#### Trinity University Digital Commons @ Trinity

Library Faculty Research

**Coates Library** 

5-22-2015

#### Building Your Local Web One Thread at a Time: Archiving Websites at Trinity University

Megan Toups Trinity University, mtoups@trinity.edu

Follow this and additional works at: https://digitalcommons.trinity.edu/lib\_faculty Part of the <u>Library and Information Science Commons</u>

#### **Repository Citation**

Toups, M. (2015, May 22). Building Your Local Web One Thread at a Time: Archiving Websites at Trinity University. Presentation as part of the Tangled Webs: Taking the Mystery out of Website Archiving panel at the Society of Southwest Archivists Annual Conference, Arlington, Texas.

This Presentation is brought to you for free and open access by the Coates Library at Digital Commons @ Trinity. It has been accepted for inclusion in Library Faculty Research by an authorized administrator of Digital Commons @ Trinity. For more information, please contact jcostanz@trinity.edu.

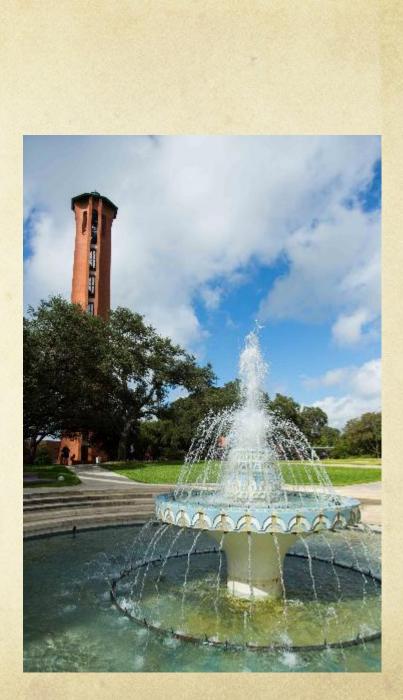
# Building Your Local Web One Thread at a Time: Archiving Websites at Trinity University

- Megan Toups, Special Collections Librarian and University Archivist, Trinity University, San Antonio, Texas
- <u>mtoups@trinity.edu</u>

SSA 2015 Conference









Classes

Digitization

Publishing

Preservation



Service

**Technical Services** 

**Public Services** 

Administration Outreach

**Exhibits** 

Training and Supervising

**Reference** Questions

Feeling **Overwhelmed**?

Processing

Housekeeping



**Events** 

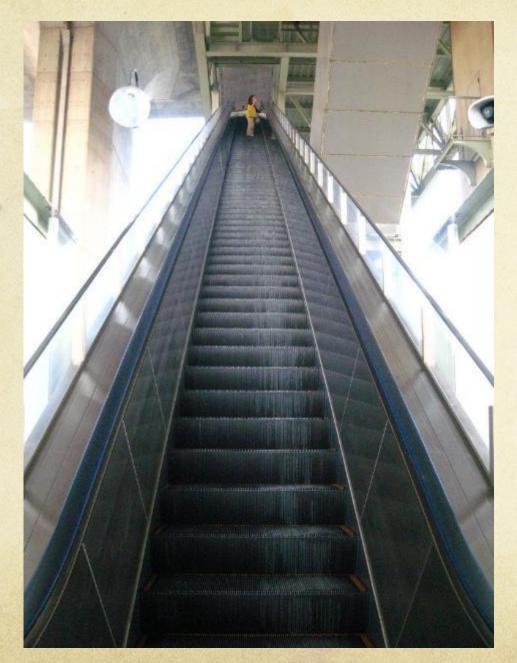
## Starting Web Archiving Felt Like:



## Can I Fit Web Archiving In?



## Web Archiving in 12 Steps



- Select Platform
- Create Policies
- Find and Compile Sites/Seeds
- Decide on Metadata Schemes
- Create Metadata
- Test Crawls
- Adjust Crawl Parameters
- Crawl Sites
- Add to Finding Aid
- Adjust/Update as Needed
- Create Procedures
- Outreach

## Select Platform: Open Source or Subscription?

### **Resources on Software:**

- <u>Practical E-Records:</u> <u>Software and Tools for</u> <u>Archivists</u>
- <u>International Internet</u>
   <u>Preservation Consortium</u>
   <u>Tools and Software</u>
- Digital Curation Centre's
   Web Archiving

## Archive-It (from the Internet Archive)





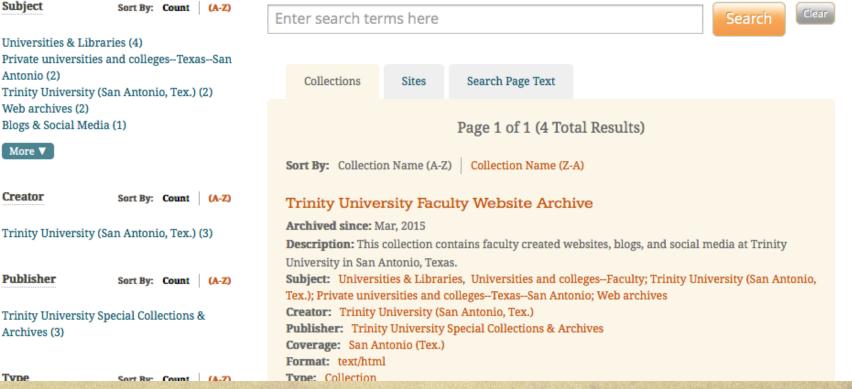
#### Trinity University

Archive-It Partner Since: Jun, 2012 **Organization Type: Colleges & Universities** Organization URL: http://www.trinity.edu/

Description: Trinity University is a transformational liberal arts and sciences institution. As a primarily residential undergraduate university with selected pre-professional and professional programs, Trinity introduces students to new ways of thinking and growing while creating and nurturing a diverse community that values multiple ways of understanding the world. Trinity University takes pride in being a catalyst for growth for its students and all members of the Trinity community.

