

Die Verwendung von Autorenidentifikatoren in wissenschaftlichen Repositorien

Ansätze, konkrete Umsetzungen und Herausforderungen

Vortrag beim 6. Bibliothekskongress 2016 in Leipzig

Friedrich Summann

ORCID ID

 orcid.org/0000-0002-6297-3348

▼ Employment (1)

Bielefeld University: Bielefeld, North Rhine-Westphalia, Germany

Head LibTec (Library)

Source: Friedrich Summann

Created: 2015-06-01

Abstract

Um eine Analyse der globalen Verwendung von Autoren-Identifikatoren in der Repository-Landschaft durchzuführen, wurde auf Basis der BASE-Infrastruktur der UB Bielefeld mit Hilfe von Email-Anfragen unter Vertretern der globalen Repository-Community und skript-basierter Metadaten-Analyse (auf OAI-PMH-Basis) eine breit angelegte Untersuchung durchgeführt. Mittlerweile ist die Integration von ORCID-Identifiern in den technischen Implementierungen vorbereitet und man kann sie auf verschiedenen Ebenen (Datenbankstruktur, Import, Benutzeroberfläche, Schnittstellen) in der Praxis zunehmend nachweisen. Auch wenn die Umsetzung noch nicht flächendeckend erfolgt, werden entsprechende Implementierungen zunehmend eingesetzt. Vergleichsweise große ORCID-Zahlen finden sich bei DataCite und Europe PMC, bei den Repository-Systemen fallen DIVA aus Schweden und OJS durch zahlreiche Installationen mit ORCID-Nachweisen auf. Auch einzelne DSpace-, Fedora-, Invenio-Repositorien und Lösungen wie Librecat (Univ. Bielefeld) und Brocade (Univ. Antwerpen) bedienen bereits durchgängig alle Ebenen und bei fast allen alternativen Plattformen (EPrints, Opus, Digital Commons) sind Teilbereiche realisiert.

Einführung und Methodik

Schon beim DINI-EPub-Workshop 2013 zum Thema Zukunft der Repositorien¹ wurde das Thema Autorenidentifizierung behandelt und der folgende Screenshot zeigt die aktualisierte Fassung einer Autorensseite aus dem Bielefelder Repository/Publicationsmanagementsystem, die schon früher als Beispiel für die Verwendung von Autorenidentifiern gezeigt wurde. Aktuell sind allerdings einige Erscheinungsformen neu hinzugekommen, wie der github-Account und der Link auf die ORCID-Seite des Autors.

The screenshot shows the header of the Bielefelder Repository website with the text 'Universität Bielefeld' and 'PUB – Publications at Bielefeld University'. Below this is the author profile for Christian Pietsch, which includes a portrait, his name, email address 'christian.pietsch [at] uni-bielefeld.de', and a list of identifiers: PEVZ-ID (7346775), Author Claim (ppi3), GitHub (pietsch), Google Scholar (Ens-DwwAAAAJ), ORCID (0000-0001-8778-1273), Researcher ID (J-1546-2012), and Scopus Author ID (55619320748). A 'Publications' tab is visible, showing 13 publications. One publication is listed: Griffiths, S., Eyssel, F. A., Philippsen, A., Pietsch, C., & Wachsmuth, S. (2015). Perception of... There is also a search bar and a 'Marked Publication(s)' indicator.

Die folgende Liste zeigt aktuell verwendete Autorenidentifizierung aus den verschiedenen Bereichen, in denen sie eine Rolle spielen: Bibliotheken, Informationsanbieter, Fachcommunities. Nicht vergessen werden dürfen dabei auch die lokal verwendeten Identifizierung.

