Vendor-Supplied Usage Data: Challenges and Opportunities

ER&L Conference

March 25, 2006

THE UNIVERSITY of TENNESSEE

MaxData Project:
UT Libraries Team

Gayle Baker, UT Libraries
Eleanor Read, UT Libraries
Maribeth Manoff, UT Libraries
Dick Kawooya, UT SIS

Carol Tenopir, Project Director, UT SIS and
Center for Information Studies
http://web.utk.edu/~tenopir/maxdata/index.htm

THE UNIVERSITY *of* TENNESSEE

MaxData

"Maximizing Library Investments in Digital Collections Through Better Data Gathering and Analysis"

Funded by Institute of Museum and Library Services (IMLS) 2005-2007

Talk Outline

- UT Libraries MaxData Team
- UT Libraries Environment
- Vendor Reports: Acquisition, Formats, Data
- Challenges
- Opportunities

UT Libraries MaxData Team

- Survey on use of e-journals during September 2005 - November 2005
- Gathered e-journal usage data from that time period
- Combined data from COUNTER JR1 reports to prepare for analysis with other local data from same time period

UT Libraries Environment

- 260+ E-resources (subscription) listed on our database menu (2004-05)
- E-journals (2004-05)
  - 17,236 titles from aggregators
  - 6,265 from publishers, etc.
- SFX used to generate E-journal list and URLs and short records for catalog

UT Libraries Environment

- 30 vendors provided COUNTER-compliant reports for UT subscriptions for 2005

- 34 other vendors provide some source of statistics

	Vendor	JR1	JR2	DB1	DB2	DB3	JR3	JR4
2	ACS Publications	Yes					Yes	Yes
3	Annual Reviews	Yes				Yes	Yes	No
4	BioOne	Yes					No	No
5	Blackwell Publishing	Yes				Yes	No	No
6	BMJ Publishing Group	Yes	Yes				Yes	Yes
7	Cambridge Scientific Abstracts	Yes		Yes		Yes	No	Yes
8	CSIRO Publishing	Yes					No	No
9	EBSCO Publishing	Yes		Yes	Yes	Yes	No	No
10	Elsevier-ScienceDirect	Yes		Yes		Yes	No	No
11	Emerald Group Publishing	Yes					No	No
12	HighWire Press	Yes	Yes	Yes	Yes	Yes	Yes	Yes
13	Ingenta	Yes				No	No	No
14	MetaPress	Yes					No	No
15	National Information Services Corporation	Yes	Yes	Yes	Yes	Yes	Yes	No
16	Nature Publishing Group							
17	(includes Palgrave Journals)	Yes					Yes	Yes
18	Oxford University Press	Yes					No	No
19	Project MUSE	Yes		Yes	No	No	Yes	Yes
20	ProQuest Information & Learning	Yes		Yes		Yes	No	No
21	Royal Society of Chemistry	Yes					No	No
22	Springer-KluwerOnline	Yes					Yes	No
23	Scitation-AIP							
24	Publishing Services (AIP/APS, ASCE, ASME)	Yes		Yes		Yes	Yes	Yes
25	Taylor & Francis Group	Yes					No	No
26	Thomson Learning/Gale	Yes (zero usage not reported)		Yes	Yes	Yes	No	No
27	Thomson Scientific			Yes	Yes	Yes		
28	Wiley (applies to *Enhanced Access License* only)	Yes					No	No
29	H W Wilson	Yes		Yes	Yes	Yes	No	No
30	Wolters Kluwer Health (Ovid)	Yes	Yes	Yes	Yes	Yes	No	No
31	Wolters Kluwer Health (SilverPlatter)			Yes	Yes	Yes		

1	Vendor Stats for UT resources that are not COUNTER-compliant
2	ABC-CLIO
3	AccuNet/AP Photo Archive
4	ASAE Technical Library
5	Beilstein Crossfire (U Wisc Minerva)
6	Bowker (Books in Print and Ulrich's)
7	Brepols (International Medieval Bibliography)
8	Caliber (U. Calif. Press)
9	CIOS (CommAbstracts)
10	Columbia Univ. Press (CIAO, Graingers, Gazetteer)
11	Congressional Quarterly (CQ)
12	ComDisDome
13	D&B Net Advantage
14	IEEE (IEL)
15	EIU (Business newsletters)
16	EiVillage2
17	Factiva
18	Haworth
19	IOP
20	JSTOR
21	LexisNexis (Academic, Congressional, etc.)
22	Marcive
23	Mergent
24	MathSciNet
25	OCLC FirstSearch
26	OCLC NetLibrary
27	NewsBank / Readex
28	Oxford U Press (OED, Groves 2 dict, DNB, OxRefOnline)
29	PsycArticles (APA)
30	RLG
31	Royal Society
32	K.G. Saur (IBZ Online, IBR Online)
33	SciFinder Scholar (AmChemSoc)
34	Thomson Gale RDS
35	Vanderbilt TV News Archive