#### Narrow Your Results

Sites and collections from this organization are listed below. Narrow your results at left, or enter a search query below to find a collection, site, specific URL or to search the text of archived webpages.



Universities & Libraries (4) Private universities and colleges--Texas--San Antonio (2) Trinity University (San Antonio, Tex.) (2) Web archives (2) Blogs & Social Media (1)

More V

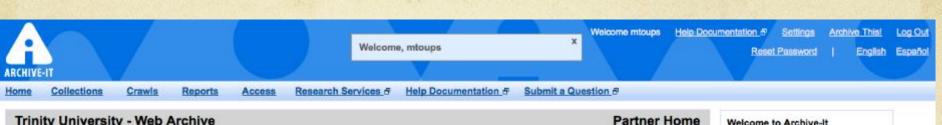
Creator

Publisher

Archives (3)

Type

Subject



**All Subscription Periods** 

Documents Crawled:

Create New Collection

Data Archived:

**Getting Started** 

Last Completed Crawl

April 29, 2015 7:07:54 AM CDT

April 29, 2015 4:12:36 PM CDT

December 10, 2014 9:02:01 AM CST

448.295

11.2%

40 GB

256 GB

15.6%

233

4,000,000

Trinity	University - Web Archive	
Partner S	ince June 2012	

Documents Crawled:

Data Archived:

Data Budget Used:

Total Active Seeds:

Active Collections

Data Budget:

Document Budget Used:

Subscription Document Budget:

Trinity University Faculty Website Archive

Trinity University Student Organizations Website Archive

Trinity University Social Media Archive

Trinity University Website Archive

Current Subscription (started Dec 15 2014)

ML					
	This home view gives you an				
	overview of your account activity				
	including subscription start date and				
	budget.				

1,301,991

115 GB

To create a new collection, click the "create new collection" link from the "collections" drop down menu at the top of the screen.

To manage existing collections, select a collection from the "collections" drop down menu at the top of the screen. You can also get to your active collections by using the links under "active collections" at the bottom of the screen. Information about current or upcoming crawls is available under the "crawls" link at the top of the screen.

- Learn more about getting started with Archive-It #
- Frequently Asked Questions about Archive-It &
- · Glossary of Web Archiving Terms 8

If you need assistance, please submit a support question #

Internet Archive - Archive-It Web UI 2 4.9-SNAPSHOT-prod-20150417-0005

Help 🗗 Settings

Archive This!

Submit a Support Question P

Next Scheduled Crawl

April 27, 2016 1:29:35 PM CDT

June 9, 2015 3:01:41 PM CDT

May 24, 2015 3:45:25 PM CDT

# Create Policies

- Coverage
- Preservation/Access
- Copyright
- Take Down

#### Trinity's Web Archiving Policy:

http://libguides.trinity.edu/ld .php?content\_id=1895810

#### **Coates Library Web Archiving Policy**

Coates Library's Special Collections and Archives is responsible for identifying and preserving the historical record of the university for future scholarship. In accordance with this mission, we gather web content about university activities using Archive-It.

#### Coverage:

Our focus is on documenting official and affiliated Trinity University websites including university administrative units, academic departments, student organizations, and university publications and events. This includes both webpages as well as social networking sites. Most sites will be crawled on a semi-annual or annual basis.

We do not ordinarily archive faculty or student personal webpages, but please contact us if you would like us to crawl your website. Also, if you are a member of a Trinity University office or student organization that does not have "trinity.edu" in the URL of your website, please let us know by emailing archives@trinity.edu.

#### **Preservation and Access:**

Coates Library has partnered with Archive-It in order to capture and preserve our campus web presence. Information on Archive-It's storage and preservation practices can be found https://webarchive.jira.com/wiki/display/ARIH/Archive-It+Storage+and+Preservation+Policy.

Archived websites are available for public viewing at https://archiveit.org/home/trinityuniversity and will be hosted by Archive-It in perpetuity. In the future we will also be adding information on our archived websites to Archon and the library catalog.

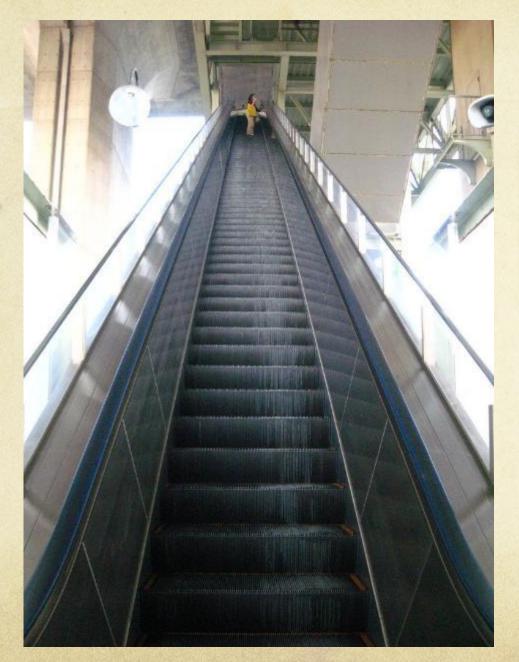
#### Copyright:

It is the sole responsibility of the user to determine the copyright status of any material in the archives before publication. If reproducing content from the web archive, we strongly suggest users identify the copyright holders of that material and contact them directly. Please contact us at archives@trinity.edu if you have any questions.