- * PND (GND), ISNI, VIAF
- * Researcher ID, Scopus ID
- * arXiv ID
- * ResearchGate, Google Scholar
- * LinkedIn, GitHub
- * AuthorClaim
- * ORCID ID
- * Lokale Ids

Um das Ziel umzusetzen, eine Übersicht über den Stand der Vorbereitung und Verwendung von Autorenidentifiern im Umfeld der Repositorien zu gewinnen, wurde auf verschiedenen Ebenen gearbeitet. Grundlage war die in Bielefeld aufgebaute BASE-Infrastruktur, sowohl in

¹ Summann F. Institutionelle Repositorien an Hochschulen - oder wie geht es weiter? 102. Deutscher Bibliothekartag (= 5. Bibliothekskongress), Leipzig 2013. [https://www.opus-bayern.de/bib-info/frontdoor.php?source_opus=1474]

organisatorischer (als Bestandteil des globalen Repository-Netzwerkes) als auch in technischer Hinsicht. Im BASE-System liegen Kontaktinformationen und Metadaten zu mehr als 4000 Repository-Systemen vor. Gleichzeitig stehen die eigentlichen Metadaten der erfassten Repositorien mit mehr als 200 Mill. Sätzen zur Verfügung, in Verbindung mit einer Infrastruktur, die für tiefgehende Analyse-Zwecke der Inhalte genutzt werden kann.

Konkret wurde mit drei Ansätzen gearbeitet:

- Versendung von Emails an Repository-Community-Vertreter mit der Bitte um Darstellung des aktuellen Stands der ORCID-Unterstützung (DSpace, DSpace IRIS, EPrints, OPUS, DigitalCommons, OJS, DataCite, Librecat, invenio)
- Skript-basierte Analyse von vorliegenden Harvest-Ergebnissen
 - Dazu wurde speziell ein aktueller, größerer Ausschnitt von DataCite (1,5 Mill. Sätze im DataCite-Format) geholt und analysiert
- Stichproben-Analyse aller BASE-Repositories
 - Per Skript wurden (in den ausgewählten Formaten MODS, MARC und XMetaDissPlus) OAI-PMH-Aufrufe bei allen aktiven Repositorien gestartet und die Antwortdateien analysiert.

Repository-Techniken

Zunächst sollen die verschiedenen Aspekte bzw. Stufen der Integration von Autorenidentifiern im Repository-Umfeld untersucht werden. Letztlich hat das Email-Feedback gezeigt, dass die Integration von Autorenidentifiern auf verschiedenen Stufen erfolgt, die teilweise aufeinander aufbauen. Diese sind im Einzelnen:

- Datenbankstruktur
- Import (automat., Autoren-GUI)
- Anzeige Benutzeroberfläche
- Schnittstellen
 - Metadatenexport (OAI-PMH)
 - Personendatenimport (ORCID-API)

Im BASE-Kontext wurde das erste Mal vor ca. einem Jahr über die Integration von ORCID-IDs gesprochen und der Anlass war das Auftauchen von entsprechenden Informationen in der durch die OAI-PMH-Schnittstelle gelieferten Metadaten von SCOAP3, speziell im MARC-Format. Um effizient vorgehen zu können, muss insbesondere genauer untersucht werden, welche Formate in welcher Weise und in welchem Umfang die relevante Information bereitstellen.

Beispiele und Lösungen

Martin Fenner hat 2014, ebenfalls auf dem Bibliothekartag, eine Übersicht² über den damaligen Stand der ORCID-Integration an Universitäten gegeben und dabei auch den Stand in den Repositorien erwähnt und dabei „diskutiert, warum Autorenidentifikation mittels ORCID in Deutschland noch wenig verbreitet ist“

Heute, zwei Jahre später, ist diese Aussage auch in Bezug auf die globale Anwendung in Repositorien weiterhin gültig. Es gibt erweiterte Lösungsansätze und zunehmende Implementierungen, die die oben genannten Stufen der Integration ausbauen. Auf der anderen Seite gibt es im Bereich der Repositorien durchaus einen ausgeprägten Verzögerungseffekt, bevor sich neue Techniken über die eingesetzte Plattformsoftware und deren Updates verbreiten.