COUNTER Code of Practice: Scope

2.2 Scope

The COUNTER Code of Practice provides a framework for the recording and exchange of online usage statistics at an international level. In doing so, it covers the following areas: data elements to be measured; definitions of these data elements; content and format of usage reports; requirements for data processing; requirements for auditing; guidelines to avoid duplicate counting when intermediary gateways and aggregators are used. In response to librarian feedback, Release 1 of the Code of Practice focuses on journals and databases, as there is more general agreement on the data elements to be measured and on their definitions than for other content types. Journals and databases also represent the largest components of most library materials budgets. A wider range of content types and more detailed levels of reporting will be covered as the Code of Practice is upgraded and extended in subsequent releases.

COUNTER: Application

2.3 Application

COUNTER is designed for librarians, vendors and intermediaries. The guidelines provided in the Code of Practice enable librarians to compare statistics from different vendors, to make better-informed purchasing decisions, and to plan infrastructure more effectively. COUNTER also provides vendors/intermediaries with the detailed specifications they need to generate data in a format useful to customers, to compare the relative usage of different delivery channels, and to learn more about online usage patterns. COUNTER also provides guidance to others interested in information about online usage statistics.

Acquisition

- Request
  - Ask account representative
  - Ask consortia administrator
  - Web
    - Administrative site
    - Reports site
    - E-resource site

THE UNIVERSITY *of* TENNESSEE

COUNTER: Report Delivery

4.3 Report delivery

Report delivery must conform to the following standards for Release 1:

- Reports must be provided either as a CSV file, as a Microsoft Excel file, or as a file that can be easily exported to Microsoft Excel
- Reports should be made available on a password-controlled website (accompanied by an e-mail alert when data is updated)
- Reports must be provided at least monthly
- Data must be updated within two weeks of the end of the reporting period
- All of last calendar year's data and this calendar year's to date must be supplied

Acqusition: Web Delivery

- Instant
- Email
  - Scheduled
  - On request
    - Immediate – 48 hours

Acquisition/Web Interface: Time Period

- Selection options
  - Specific From/To date range
  - Single month
  - Calendar year
  - Previous/Prior Year(s)
  - Year-to-date

THE UNIVERSITY *of* TENNESSEE

Acquisition/Web Interface: Time Period

- Output time slices
  - Cumulative, year-to-date
  - Monthly, year-to-date (COUNTER)
  - Specific From/To date range

THE UNIVERSITY *of* TENNESSEE

Acquisition/Web Interface: Selection of Data

- Some vendors provided a choice:
  - All publications
  - Publications with usage
  - Subscribed publications

Formats/Output

- CSV
- Excel
- HTML
- PDF
- Tab-delimited
- XML
- Other

COUNTER: JR1 Format

Journal Report 1: Number of Successful Full-Text Article Requests by Month and Journal
(Full journal name, print ISSN and online ISSN are listed.)

	Print ISSN	Online ISSN	Jan-01	Feb-01	Mar-01	Calendar YTD
Total for all journals			**6637**	**8732**	**7550**	**45897**
Journal of AA	1212-3131	3225-3123	456	521	665	4532
Journal of BB	9821-3361	2312-8751	203	251	275	3465
Journal of CC	2464-2121	0154-1521	0	0	0	0
Journal of DD	5355-5444	0165-5542	203	251	275	2978

Formats/Data Fields

- Titles – all uppercase, mixed, leading articles, subscriber id

ACTA AGRICULTUR SCANDINAVICA. SECTION A, ANIMAL SCIENCE.	
Acta Agriculturae Scandinavica: Section B, Soil & Plant Science	

Journal	Print Issn	Online Issn
Systematic Botany	"03636445"	""
The American Biology Teacher	"00027685"	""

Title
The EMBO Journal (Subscriber ID: INST_00512897)
EMBO Reports (Subscriber ID: INST_00512897)

Formats/Data Fields

- ISSN – leading 0's , starting with ", enclosed in "" , with/without hyphen, ISBN appearing in ISSN field

Journal	Print Issn	Online Issn
Systematic Botany	"03636445"	""
The American Biology Teacher	"00027685"	""