# Case Study: Archiving Student Organizations

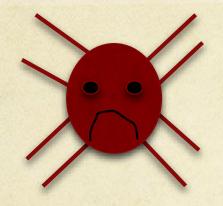
"Despite the importance of such organizational records to the mission of academic archives, they present many persistent difficulties for archival management. These stem from several conditions common to their existence. Student and faculty organizations, by their very nature, operate largely outside the administrative framework of the institution... Moreover, the mechanisms an archivist might normally use to solicit institutional records and arrange for regular transfers will only be marginally effective for these groups...their membership, leadership, program scope, and even their very existence are constantly changing. The fact that many of these organizations may become defunct before the archivist first learns of their existence makes difficult not only the acquisition of their records, but also their proper description." -William Maher, The Management of College and University Archives, p. 234-235 (1992)

## Web Archiving in 12 Steps



- Select Platform
- Create Policies
- Find and Compile Sites/Seeds
- Decide on Metadata Schemes
- Create Metadata
- Test Crawls
- Adjust Crawl Parameters
- Crawl Sites
- Add to Finding Aid
- Adjust/Update as Needed
- Create Procedures
- Outreach

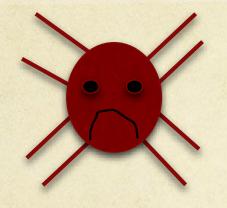
# Introducing...Spidey

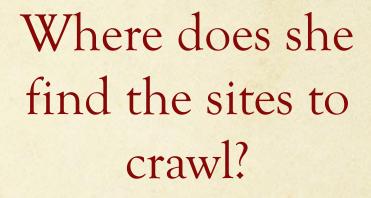




Spidey's web didn't make itself...to build it she needed certain threads...

## Thread #1



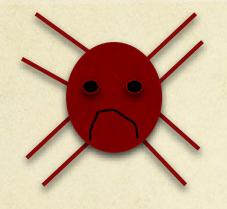




Google	trinity acabellas	megtoups@gmail.com +
	Web Videos News Shopping Images More - Search tools	± © ‡
	About 2,440 results (0.33 seconds)	
	Trinity AcaBellas   Facebook https://www.facebook.com/TrinityAcaBellas - Trinity AcaBellas. 956 likes · 3 talking about this. We are an all-female collegiate a cappella group based at Trinity University in San Antonio, Texas.	
	Trinity AcaBellas - YouTube https://www.youtube.com/user/TrinityAcaBellas - Trinity University's all-female collegiate a cappella group! Current AcaMembers: Natalie Feto '12 Kendra Doshier '12 Alexandra Hunt 12 Jessica Rodriguez '12.	
	"Valerie" - Trinity AcaBellas - Official Music Video - YouTube         www.youtube.com/watch?v=KNc8nRfH3BM         May 22, 2012 - Uploaded by Natalie Toto         Video written, directed and edited by Natalie Toto Cinematography         by Andrew Rieger Acabellas in Video: Ally	
(	Trinity University   Music   The AcaBellas web.trinity.edu > > Performing Ensembles  Trinity University  The AcaBellas are Trinity University's student-rup female a cappella group. It was founded in the fail of 2007 by two students who loved singing and wanted to	
	Something Monday: Trinity's AcaBellas to Carol in Main Plaza therivardreport.com/something-monday-the-acabellas-carol-in-main-plaza/ Dec 14, 2013 - AcaBellas are student-led organization from Trinity University with a to Main Plaza where The AcaBellas will sing favorite holiday carols	
	Trinity AcaBellas by leticia argueta on Prezi prezi.com/lw-ym9h5zbbd/trinity-acabellas/ Prezi ~ Nov 21, 2014 - Who makes the CUT ? Diversity Culture Range Style Can you BELL-ieve it? Element Aspect Persona Pitch Perfect Auditions Selections	
	Copy of Trinity AcaBellas by leticia argueta on Prezi	