Die folgende Tabelle zeigt gängige Plattformen und den Grad der Autorenidentifizier-Unterstützung

System	Datenbank- struktur	Import	Benutzer- Oberfläche	OAI-PMH nachgewiesen
DSpace	x	x	x	(x)
DSpace IRIS	x	x	x	
EPrints	x	x	x	
Invenio	x	x	x	X
OJS	x	x	x	x
Digital Commons	x	x	x	
DIVA	x	x	x	x
DataCite	x	x	x	x
Librecat	x	x	x	x

DataCite-Auswertung

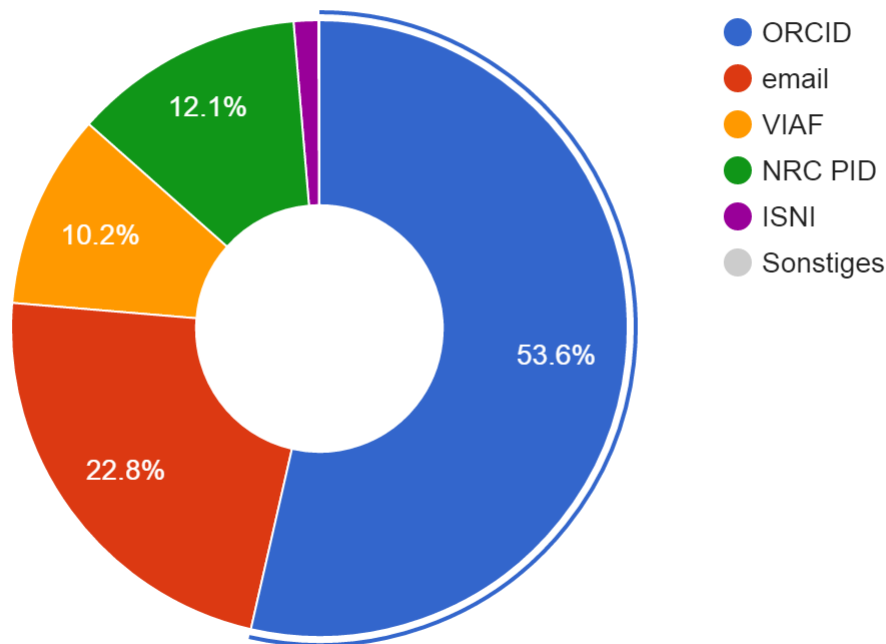
Bei Europe PMC und DataCite werden ORCID-IDs in den Metadaten via OAI-PMH bereits heute in großem Umfang bereitgestellt. Daher wurden bei DataCite aktuelle 1,5 Mill. Datensätze erfaßt, bei denen im begrenzten Umfang im entsprechenden Attribut verschiedenartige Autorenidentifizier angeliefert wurden, dokumentiert in der folgenden Liste (die Benennungen sind die Originalbezeichner)

- VIAF – 1652
- ISNI – 208
- ORCID – 8664
- email – 3681
- NRC PID – 1960
- personIdUniBi – 9

² Fenner, Martin ORCID an Universitäten: Statusbericht mit internationaler Perspektive 103. Deutscher Bibliothekartag Bremen 2014
[<http://nbn-resolving.org/urn:nbn:de:0290-opus-15562>]

Die folgende Graphik zeigt die daraus resultierende prozentuale Verteilung und damit wird auch klar, dass zumindest im Datacite-Umfeld die ORCID-ID die größte Verbreitung hat. (Segment mit datestamp-Parameter ,from=2016-01-01', basierend auf OAI-PMH Download, Stand 8.3.2016)

Autorenidentifikatoren - Verteilung bei DataCite



Resultate der automatischen Analyse

Das oben bereits vorgestellte Analyse-Skript geht im Detail in folgender Weise vor:

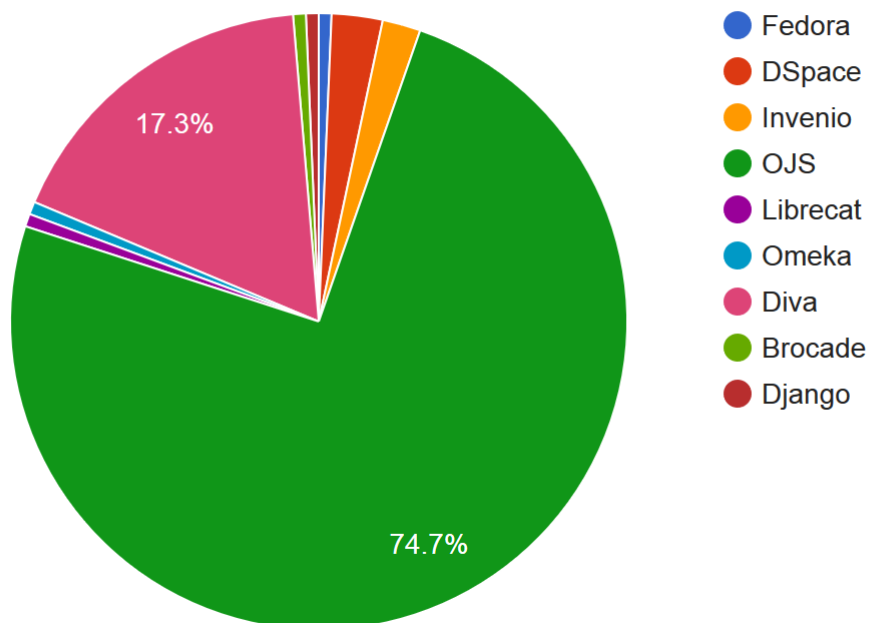
Es wurde bei allen aktiven in BASE notierten Repositories eine OAI-PMH-Anfrage in den dafür definierten Formate (das waren zum Zeitpunkt März 2016 MODS bzw. swepubmods, MARC, XMetaDissPlus) gestartet. Ausgewählt wurden solche Formate, bei denen eine Anlieferung von ORCID-IDs nachgewiesen war und daher ein Vorkommen zu erwarten war. Als Startparameter für den betrachteten Zeitraum wurde der 1.1.2016 gewählt, um möglichst aktuelle Datensätze erfassen zu können. In den gelieferten Datensätzen wurde per Pattern Matching nach ORCID-IDs und ResearcherID gesucht. Nachgewiesen werden konnten aber nur ORCID-IDs und das in 156 von über 5500 betrachteten Repositories, bei denen allerdings nur in 2489 Fällen die betrachteten Formate unterstützt wurden.

Via OAI-PMH werden als Ergebnis der Analyse ORCID-Metadaten von den folgenden Plattformsystemen bereitgestellt (Stand 8.3.2016):

System	Verbreitung	Anzahl Installationen	Anzahl ORCID-IDs
DIVA	(flächendeckend)	26	869
OJS	(zahlreiche Installationen)	112	727
Invenio	(u.a. SCOAP3)	3	466
DSpace	(einzelne Installationen)	4	126
Librecat	(UB Bielefeld)	1	23
Fedora	(einzelne Version)	1	498
Brocade	(Univ. Antwerpen)	1	1
Django	(Univ. Bochum)	1	2
Omeka	(Univ. Otago)	1	1