	Print ISSN
ABA Journal	7470088
Abacus	13072
ABBA	0361686X

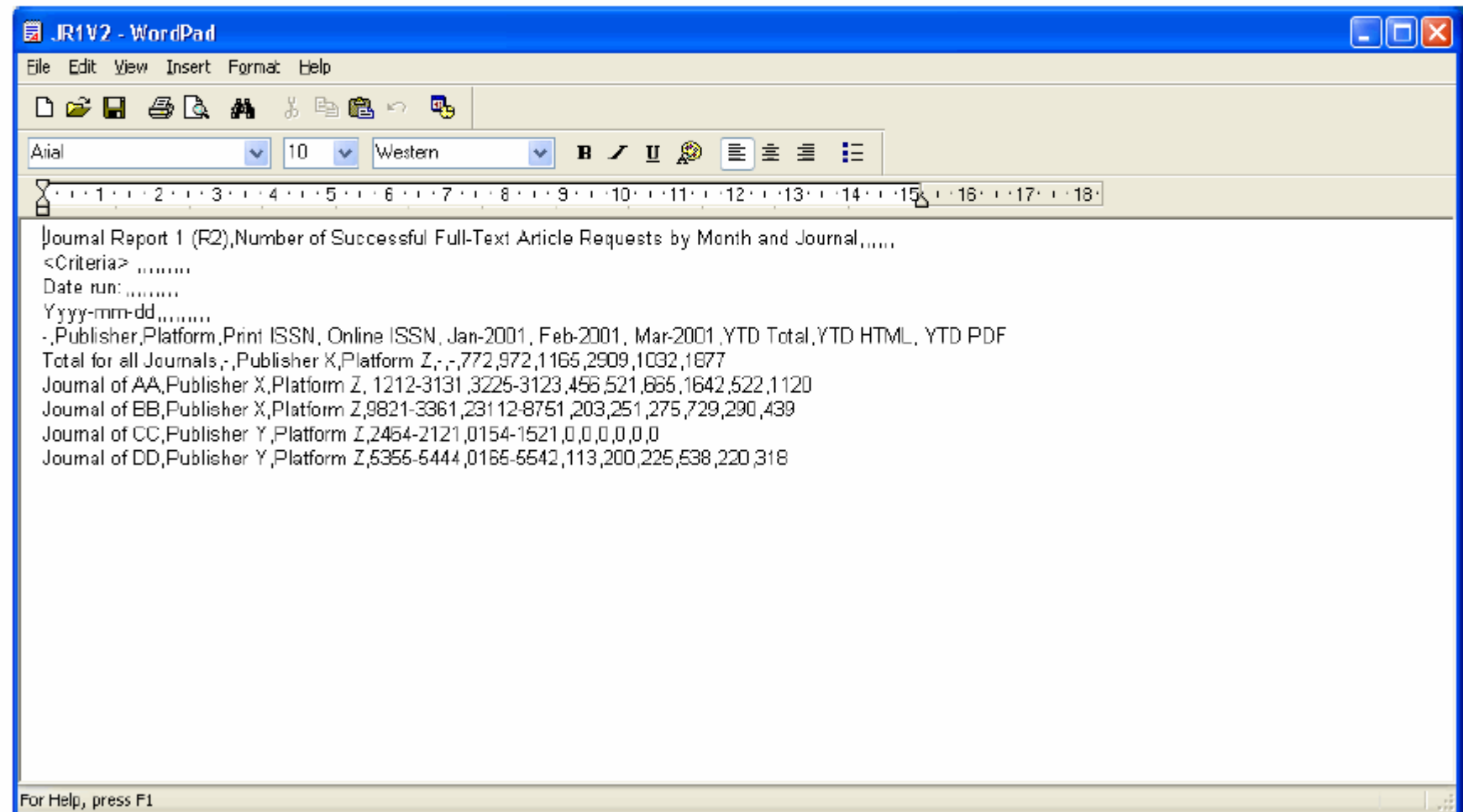|	Publisher	Print ISSN
Total for all journals		
100 Best Careers for the 21st Century	Peterson's	0028605950
19TH CENTURY MUSIC	University of C	0148-2076

COUNTER: ISSN

3.1.1.12	Print ISSN	Free text format (up to 13 characters in future)	Unique International Standard Serial Number assigned to the print version of a journal by the national ISSN agency of the country from which the journal is published
3.1.1.13	Online ISSN	Free text format (up to 13 characters in future)	Unique International Standard Serial Number assigned to the online version of a journal by the national ISSN agency of the country from which the journal is published.