🗄 🕤 - Č - =	SeedsList_2014 - Excel						? 🗹 — 🗆
LE HOME INSERT PAGE LAYOUT FORMULAS DATA	REVIEW VIEW						Toups, Megan 🔻
Cut $Calibri \cdot 12 \cdot A^* = =$	🖓 - 🐺 Wrap Text 🛛 General -		Normal Bad	· -		∑ AutoSum - A	
B Copy -					ert Delete Format	Fill → Z ■ ■ Sort & Find &	
Format Painter $\mathbf{B}  I  \underline{U}  \mathbf{P}  \mathbf{P}  \mathbf{A}  \mathbf{P} \equiv \mathbf{F} \equiv \mathbf{F}$		able -	GOOD Nei			✓ Clear ▼ Filter ▼ Select ▼	
Clipboard 🕞 Font 🕞	Alignment 🗔 Number 🗔		Styles		Cells	Editing	
$\bullet$ : $ imes$ $\checkmark$ $f_x$ Semi Annual							
А	В	С	D	E	F	G	
eeds							
	Exclude in next crawl if set up Semi Annual because already c	awled o	quNote: Need to inc	rease Facebook an	d Youtube docume	nt limit to have 2000 x # of	seeds and do t
Current Seeds							
ïtle	Seed (URL)	A	/Ir Last Update	Francisco	of c Crawl As?	Group?	
Greek Council	https://web.trinity.edu/x6187.xml	Active/	in Last Opdate	2014 Semi Annual	or crerawinks:	Greek Council	
rinity University Academic Honor Code	http://www.trinity.edu/departments/academic_affairs/honor	code/		2013 One-Time		Honor Council	
tudent Ambassadors at Trinity University	http://www.trinity.edu/student_org/ambassadors/	,		2014 Semi Annual		Student Ambassadors	
rinity University Student Ambassadors   Facebook	https://www.facebook.com/TrinityUniversityStudentAmbassac	lors/		2014 Semi Annual		Student Ambassadors	
rinity Diversity Connection   Facebook	https://www.facebook.com/TrinityDiversityConnection/			2012 One-Time		Trinity Diversity Connectio	n
rinity Diversity Connection Blog	http://trinitydiversityconnection.blogspot.com/			2012 One-Time		Trinity Diversity Connectio	n
rinity University Volunteer Action Community	https://sites.google.com/a/trinity.edu/tuvac/home/			2014 Semi Annual		Trinity University Voluntee	r Action Com
Frinity University Volunteer Action Community (TUVAC)   Facebo	https://www.facebook.com/pages/Trinity-University-Voluntee	r-Action	I-(	2014 Semi Annual		Trinity University Voluntee	r Action Comr
The AcaBellas	http://web.trinity.edu/x6519.xml		Unclear	One-Time	One Page Only	/ AcaBellas	
Trinity AcaBellas   Facebook	https://www.facebook.com/TrinityAcaBellas/			2014 Semi Annual		AcaBellas	
Trinity AcaBellas   Youtube	http://www.youtube.com/user/TrinityAcaBellas/			2012 One-Time		AcaBellas	
Trinity AED	https://sites.google.com/site/trinityuniversityaed/home/			2014 Semi Annual		Alpha Epsilon Delta (Pre-N	
Trinity University AED   Facebook	https://www.facebook.com/trinityAED/			2014 Semi Annual		Alpha Epsilon Delta (Pre-N	
Alpha Kappa Psi - Nu Pi Chapter	http://www.trinityakpsi.com/			2014 Semi Annual		Alpha Kappa Psi (Business	
Alpha Kappa Psi - Trinity University Spring 2014 Rush   Facebook				2014 One-Time		Alpha Kappa Psi (Business	and Economi
Ipha Lambda Delta - Trinity University	http://www.trinity.edu/student_org/ald/			2009 One-Time			
Alpha Phi Omega - Delta Pi Chapter Trinity ASME (American Society of Mechanical Engineers)   Facel	http://apoonline.org/deltapi/			2014 Semi Annual 2014 Semi Annual			
TU Bad Movie Club	https://www.facebook.com/TuBadMovieClub/			2014 Semi Annual 2013 One-Time	One Page Only	,	
3ARE (Beauty Advocates with Realistic Expectations)   Facebook		tions/		2013 Semi Annual	one rage on	BARE (Beauty Advocates v	vith Realistic F
BARE   Beauty Advocates with Realistic Expectations Blog	http://beautyadvocates.wordpress.com/			2013 One-Time		BARE (Beauty Advocates v	
Black Student Union	http://www.trinity.edu/student_org/bsu/			2007 One-Time		Black Student Union	
Black Student-Union   Facebook	https://www.facebook.com/TUbsu11/			2014 Semi Annual		Black Student Union	
Catholic Student Group at Trinity University	https://sites.google.com/a/trinity.edu/csg/			2014 Semi Annual		Catholic Student Group	
rinity University's Catholic Student Group   Facebook	https://www.facebook.com/trinityuniversitycsg/			2014 Semi Annual		Catholic Student Group	
Catholic Student Group	http://www.trinity.edu/student_org/csg/			2011 One-Time		Catholic Student Group	
The Chemistry Club	http://www.trinity.edu/student_org/chemistry/		Unclear	One-Time		Chemistry Club (American	Chemical Soci
Chemistry Club Trinity University	http://www.trinity.edu/student_org/chemistry/chemclub/			2004 One-Time		Chemistry Club (American	Chemical Soci
Chinese Culture Club	http://www.trinity.edu/student_org/ccc/		Unclear	One-Time			

## Thread #2





We have 170 websites for student organizations...

some are updated hourly, others weekly, others yearly...

so how frequently does she crawl them?

# We are taking "Snapshots" of the Web



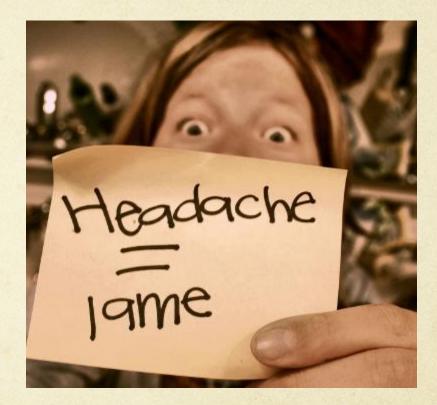
# What's the problem?

#### We have 170 sites

+ updated at various intervals

+ that may change updating schedule at any time

+ limited data storage

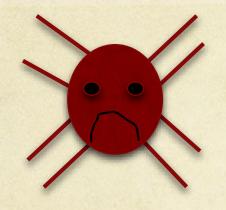


# Solution: Simple is Sustainable Website Updated Frequency of Our Crawls

- One Time
- Hourly
- Weekly
- Monthly
- Quarterly
- Semi-Annually
- O Annual

- O One Time
- Hourly
- Weekly
- Monthly
- Quarterly
- Semi-Annually
- Annual

## Thread #3





What if an organization has 2 websites or more? How do she keep them together?

