

## VIRTA Publication Information Service



*CSC – Finnish research, education and public administration ICT knowledge centre*

## Background: Publication information collection in Finland



- Ministry of Education and Culture collects bibliographic information on scientific publications annually from
  - 14 universities and 5 university hospital districts (since 2011)
  - 23 universities of applied sciences (since 2012)
  - 12 state research institutes (gradually since 2014)
- In total, ~60 000 publications per year  
= books, journal articles, conference papers, non-scholarly publications
- The data are publicly available
  - Statistical data: [www.vipunen.fi](http://www.vipunen.fi)
  - Bibliographic data: [www.juuli.fi](http://www.juuli.fi)

## Background: Use of publication information in Finland

1. Monitoring research and criteria in the performance-based funding model of universities
  - 13 % of core funding is allocated via publication points
  
2. Use of information through various services and other processes
  - VIRTA Publication Information Service implemented in 2016
  - All publication information available to other services via APIs
  - Publication information automatically to research funding applications etc.
  - Implementation in Academy of Finland's online service in 2017
  
3. Next step: VIRTA *Research* Information Service
  - = Extension of VIRTA to other research output (projects, researchers, research data, infrastructures etc.)

## VIRTA Publication Information Service

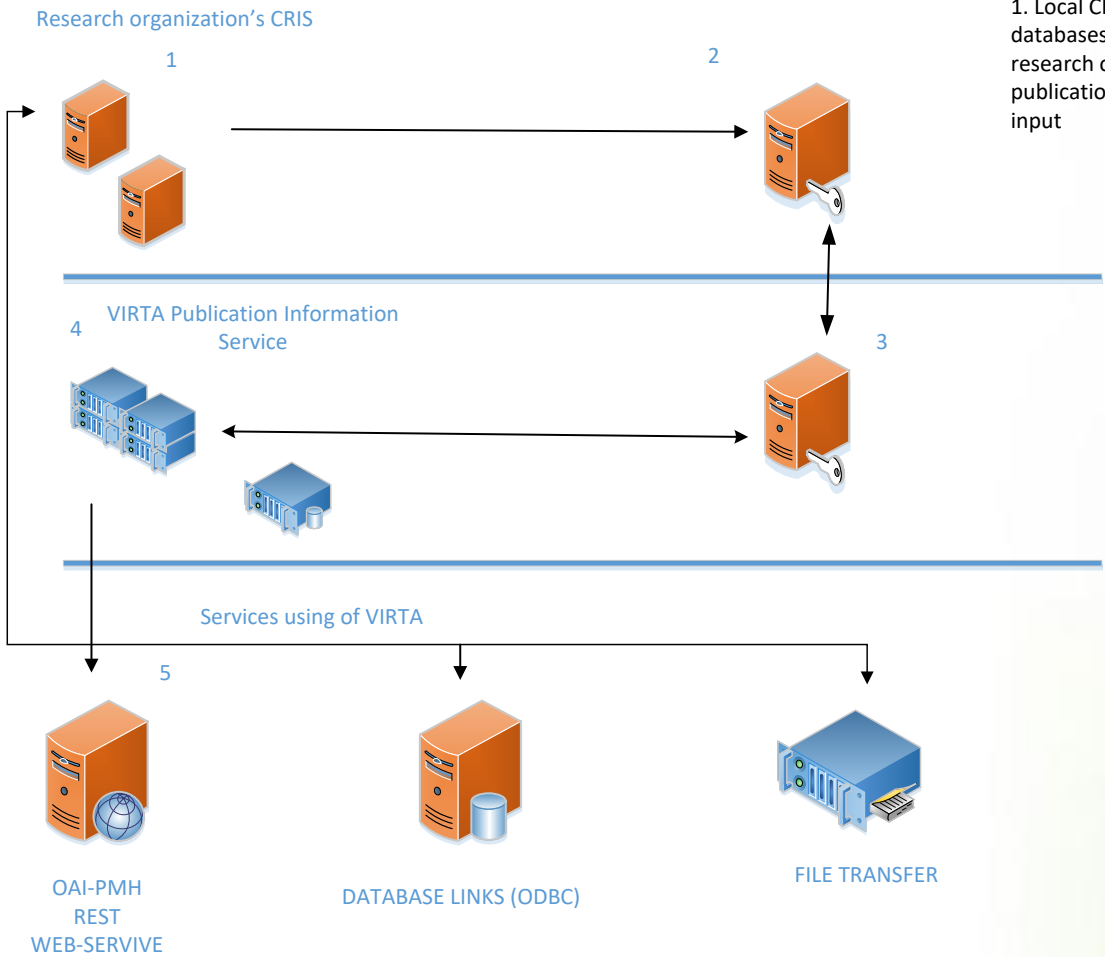


- The Finnish research organizations use VIRTA to store a copy of publication information in their institutional CRISes or other publication databases
- VIRTA publication information service is a data warehouse, “a data hub”, making publication information available for other services
- The information can be transferred to research funders, publication or research data repositories, research infrastructure services, or other services used by researchers
- VIRTA provides the higher education institutions and other research organizations with up-to-date, comparative data on their own publications, as well as those of other institutions.
- The concept originates from *VIRTA higher education achievement register* (compiling higher education degrees, study records and grades) implemented in 2012

# VIRTA Publication Information Service in a nutshell

<b>Data sources</b>	Master data in local CRISes or publication databases of HEIs, university hospitals, state research institutes
<b>Data format</b>	XML files (CSV converter provided for small organizations)
<b>Data contents</b>	The data must include required fields and fulfil certain technical criteria.
<b>Data transfer</b>	From organizations via a secure and certified connection by using SFTP protocol and SSH authentication keys.