Journal Report 1: Number of Successful Full-Text Article Requests by Month and Journal
(Full journal name, print ISSN and online ISSN are listed.)


```
JR1V2 - WordPad
File   Edit   View   Insert   Format   Help

Arial          10      Western        B  I  U

Journal Report 1 (R2),Number of Successful Full-Text Article Requests by Month and Journal,,,,,,
<Criteria> ,,,,,,,,,
Date run: ,,,,,,,,,
Yyyy-mm-dd,,,,,,,,,
-,Publisher,Platform,Print ISSN, Online ISSN, Jan-2001, Feb-2001, Mar-2001,YTD Total,YTD HTML, YTD PDF
Total for all Journals,-,Publisher X,Platform Z,-,-,772,972,1165,2909,1032,1877
Journal of AA,Publisher X,Platform Z, 1212-3131,3225-3123,456,521,665,1642,522,1120
Journal of BB,Publisher X,Platform Z,9821-3361,23112-8751,203,251,275,729,290,439
Journal of CC,Publisher Y,Platform Z,2464-2121,0154-1521,0,0,0,0,0,0
Journal of DD,Publisher Y,Platform Z,5355-5444,0165-5542,113,200,225,538,220,318

For Help, press F1
```

Formats/Data Fields

- Numeric fields with "n/a" or spaces or dash for missing data or no use or not yet available

Publication	Print ISSN	Online ISS	ISBN	Publisher	Jan-05	Feb-05
Agriculture & Food	9998-0002			XXXXX	n/a	n/a

Title	Sep-2005	Oct-2005	Nov-2005
Arthritis Care & Research	-	-	0
Bioelectromagnetics 1980-1995	-	0	0

The Counts (JR1): Issues

- Vendor only lists publications with usage
- Accessible titles, but paid for by another local unit (medical/law library)
- Titles from trials
- Excessive downloading

	Sep-05	Oct-05	Nov-05	Dec-05	Calendar YTD
Total for all journals	1351	2099	2001	2484	16906
Veterinary Radiology & Ultrasound	97	87	39	1124	1990

THE UNIVERSITY *of* TENNESSEE

COUNTER: JR1

- ## Zero-Use Note

2. Journals for which the number of full-text article requests is zero in every month should be included in Journal Report 1

- ## Customer Data

4.2 Customer categories for Usage Reports:
Customer accounts, access and entitlements to vendor sites are organized in a number of different ways, but most commonly by IP addresses or by username/password. …

Augmenting UT Vendor Data

- Convert ISSN to xxxx-xxxx format

- Remove subscriber ID from some titles

- Add Publisher and Platform columns (COUNTER, rev. 2)

- Add collection name to Platform when possible

- Separate instances where all journal reports (JR1, JR2, …) appear on the same page

THE UNIVERSITY *of* TENNESSEE

Challenges: Vendor

- Minimize post-download processing of files by library staff
- One programmer can save labor for hundreds of library staffers

Suggestions for Project COUNTER

- Add a title sorting field: uppercase, leading articles and commas removed
- Provide a standard for collections with multiple document types
- To "make better-informed purchasing decisions"
  - Add the subscribed E-resource collection/service (concatenate to Platform name)
  - Identify subscribed vs. non-subscribed journals

Suggestions for Project COUNTER

- Deal with errors of omission and commission in audit

  - Check for zero-use titles

  - Check for titles that we may not have access to

- Establish an ISSN standard – most important field

THE UNIVERSITY *of* TENNESSEE

Opportunities

- Helpful tools and initiatives
  - MPS *ScholarlyStats*
  - Thomson Scientific *Journal Use Reports*
  - ERUS
  - ERMS
  - SUSHI
  - UKSG / Project Counter Usage Factor Feasibility Study

THE UNIVERSITY *of* TENNESSEE

SUSHI

- John Martin, chief commercial director at Swets, says that the main impact of SUSHI will be to allow Swets to improve its service to customers. "**At the moment, publishers base their usage on the number of users, which is an estimate. If they could actually demonstrate what the usage is they could come up with the perfect model for pricing**," he said.

- Tracey Caldwell, "Sushi to Satisfy Usage Demands," Information World Review, 17 January 2006, p. 3. (Retrieved from LexisNexis Academic March 19, 2006)

THE UNIVERSITY *of* TENNESSEE

Opportunities

- Notable non-COUNTER reports
  - Referring URLs
  - Top 10 articles viewed
  - Daily usage along with turnaways

THE UNIVERSITY *of* TENNESSEE

Opportunities

- COUNTER, Release 2 (JR1)
  - Platform and Publisher columns added
  - Dates Month-YYYY
  - Separate totals for HTML and PDF downloads
  - Examples of CSV formats for usage

THE UNIVERSITY *of* TENNESSEE

MaxData

- Future
  - Look at related use data from SFX for September through November 2005
  - Survey about vendor use data: amount of time, what data is most useful, etc.

Questions?

Resources

- Project COUNTER: www.project.counter.org
- ERUS (Electronic Resources Usage Statistics): http://web.simmons.edu/~andersoc/erus/
- SUSHI (Standardized Usage Statistics Harvesting Initiative): http://www.niso.org/committees/SUSHI/SUSHI_comm.html
- UKSG / Project Counter Usage Factor Feasibility Study: http://www.uksg.org/rfp.pdf