### / Group

					10.00 E Ta	
	SeedsList_2014 - Excel					? 🗹 — 🗆 🗙
FILE HOME INSERT PAGE LAYOUT FORMULAS DATA	REVIEW VIEW					Toups, Megan 👻 🔍
$\land$ Cut Calibri $12 \sim A^{*} = =$	🖓 - 🗟 Wrap Text General -		ormal Bad	<b>€ □□</b>		∑ AutoSum - A
E Copy -				· · · ·		▼ Fill ▼
Paste 💞 Format Painter 🖪 B I 🙂 🗸 🖾 🗸 📥 🖛 🚍 🚍	E 🚈 🛱 Merge & Center ▾ 💲 ▾ % ㆍ 😚 🕺 Conditional Form	rmat as Go able ▼	ood Neutral	Ţ Insert	Delete Format	ert & Find & ≪ Clear * Filter * Select *
Clipboard 🗔 Font 🗔	Alignment IN Number IN	ubic	Styles		Cells	Editing
			,			
E44 🔻 🗄 🗙 🗹 $f_X$ Semi Annual						~ ~
А	В	С	D	E	F	G
1 Seeds						
2	Exclude in next crawl if set up Semi Annual because already cr	rawled qu	Note: Need to increase	Facebook and Y	outube docume	nt limit to have 2000 x # of seeds and do this
3 Current Seeds						
4						
5 Title	Seed (URL)	Active/In	r Last Update	Frequency of a	Crawl As?	Group?
41 Greek Council	https://web.trinity.edu/x6187.xml		2014	Semi Annual		Greek Council
42 Trinity University Academic Honor Code	http://www.trinity.edu/departments/academic_affairs/honor_	code/	2013	One-Time		Honor Council
43 Student Ambassadors at Trinity University	http://www.trinity.edu/student_org/ambassadors/		2014	Semi Annual		Student Ambassadors
44 Trinity University Student Ambassadors   Facebook	https://www.facebook.com/TrinityUniversityStudentAmbassad	dors/	2014	Semi Annual	]	Student Ambassadors
45 Trinity Diversity Connection   Facebook	https://www.facebook.com/TrinityDiversityConnection/		2012	One-Time		Trinity Diversity Connection
46 Trinity Diversity Connection Blog	http://trinitydiversityconnection.blogspot.com/		2012	One-Time		Trinity Diversity Connection
47 Trinity University Volunteer Action Community	https://sites.google.com/a/trinity.edu/tuvac/home/		2014	Semi Annual		Trinity University Volunteer Action Communi
48 Trinity University Volunteer Action Community (TUVAC)   Faceboo	https://www.facebook.com/pages/Trinity University Volunteer	r Action (	2014	Semi Annual		Trinity University Volunteer Action Commun
49 The AcaBellas	http://web.trinity.edu/x6519.xml		Unclear	One-Time	One Page Only	/ AcaBellas
50 Trinity AcaBellas   Facebook	https://www.facebook.com/TrinityAcaBellas/		2014	Semi Annual		AcaBellas
51 Trinity AcaBellas   Youtube	http://www.youtube.com/user/TrinityAcaBellas/		2012	One-Time		AcaBellas
52 Trinity AED	https://sites.google.com/site/trinityuniversityaed/home/		2014	Semi Annual		Alpha Epsilon Delta (Pre-Med)
53 Trinity University AED   Facebook	https://www.facebook.com/trinityAED/		2014	Semi Annual		Alpha Epsilon Delta (Pre-Med)
54 Alpha Kappa Psi - Nu Pi Chapter	http://www.trinityakpsi.com/		2014	Semi Annual		Alpha Kappa Psi (Business and Economics)
55 Alpha Kappa Psi - Trinity University Spring 2014 Rush   Facebook	https://www.facebook.com/trinityakpsispring2014rush/		2014	One-Time		Alpha Kappa Psi (Business and Economics)
56 Alpha Lambda Delta - Trinity University	http://www.trinity.edu/student_org/ald/		2009	One-Time		
57 Alpha Phi Omega - Delta Pi Chapter	http://apoonline.org/deltapi/		2014	Semi Annual		
58 Trinity ASME (American Society of Mechanical Engineers)   Faceb	https://www.facebook.com/TrinityAsme/		2014	Semi Annual		
59 TU Bad Movie Club	https://www.facebook.com/TuBadMovieClub/			One-Time	One Page Only	
60 BARE (Beauty Advocates with Realistic Expectations)   Facebook		tions/		Semi Annual		BARE (Beauty Advocates with Realistic Expe
61 BARE   Beauty Advocates with Realistic Expectations Blog	http://beautyadvocates.wordpress.com/			One-Time		BARE (Beauty Advocates with Realistic Expe
62 Black Student Union	http://www.trinity.edu/student_org/bsu/			One-Time		Black Student Union
63 Black Student-Union   Facebook	https://www.facebook.com/TUbsu11/			Semi Annual		Black Student Union
64 Catholic Student Group at Trinity University	https://sites.google.com/a/trinity.edu/csg/			Semi Annual		Catholic Student Group
65 Trinity University's Catholic Student Group   Facebook	https://www.facebook.com/trinityuniversitycsg/			Semi Annual		Catholic Student Group
66 Catholic Student Group	http://www.trinity.edu/student_org/csg/			One-Time		Catholic Student Group
67 The Chemistry Club	http://www.trinity.edu/student_org/chemistry/		Unclear	One-Time		Chemistry Club (American Chemical Society)
68 Chemistry Club Trinity University	http://www.trinity.edu/student_org/chemistry/chemclub/			One-Time		Chemistry Club (American Chemical Society)
69 Chinese Culture Club	http://www.trinity.edu/student_org/ccc/		Unclear	One-Time		
<ul> <li>ReadMeFirst Metadata Fields Collections Metadat</li> </ul>	a List Seeds Metadata List 🕂					•
READY						III III