<b>Updates</b>	New publications and corrections in local systems are updated automatically to VIRTA e.g. once a day. The frequency depends on the organizations, minimum being once a year.
<b>Temporal coverage</b>	All data from previous years to present can be transferred. Statistics compiled once a year.
<b>Data validation</b>	Duplicates, faults as well as inter-organizational co-publications identified automatically and real time. Error reports available for research organizations in an online service.
<b>Data use and availability</b>	All metadata are published real time in JUULI portal: <a href="http://www.juuli.fi">www.juuli.fi</a> . Tiedot hyödynnettävissä muihin järjestelmiin ja palveluihin REST- ja OAI-PMH – rajapintojen kautta.

# VIRTA system architecture



1. Local CRISes or publication databases of universities and other research organizations in which the publication information is primarily input

2-3. Servers and services used for data transfer. Publication information is transferred by using encrypted SFTP-protocol.

4. VIRTA Publication Information Service compiles XML files from various organizations

5. Publication information in VIRTA is used securely by multiple methods and protocols.

## Data contents of VIRTA publication information service

- Research organizations need to provide certain required data fields for each publication
- Information can be supplemented with optional data fields
- The data must fulfil certain technical criteria (correct form of ISSN, ISBN, etc).
- All the definitions and requirements are introduced in detail in the Ministry's Publication data collection Guide to ensure uniformity and quality of data
- VIRTA enables the research organizations to present, correct and complete the data online.

# Data contents of VIRTA publication information service

Required in the Ministry's publication data collection				
Required in data transfer <sup>a</sup>	Optional in data transfer <sup>b</sup>	Optional in Ministry's data collection	Other information not included in Ministry's data collection	Information generated automatically in VIRTA
Organization ID Publication ID in organization Publication year Title Authors Publication type Authors in organization ISBN* ISSN* Conference title*	Number of authors Scientific field (1-6) International (yes/no) International co-publication (yes/no) Co-publication with a company Open Access Publisher*	Volume Number Pages Article number Journal/series title Country Unit or department Place of publishing Host publication's title Host publication's editors Language DOI Permanent address Source ID Keywords ORCID	Project ID Funder ID Status of the publication	Reporting year Publication's VIRTA ID Co-publication's VIRTA-ID Publication Forum ID Publication Forum ranking
*) Requirements depend on publication type				