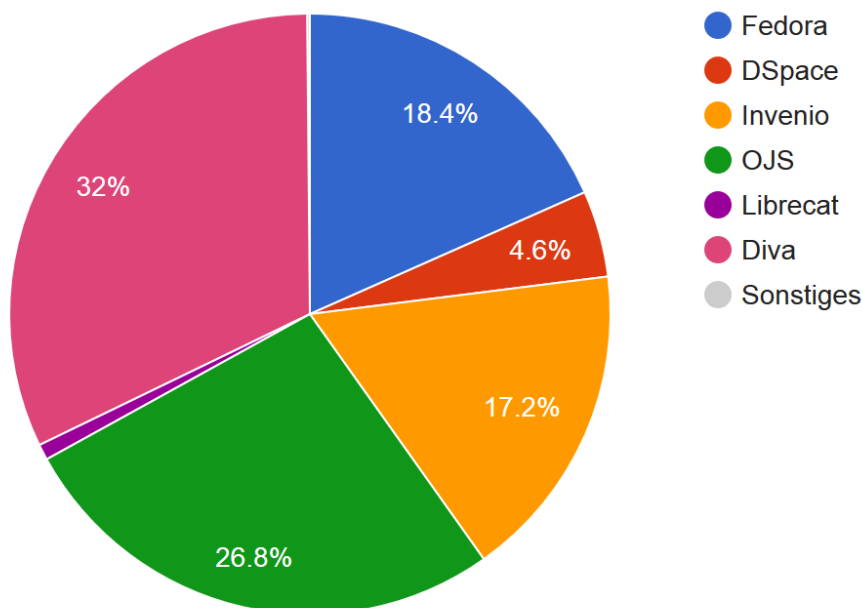
Die folgenden Abbildungen zeigen die prozentuale Verteilung der Spalten 3 und 4:

Autorenidentifikatoren - Verteilung Repository-Systeme mit ORCID-IDs



Offensichtlich ist beim DIVA-Service in Schweden bei allen gehosteten Institutionen ORCID umfanglich im Einsatz. Bei den weit verbreiteten Plattformen fällt OJS durch zahlenmäßig viele Installationen auf, während bei DSpace und Fedora nur vereinzelte Einsätze vorliegen. Daneben tauchen weniger verbreitete Systeme wie Librecat (Bielefeld, Gent), Django (Bochum) und Brocade (Antwerpen) auf.

Autorenidentifikatoren - Verteilung Anzahl ORCID-IDs in Repository-Systemen



Projekt ORCID DE

Als dieser Vortrag geplant wurde, war der Ausgang des ORCID-DE-Projektantrages noch ungewiss. Nun, da die Förderung feststeht (Teilnehmer sind das **Helmholtz Open Science Koordinationsbüro am Deutschen GeoForschungsZentrum GFZ**, die Deutsche Nationalbibliothek und die Universitätsbibliothek Bielefeld) kann die vorliegende Untersuchung auch im Projekt-Kontext genutzt werden.

Schwerpunkte der UB Bielefeld im Projekt sind:

- ORCID-Claiming-Service für den OAI-Dokumentenraum
- Standardisierung
- ORCID-Integration jenseits von Zeitschriftenartikeln
- Unterstützung der ORCID-Implementierung

Ausblick/Zusammenfassung

Folgendes lässt sich zusammenfassend feststellen:

- ORCID-Strukturen sind vorbereitet!
- ORCID-Informationen werden zunehmend eingespielt!
- Die Sichtbarkeit in den Repository-Oberflächen wird zunehmen!
- Die Sichtbarkeit in den Schnittstellen wird zunehmen!
- Die Nutzung der Schnittstellen für Nachnutzung der Informationen wird zunehmen (LOD)!

Anhang Konkrete Beispiele der Verwendung von Autorenidentifikatoren

PUB-System UB Bielefeld (Librecat)



[← Back](#)

Towards Linked Research Data: An Institutional Approach

Wijes C, Jahn N, Lier F, Paul-Stueve T, Vompras J, Pietsch C, Cimiano P (2013)
In: 3rd Workshop on Semantic Publishing (SePublica). García Castro A, Lange C, Lord P, Stevens R (Eds); CEUR Workshop Proceedings, Aachen: 27-38.