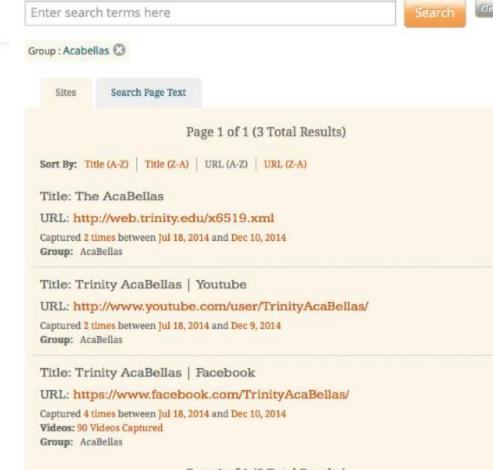
Group Sort By: Count (A-Z)	Enter search terms here Search				
Acabellas (3) Alpha Chi Lambda (3)					
Alpha Epsilon Delta (Pre-Med) (2) Alpha Kappa Psi (Business and Economics) (2)	Sites Search Page Text				
BARE (Beauty Advocates with Realistic Expectations) (2) Black Student Union (2) Catholic Student Group (3) Chemistry Club (American Chemical	Page 1 of 2 (170 Total Results)       Next Page ►         Sort By: Title (A-Z)   Title (Z-A)   URL (A-Z)   URL (Z-A)				
Society) (2) Chi Beta Epsilon (5)	Title: Alpha Chi Lambda Trinity University, San Antonio				
College Democrats (3) Delta Epsilon Iota (Career Services) (2) Equestrian Team (5) Gamma Chi Delta (2)	URL: http://alphachilambda.net/ Captured once on Dec 9, 2014 Group: Alpha Chi Lambda				
Geology Club (2) Greek Council (4) InterVarsity Christian Fellowship (2) Jewish Students Association (2) Krtu (3) Loon-E Crew (Hip Hop Dance Crew) (3)	Title: Alpha Phi Omega - Delta Pi Chapter URL: http://apoonline.org/deltapi/ Captured once on Dec 9, 2014				
Men's Lacrosse (2) Mirage (2) Model United Nations (1) Mu Phi Epsilon (Music) (3) National Society of Collegiate Scholars (2)	Title: BARE   Beauty Advocates with Realistic Expectations Blog URL: http://beautyadvocates.wordpress.com/ Captured 3 times between Jul 18, 2014 and Dec 9, 2014 Group: BARE (Beauty Advocates with Realistic Expectations)				
Neuroscience Club (2) Phi Delta Kappa (2) Phi Sigma Chi (2) Phi Sigma Pi (Honor Fraternity) (2) Pi Kappa Alpha (2) Reformed University Fellowship (RUF) (3)	Title: Chi Delta Tau URL: http://chideltatau.weebly.com/ Captured once on Jul 18, 2014				
Sexual Diversity Alliance (2) Sigma Theta Tau (3) Society of Women Engineers (2) Student Ambassadors (2) Student Government Association (2)	Title: The Rho Iota Chapter of Delta Epsilon Iota at Trinity University Blog URL: http://dei-trinityuniversity.blogspot.com/				
More ▼ Less ▲	Captured 2 times between Jul 18, 2014 and Dec 9, 2014 Group: Delta Epsilon Iota (Career Services)				

-----

Narrow Your Results

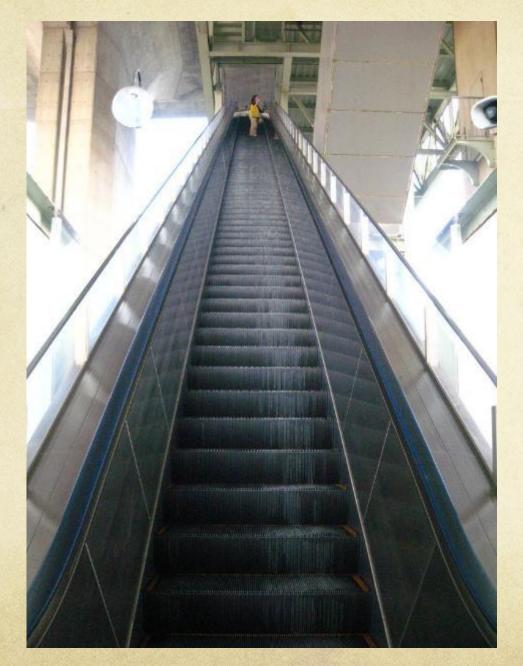
There are no further ways to narrow your results.

Sites for this collection are listed below. Narrow your results at left, or enter a search query below to find a site, specific URL or to search the text of archived webpages.



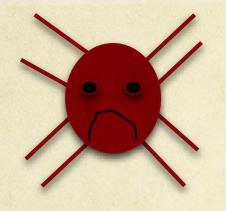
Page 1 of 1 (3 Total Results)

## Web Archiving in 12 Steps



- Select Platform
- Create Policies
- Find and Compile Sites/Seeds
- Decide on Metadata Schemes
- Create Metadata
- Test Crawls
- Adjust Crawl Parameters
- Crawl Sites
- Add to Finding Aid
- Adjust/Update as Needed
- Create Procedures
- Outreach

## Thread #4



How does she crawl sites so they look the way they do on the live web

AND



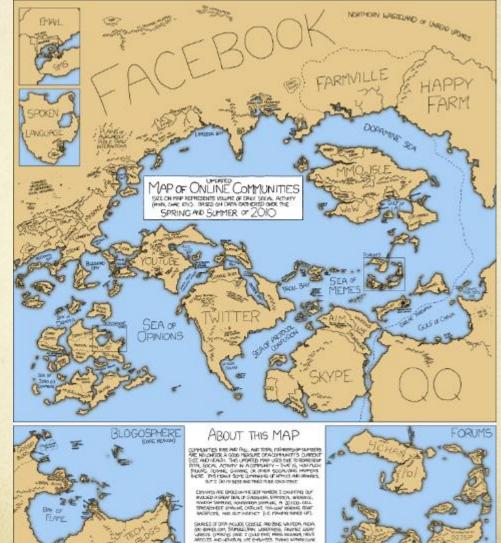
stay within her data budget?