The data are transferred from research organizations as XML files defined in [VIRTA XML schema](#)

```
<?xml version="1.0" encoding="UTF-8"?>
<Julkaisut xsi:schemaLocation="urn:mace:funet.fi:julkaisut/2015/03/01" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns="urn:mace:funet.fi:julkaisut/2015/03/01">
  <Julkaisu>
    <OrganisaatioTunnus>99999</OrganisaatioTunnus>
    <IlmoitusVuosi>2014</IlmoitusVuosi>
    <JulkaisunTilaKoodi>1</JulkaisunTilaKoodi>
    <JulkaisunOrgTunnus>3123213</JulkaisunOrgTunnus>
    <JulkaisunOrgYksikot>
      <YksikkoKoodi>E01</YksikkoKoodi>
      <YksikkoKoodi>E02</YksikkoKoodi>
      <YksikkoKoodi>E03</YksikkoKoodi>
      <YksikkoKoodi>E04</YksikkoKoodi>
      <YksikkoKoodi>E05</YksikkoKoodi>
    </JulkaisunOrgYksikot>
    <JulkaisuVuosi>2014</JulkaisuVuosi>
    <JulkaisunNimi>19q13 amplification is associated with high grade and stage in pancreatic cancer</JulkaisunNimi>
    <TekijätiedotTeksti>Kuuselo, Riina; Simon, Ronald; Karhu, Ritva; Tennstedt, Pierre; Marx, Andreas H.; Izbicki, Jakob R.; Yekebas, Emre; Sauter, Guido; Kallioniemi, Anne</TekijätiedotTeksti>
    < SivunumeroTeksti>569-575</ SivunumeroTeksti>
    < Artikkelinumero>III</ Artikkelinumero>
    < JufoTunnus>56446</ JufoTunnus>
    < JufoLuokkaKoodi>1</ JufoLuokkaKoodi>
    < JulkaisumaaKoodi>826</ JulkaisumaaKoodi>
    < LehdenNimi>Genes, Chromosomes & Cancer</ LehdenNimi>
    < ISSN>1045-2257</ ISSN>
    < VolyymiTeksti>49</ VolyymiTeksti>
    < LehdenNumeroTeksti>6</ LehdenNumeroTeksti>
    < JulkaisutyyppiKoodi>A1</ JulkaisutyyppiKoodi>
  </Julkaisu>
  <TieteenalaKoodit>
    <TieteenalaKoodi JNro="1">3111</TieteenalaKoodi>
    <TieteenalaKoodi JNro="2">3122</TieteenalaKoodi>
  </TieteenalaKoodit>
  <KoulutusalaKoodit>
    <KoulutusalaKoodi JNro="1">90</KoulutusalaKoodi>
  </KoulutusalaKoodit>
  <YhteisjulkaisuKVKytin>1</YhteisjulkaisuKVKytin>
  <YhteisjulkaisuSHPKytin>0</YhteisjulkaisuSHPKytin>
  <YhteisjulkaisuTutkimuslaitosKytin>0</YhteisjulkaisuTutkimuslaitosKytin>
  <YhteisjulkaisuMuuKytin>0</YhteisjulkaisuMuuKytin>
  <JulkaisunKansainvalisyysKytin>1</JulkaisunKansainvalisyysKytin>
  <JulkaisunKieliKoodi>en</JulkaisunKieliKoodi>
  <AvoinSaatavuusKoodi>2</AvoinSaatavuusKoodi>
  <EVOjulkaisuKytin>1</EVOjulkaisuKytin>
  <DOI>10.1002/gcc.20767</DOI>
  <PysyvaOsoiteTeksti>http://onlinelibrary.wiley.com/doi/10.1002/gcc.20767/abstract</PysyvaOsoiteTeksti>
  <Tekijat>
```

## Transfer of data

- The XML files are transferred to VIRTA from research organizations via a secure and certified connection by using Secure File Transfer Protocol (SFTP) and SSH authentication keys
- Access to the server is limited by firewalls and allowed only to designated IP addresses
- Either the imported file can contain only new or edited publications (recommended), or alternatively, the organization can transfer all its data which is then read over the previous data
- When a publication is edited it is identified on the basis of Publication ID unique per publication.

## Validation of data

- The missing and erroneous information and duplicate publications are detected automatically
- Each organization can check the publications' status and error reports in Extra-Vipunen online service
- Co-publications from two or more organizations are identified and connected automatically
- Publication Forum ID and ranking ([www.julkaisufoorumi.fi](http://www.julkaisufoorumi.fi)) are detected automatically for each scientific publication

# Status reports in Extra Vipunen online service



Suodattimet <

- Organisaatio
- Julkaisuvuosi
  - 2010
  - 2013
  - 2014
  - 2015
  - 2016
  - 2017
- Ilmoitusvuosi
- Julkaisutyyppi
- Julkaisun nimi
- Luontipvm
- Päivityspvm
- JulkaisunTunnus
- OrganisaationJulkaisuT...