Download [sepublica-2013_institutional_approach.pdf](#)
URN [urn:nbn:de:0070-pub-25806212](#)
URL <http://ceur-ws.org/Vol-994/paper-03.pdf>
Conference Paper | Published | English

Details | **Files** | **Related Material**

Author **Wijes, Cord**^{UniBi} ; **Jahn, Najko**^{UniBi} ; **Lier, Florian**^{UniBi} ; **Paul-Stueve, Thilo**^{UniBi} ; **Vompras, Johanna**^{UniBi} ; **Pietsch, Christian**^{UniBi} ; **Cimiano, Philipp**^{UniBi}
Editor García Castro, Alexander ; Lange, Christoph ; Lord, Phillip ; Stevens, Robert
Department **Universitätsbibliothek**
Technische Fakultät, AG Semantic Data Science und Wissensmanagement

Der Link auf die ORCID-Seite des Autors liefert:

ORCID ID
 orcid.org/0000-0001-8778-1273

Saarbrücken, Milton Keynes, Bielefeld

Education (1) Sort

Universität des Saarlandes: Saarbrücken, Saarland, Germany
Diplom-Linguist (Computerlinguistik) (Computational Linguistics)
Source: Christian Pietsch Created: 2014-03-24

Employment (2) Sort

Open University: Milton Keynes, Milton Keynes, United Kingdom
project officer (Computing)
Source: Christian Pietsch Created: 2014-03-24

Universität Bielefeld: Bielefeld, Nordrhein-Westfalen, Germany
Source: Christian Pietsch Created: 2014-03-24

Works (7) Sort

Why We Develop Intelligent Conversational Agents with Python. In: What Language Do You Use to Create Your AI Programs and Why?

BORIS Univ. Bern (EPrints)

Deutsch | English | Français

BORIS

Bern Open Repository and Information System

Home Statistik

Durchsuchen nach ORCID

Klicken Sie auf einen Link, um sich die Publikationen anzeigen zu lassen.

(1768)	0000-0002-0858-3963 (1)	0000-0002-7709-4127 (2)
(148)	0000-0002-0928-4084 (1)	0000-0002-7754-9590 (1)
0000-0001-5016-9822 (3)	0000-0002-0928-5722 (26)	0000-0002-7862-736X (5)
0000-0001-5040-7448 (1)	0000-0002-0956-5329 (9)	0000-0002-7880-7205 (1)
0000-0001-5044-5057 (1)	0000-0002-0990-9646 (2)	0000-0002-7906-5629 (1)
0000-0001-5117-6047 (1)	0000-0002-1106-6123 (5)	0000-0002-7910-0324 (4)
0000-0001-5141-683X (3)	0000-0002-1129-3406 (1)	0000-0002-7923-3741 (2)
0000-0001-5218-1649 (1)	0000-0002-1140-8664 (2)	0000-0002-7926-5874 (14)
0000-0001-5317-3990 (2)	0000-0002-1159-484X (2)	0000-0002-7980-3601 (1)
0000-0001-5343-5954 (1)	0000-0002-1165-6377 (31)	0000-0002-7980-3601 (1)
0000-0001-5368-8169 (34)	0000-0002-1190-9880 (2)	0000-0002-7980-3601 (1)
0000-0001-5416-5044 (1)	0000-0002-1200-9577 (1)	0000-0002-7980-3601 (1)

Haben Sie eine Anr...

King Abdullah Univ. of Science and Technology, Saudi-Arabien (DSpace)



KAUST Repository

Home Browse About Submit an Item Visualize Advanced Search

KAUST Repository > Research > Articles >

Complete Genome Sequence of Mycobacterium vaccae Type Strain ATCC 25954

This item has been cited 2 time(s). (data provided by Scopus)

HANDLE URI:
<http://hdl.handle.net/10754/554099>

TITLE:
Complete Genome Sequence of Mycobacterium vaccae Type Strain ATCC 25954

AUTHORS:
Ho, Y. S.; Adroub, S. A.; Abadi, Maram (0000-0003-3802-8669) ; Al Alwan, B.; Alkhateeb, R.; Gao, G.; Ragab, A.; Ali, S.; van Soolingen, D.; Bitter, W.; Pain, Arnab (0000-0002-1755-2819) ; Abdallah, A. M.