## Test Crawls

- Depending on how much data you have to archive, good idea to run test crawls
- These crawls tell you how much data you will archive, but do not count that data against your subscription
- We ran test crawls to see how much data we would likely crawl

# Adjusting Crawl Parameters

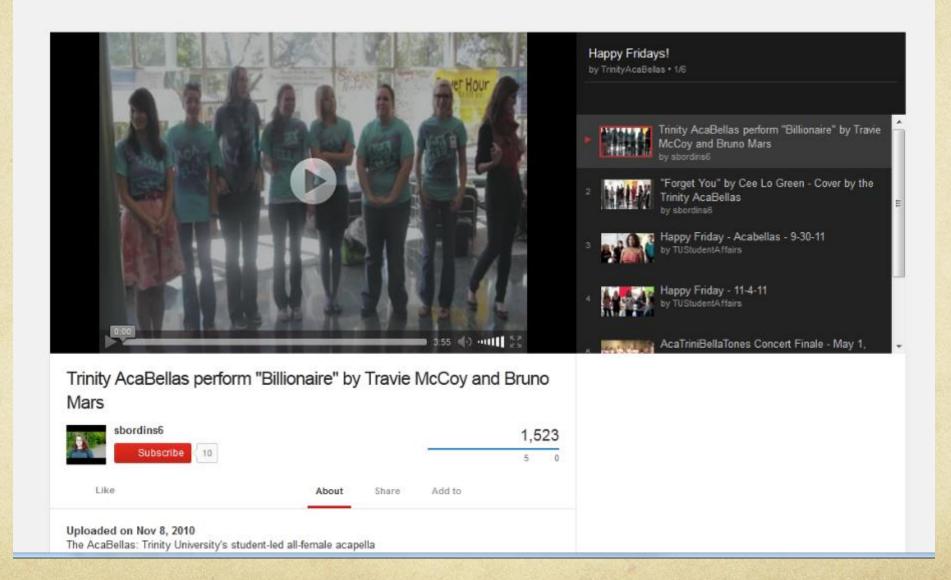
- Great idea for social media sites
- Without it, might capture poor quality or too much web
- Archive-It provides suggested parameters for social media sites including Facebook, Youtube, Twitter, Vimeo



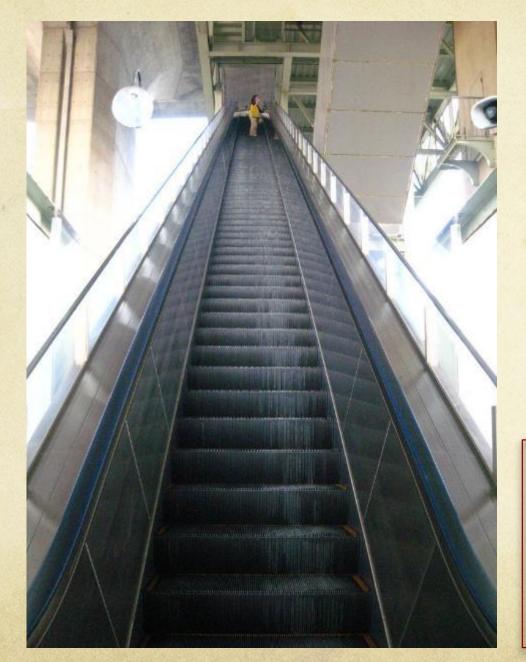
VINCTION AND IN SYMPHONE STARS



You are viewing an archived web page, collected at the request of Trinity University using Archive-It. This page was captured on 14:59:09 Jul 18, 2014, and is part of the Trinity University Student Organizations Website Archive collection. The information on this web page may be out of date. See All versions of this archived page. Videos

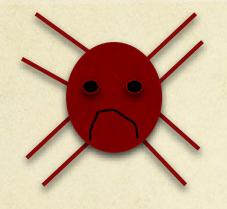


## Web Archiving in 12 Steps



- Select Platform
- Create Policies
- Find and Compile Sites/Seeds
- Decide on Metadata Schemes
- Create Metadata
- Test Crawls
- Adjust Crawl Parameters
- Crawl Sites
- Add to Finding Aid
- Adjust/Update as Needed
- Create Procedures
- Outreach

## Thread #5



How does she integrate archived sites into a finding aid?



Does she need to add a link for all 170 sites?

- 01 Corporate Administration (Board of Trustees)
- 02 University Administration
- 03 Faculty
- 04 Representative Bodies
- 05 Publications
- 06 Memorabilia
- 07 Audiovisuals
- 08 Special Collections
- 09 Websites

# Adding to Finding Aid

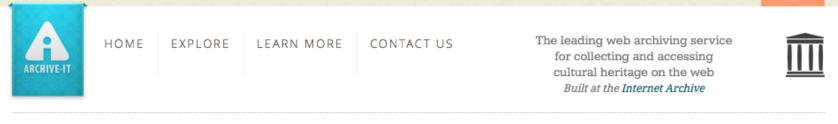
09 001 <u>Trinity University Website Archive</u>
09 002 <u>Trinity University Student Organizations Website Archive</u>
09 003 <u>Trinity University Faculty Website Archive</u>

**Scope and Contents:** This collection contains websites, blogs, publications, and social media of student organizations and groups at Trinity University in San Antonio, Texas. These groups include university-sponsored organizations, registered student organizations, and fraternities and sororities. Archiving began in May 2014.