## Julkaisutietopalvelun julkaisut

Julkaisujen määrä	Vuosi		
	2014	2015	2016 Yhteensä
Aalto yliopisto	139	3944	4083
Arcada - Nylands svenska yrkeshögskola		116	116
Centria-ammattikorkeakoulu		137	137
Diakonia-ammattikorkeakoulu		204	204
Geologian tutkimuskeskus			56
Haaga-Helia ammattikorkeakoulu	3	252	255
Helsingin yliopisto	1144	11044	12188
Humanistinen ammattikorkeakoulu		297	297
Hämeen ammattikorkeakoulu		267	267
Itä-Suomen yliopisto	151	3186	3337
Jyväskylän ammattikorkeakoulu	38	369	407
Jyväskylän yliopisto	118	3094	3212
Kajaanin ammattikorkeakoulu	1	101	102
Karelia-ammattikorkeakoulu		195	195
Kuopion yliopistollisen sairaalan erityisvastuualue		699	699
Kymenlaakson ammattikorkeakoulu	2	363	365
Lahden ammattikorkeakoulu	3	293	1
Lapin ammattikorkeakoulu		392	392
Lapin yliopisto	33	676	709
Lappeenrannan tekn. yliopisto	38	998	1036
Laurea-ammattikorkeakoulu	1	308	309
Luonnonvarakeskus		1556	1556
Lääkealan turvallisuus- ja kehittämiskeskus		137	137
Maanmittauslaitos		161	161
Metropolia ammattikorkeakoulu	2	292	294

