	Show publications with links to given UniProt ID	UNIPROT_PUBS:q1rdg3
EMBL_PUBS:	Show publications with links to given EMBL ID	EMBL_PUBS:KJ634683
PDB_PUBS:	Show publications with links to given PDBe ID	PDB_PUBS:2w3z

INTACT_PUBS:	Show publications with links to given IntAct ID	INTACT_PUBS:ebi-493556
INTERPRO_PUBS:	Show publications with links to given InterPro ID	INTERPRO_PUBS:ipr013998
CHEBI_PUBS:	Show publications with links to given ChEBI ID	CHEBI_PUBS:62806
CHEBITERM:	Search for mined chemical terms	CHEBITERM:dihydrocapsaicin
CRD_LINKS:	Show publications with links to given related content ID	CRD_LINKS:22001008219
HAS_ARXPR:	Limit search results to publications with(out) links to the ArrayExpress catalog	HAS_ARXPR:y
HAS_UNIPROT:	Limit search results to publications with(out) links to the UniProt catalog	HAS_UNIPROT:y
HAS_EMBL:	Limit search results to publications with(out) links to the EMBL database	HAS_EMBL:y
HAS_PDB:	Limit search results to publications with(out) links to the PDBe database	HAS_PDB:y
HAS_INTACT:	Limit search results to publications with(out) links to the IntAct database	HAS_INTACT:y
HAS_INTERPRO:	Limit search results to publications with(out) links to the InterPro database	HAS_INTERPRO:y
HAS_CHEBI:	Limit search results to publications with(out) links to the ChEBI dictionary	HAS_CHEBI:y
HAS_CHEMBL:	Limit search results to publications with(out) links to ChEMBL	HAS_CHEMBL:y
HAS_OMIM:	Limit search results to publications with(out) links to OMIM	HAS_OMIM:y

CITES:	Search for publications that cite a given article; article to be specified in the format, <i>ID_source</i> (see Section 3 for sources)	CITES:8521067_med , CITES:IND43783977_agr
CITED:	Search for publications that have been cited N times	CITED:100
REFFED_BY:	Search for publications that cite the specified article; format, <i>ID_source</i> (see Section 3 for sources)	REFFED_BY:9497246_med
Database citations		
ACCESSION_ID:	Finds articles containing the specified accession number	ACCESSION_ID:A12360
ACCESSION_TYPE:	Find articles that cite ArrayExpress records	ACCESSION_TYPE:arrayexpress
	Find articles that cite BioProject records in the European Nucleotide Archive	ACCESSION_TYPE:bioproject
	Find articles that cite BioSamples records	ACCESSION_TYPE:biosample
	Find articles that cite data Document Object Identifiers	ACCESSION_TYPE:doi
	Find articles that cite European Genome-phenome Archive records	ACCESSION_TYPE:ega
	Find articles that cite EM resources in the Electron Microscopy data Bank (EMDB) records	ACCESSION_TYPE:emdb
	Find articles that cite Ensembl records	ACCESSION_TYPE:ensembl
	Find articles that cite EudraCT records from the EU Clinical Trials Register	ACCESSION_TYPE:eudract
	Find articles that cite European Nucleotide Archive records	ACCESSION_TYPE:gen
	Find articles that cite Gene Ontology records	ACCESSION_TYPE:go

	Find articles that cite InterPro records	ACCESSION_TYPE:interpro
	Find articles that cite NCT clinical studies records from the US NIH ClinicalTrials.gov registry	ACCESSION_TYPE:nct
	Find articles that cite OMIM records	ACCESSION_TYPE:omim
	Find articles that cite Protein Data Bank in Europe (PDB)records	ACCESSION_TYPE:pdb
	Find articles that cite protein families (Pfam) records	ACCESSION_TYPE:pfam
	Find articles that cite ProteomeXchange records	ACCESSION_TYPE:pxd
	Find articles that cite Reference Sequence Database (RefSeq) records	ACCESSION_TYPE:refseq
	Find articles that cite records in the dbSNP Short Genetic Variations database (RefSNP)	ACCESSION_TYPE:refsnp
	Find articles that cite the resource of protein sequence and functional information (UniProt)	ACCESSION_TYPE:sprot
	Find articles that cite TreeFam (database of animal gene trees) records	ACCESSION_TYPE:treefam

Section-level search		
ABBR:	Find articles with word “mRNA” in the Abbreviations section	ABBR:mRNA
ACK_FUND:	Find articles with word “ERC” in the Acknowledgements & Funding section	ACK_FUND:ERC
APPENDIX:	Find articles with word “ethics” in the Appendix section	APPENDIX:ethics
AUTH_CON:	Find articles with phrase “Smith” in the Author Contribution section	AUTH_CON:“Smith”
CASE:	Find articles with word “leukemia” in the Case Study section	CASE:leukemia
COMP_INT:	Find articles with phrase “no conflict” in the Competing Interest section	COMP_INT:“no conflict”
CONCL:	Find articles with word “osteoporosis” in the Conclusion section	CONCL:osteoporosis
DISCUSS:	Find articles with word “cardiovascular” in the Discussion section	DISCUSS:cardiovascular
FIG:	Find articles with phrase “in vitro” in the Figures section	FIG:“in vitro”
INTRO:	Find articles with phrase “protein interactions” in the Introduction & Background section	INTRO:“protein interactions”
KEYWORD:	Find articles with word “isoform” in the Keywords section	KEYWORD:isoform
METHODS:	Find articles with phrase “yeast two-hybrid” in the Materials & Methods section	METHODS:“yeast two-hybrid”
OTHER:	Find articles with phrase “transgenic mice” in the Others section	OTHER:“transgenic mice”
REF:	Find articles with word “COX2” in the References section	REF:COX2
RESULTS:	Find articles with phrase “in vivo” in the Results section	RESULTS:“in vivo”

SUPPL:	Find articles with word “supplementary” in the Supplementary Information section, indicating where supplementary data might be available.	SUPPL:supplementary
TABLE:	Find articles with word “comparison” in the Tables section	TABLE:comparison