ABSTRACT:
Mycobacterium vaccae is a rapidly growing, nontuberculous Mycobacterium species that is generally not considered a human pathogen and is of major pharmaceutical interest as an

Files 1

J. Bacteriol.-2012-H...
88KB

Download

BENZIE Iris FF

- Profile
- Research Outputs
- Bibliometrics

Update my Profile

- Network Lab
- View Statistics
- RSS Feed

Author Identifications

ORCID 0000-0002-2312-0318
Author ID (Scopus) 7004481510

Profile



BENZIE Iris FF

Chair Professor of Biomedical Science
Department of Health Technology and Informatics
Faculty of Health and Social Sciences

OJS-Testinstallation (UB Bielefeld)

Bibliotheksebene

ÜBER DIE AUTORIN/DEN
AUTOR

Sebastian Wolf
[ORCID-iD](#)

Startseite > Bd. 28, Nr. 1 (2014) > **Wolf**

10 Jahre BASE

Sebastian Wolf

Die Entwicklung unserer Suchmaschine BASE von 2004-2014

Via OAI-PMH liefert die Installation die ORCID-ID im Marc-Format wie folgt aus:

```
<varfield tag="042" ind1=" " ind2=" " ><subfield label="a" >dc</subfield>
</varfield>
<varfield tag="245" ind1="0" ind2="0" ><subfield label="a" >10 Jahre BASE</subfield>
</varfield>
<varfield tag="100" ind1="1" ind2=" " ><subfield label="a" >Wolf, Sebastian</subfield>
<subfield label="0" >http://orcid.org/0000-0002-6297-3348</subfield>
</varfield>
<varfield tag="520" ind1=" " ind2=" " ><subfield label="a" >Die Entwicklung unserer Suchmaschine BASE von
2004-2014</subfield>
</varfield>
<varfield tag="260" ind1=" " ind2=" " ><subfield label="b" >Bibliotheksebene</subfield>
</varfield>
<dataField tag="260" ind1=" " ind2=" " ><subfield label="c" >2014-10-24 08:12:46</subfield>
</dataField>
```

James Madison University Scholarly Commons (Digital Commons)

DISSERTATIONS

The Effects of Academic Libraries' Resource, Expenditure, and Service Decisions on Library Use: An Analysis of ACRL and NCES Data

[Jody C. Fagan, James Madison University](#)

Follow

Download

201 Downloads
Since February 09, 2015

Preferred Name

Jody Condit Fagan

ORCID

0000-0003-2861-3477

Date of Award

Fall 2014

Document Type


Dissertation/Thesis

Degree Name

Included in
[Higher Education Commons, Library and Information Science Commons](#)

SHAR





BTH

BLEKINGE
INSTITUTE OF
TECHNOLOGY

Simple search **Advanced search -
Research publications** Advanced search -
Student theses Statistics

English Svenska Norsk

Change search

Research publications > Result list (...) > Real-time landscape-size convective clouds simulation

References 1 2 3 4 2 of 4 Link to record

<< Back to result list Share

Real-time landscape-size convective clouds simulation

- Goswami, Prashant
Blekinge Institute of Technology, Faculty of Computing, Department of Creative Technologies. BTH. (DIKR)
ORCID id: 0000-0003-0891-2859

2015 (English)
In: Proceedings of the 19th ACM Symposium on Interactive 3D Graphics, ACM, 2015, 135- p.
Conference paper, Poster (Refereed)



Place, publisher, year, edition, pages
ACM, 2015. 135- p.

National Category
Computer Science

Identifiers
URN: urn:nbn:se:bth-736
ISBN: 978-1-4503-3392-4
OAI: oai:DIVA.org:bth-736
DIVA: diva2:815187

Conference
ACM Symposium on Interactive 3D Graphics

Available from: 2015-05-29 **Created:** 2015-05-29 **Last updated:** 2015-06-03
Bibliographically approved

Open Access in DiVA
PDF fulltext (3456 kB)  
24 downloads

Other links
<http://dl.acm.org/citation.cfm?id=2721396>

Search in DiVA

By author/editor
Goswami, Prashant

By organisation
Department of Creative Technologies

On the subject
Computer Science

Search outside of DiVA
Google
Google Scholar

Downloads of Publication
20