## Näytä >



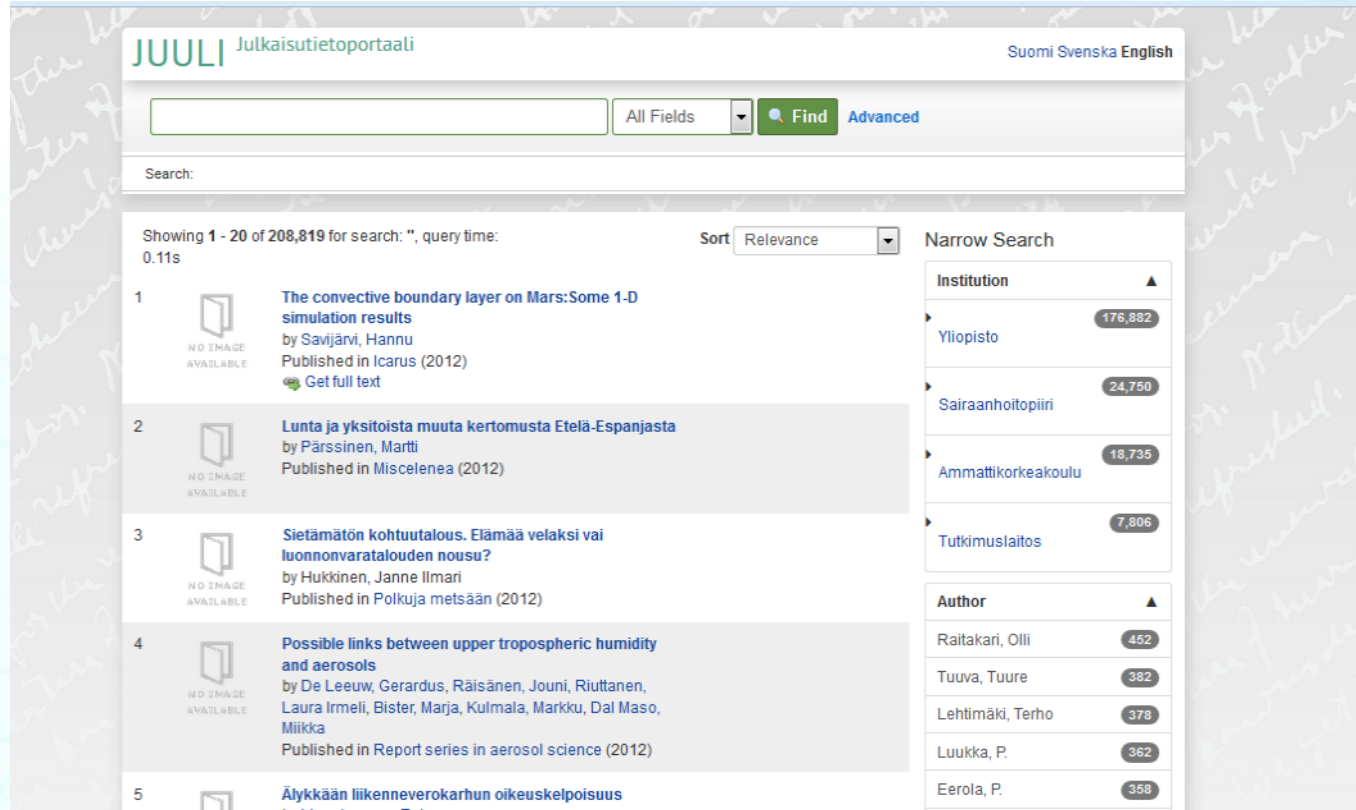
Julkaisutietopalvelu



Graafi

Tilanneraportti

# JUULI publication portal








**JUULI** Julkaisutietoportaaali Suomi Svenska English

Search:  All Fields  [Advanced](#)

Showing 1 - 20 of 208,819 for search: "", query time: 0.11s Sort: Relevance

Narrow Search	
<b>Institution</b>	
Yliopisto	176,882
Sairaanhoidopiiri	24,750
Ammattikorkeakoulu	18,735
Tutkimuslaitos	7,806
<b>Author</b>	
Raitakari, Olli	462
Tuuva, Tuure	382
Lehtimäki, Terho	378
Luukka, P.	362
Eerola, P.	358

1	 NO IMAGE AVAILABLE	<b>The convective boundary layer on Mars: Some 1-D simulation results</b> by Savijärvi, Hannu Published in Icarus (2012) <a href="#">Get full text</a>
2	 NO IMAGE AVAILABLE	<b>Lunta ja yksitoista muuta kertomusta Etelä-Espanjasta</b> by Pärssinen, Martti Published in Miscelanea (2012)
3	 NO IMAGE AVAILABLE	<b>Sietämätön kohtuutus. Elämää velaksi vai luonnonvaratalouden nousu?</b> by Hukkinen, Janne Ilmari Published in Polkuja metsään (2012)
4	 NO IMAGE AVAILABLE	<b>Possible links between upper tropospheric humidity and aerosols</b> by De Leeuw, Gerardus, Räisänen, Jouni, Riuttanen, Laura Irmeli, Bister, Marja, Kulmala, Markku, Dal Maso, Miikka Published in Report series in aerosol science (2012)
5	 NO IMAGE AVAILABLE	<b>Älykkään liikenneverokarhun oikeuskelpoisuus</b> by ...

JUULI ([www.juuli.fi](http://www.juuli.fi)) displays bibliographic information of all publications in VIRTAs

The publications in VIRTAs are updated to JUULI daily

Opetushallinnon tilastopalvelu > Korkeakoulutus ja TK-toiminta - Julkaisut - Organisaatio.dsb

### Julkaisut vuonna 2015

Julkaisu	A1 Alkuperäisartikkeli tieteellisessä aikakauslehdessä	A2 Katsausartikkeli tieteellisessä aikakauslehdessä	A3 Kirjan tai muun kokoomateoksen osa	A4 Artikkelit konferenssijulkaisuissa	B1 Kirjoitus tieteellisessä aikakauslehdessä
Ammattikorkeakoulu	358	17	74	214	95
Tutkimuslaitos	2 053	117	123	163	144
Yliopisto	17 829	921	2 712	2 934	2 438
Yliopistollisen sairaalan erityisvastuualue	4 358	454	41	5	470
<b>Yhteensä</b>	<b>24 598</b>	<b>1 509</b>	<b>2 950</b>	<b>3 316</b>	<b>3 147</b>

Statistical analyses on publishing activity are openly available in the Vipunen portal ([www.vipunen.fi](http://www.vipunen.fi)) that contains not only data on publications but also other information related to the operations of research universities and universities of applied sciences.

## Use of VIRTAs publication information

- Publication information in VIRTAs are usable for other systems or services through a REST API, OAI-PMH or file transfers .
- VIRTAs will be used in the online service of the major Finnish research funding organization, Academy of Finland from 2017 on
  - No longer need for manual input of publications when reporting the project's output while the researcher can pick up publication data automatically from VIRTAs.
  - The project information (project and funder ID) related to the publications are returned into VIRTAs
- The publication data can also be used in researchers' CVs in research funding applications etc.
- The near future goal is to extend the concept to other funding organizations as well as other services where researchers need to input data on their publications, such as research data services, research infrastructure services



More information:

[virta-julkaisut@postit.csc.fi](mailto:virta-julkaisut@postit.csc.fi)