On-line Images/Records

Remember: Simple is Sustainable

Available:	https://archive-it.org/collections/4570
Title:	Trinity University Student Organizations Website Archive
Date:	2014-Present
<b>Repository:</b>	Trinity University Archives and Special Collections
Found in:	Trinity University Student Organizations Website Archive



Explore >> Trinity University >> Trinity University Student Organizations Website Archive



#### Trinity University Student Organizations Website Archive

Collected by: Trinity University

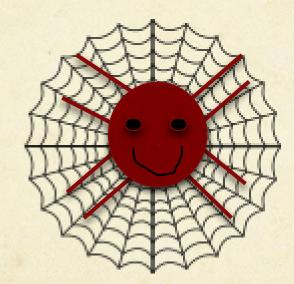
#### Archived since: May, 2014

Description: This collection contains websites, blogs, publications, and social media of student organizations and groups at Trinity University in San Antonio, Texas. These groups include universitysponsored organizations, registered student organizations, and fraternities and sororities. Subject: Universities & Libraries, Trinity University (San Antonio, Tex.), College students--Texas--San Antonio, Student activities--Texas--San Antonio, Private universities and colleges--Texas--San Antonio, Web archives Creator: Trinity University (San Antonio, Tex.) Publisher: Trinity University Special Collections & Archives Coverage: San Antonio (Tex.) Format: text/html Type: Collection Language: eng **Collector:** Trinity University Special Collections & Archives **Rights:** It is the sole responsibility of researchers to comply with federal copyright regulations and laws for fair use and in the cases of third party copyright. If reproducing content from the web archive, we strongly suggest users identify the copyright holders of that material and contact them directly. Please contact Special Collections and Archives (http://lib.trinity.edu/lib2/spec\_coll.php) if you have any

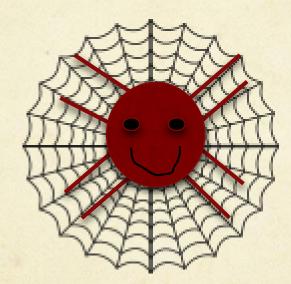
#### Narrow Your Results

Sites for this collection are listed below. Narrow your results at left, or enter a search query below to find a site, specific URL or to search the text of archived webpages.

# Spidey made it and built her web!



# Where Will Spidey Go Next?



# Faculty Websites

You are viewing an archived web page, collected at the request of Trinity University using Archive-It. This page was captured on 15:25:28 Mar 20, 2015. The information on this web page may be out of date. See All versions of this archived page.

Thirty years ago columnist Lewis Lapham made the following observation:

There no longer exists a theater of ideas in which artists or philosophers can perform the acts of the intellectual or moral imagination. In nineteenth-century England Charles Darwin could expect On The Origin of Species to be read by Charles Dickens as well as by Disraeli and the vicar in the shires who collected flies and water beetles. Dickens and Disraeli and the vicar could assume that Mr. Darwin might chance to read their own observations. But in the United States in 1979 what novelist can expect his work to be read by a biochemist, a Presidential candidate, or a director of corporations; what physicist can expect his work to be noticed, much less understood, in the New York literary salons? ("A Juggernaut of Words," *Harper's Magazine*, June 1979; pp. 12-13).

Conditions have hardly improved three decades later. Now in the supposed "Information Age" six out of ten American households do not purchase a single book and one-half of American adults do not read one. Forty-three years ago in 1965 when the Gallup Organization asked young people if they read a daily newspaper, 67 percent said yes; in 2006, according to the NORC General Social Survey, only 11 percent of those 18-24 answered affirmatively. And yet "they" say we are saturated with informational overload!

I am most interested in the potential of this cyberspace medium to inform and to generate discourse, to enhance information literacy, and to truly be a "theater of ideas." This site features commentary, data analyses (hey, we've become a "factoid" culture), occasional essays, as well as the requisite links, put together for courses taught by myself and my colleagues. Additions and updates are made daily. If you do give feedback on one of the message pads scattered across these pages and wish a reply, please include your e-mail address.

And now for some sites to stimulate the <u>sociological imagination</u> (or, at a minimum, prepare one for <u>Sociology Jeopardy</u>).

# Adjust/Update Schedule

#### • Ideas:

- Every other year have students check our list for new sites
- Prioritize organizations with no known site for rechecking
- I'll also add sites as they become known through campus listservs and word of mouth

# Current "Outreach" Mostly Passive and Hidden (on Libguide and Archon Finding Aid)



The 1923 Pep Squad led cheers during events on campus. For more information on this photograph and to browse others, please visit our Historic Images Archive.

#### Additional Digital Collections at Trinity University

#### Historic Images Archive

The historic images in the collection range in date throughout the University's first 100 years and are divided into 4 subject categories (based on the University's geographic location): the Tehuacana Images (1869-1902), the Waxahachie Images (1902-1942), the Woodlawn: San Antonio Images (1910-1952), and the Skyline Campus: San Antonio Images (mid 1950s-present).

#### Theatre Arts Images Archive

Trinity's outstanding Drama Program has continued to preserve and collect images of theatrical productions from the last 50 years of its history.

#### Digital Trinitonian

The digitized version of the campus newspaper. Available dates include some issues from 1919-1943 and 1959-present.

#### Digital Mirage

The digitized version of the yearbook of Trinity University. Available years include 1912-1916, 1919-1942, 1944-1978, 1980 to present.

#### Digital Commons @ Trinity

The Digital Commons is a growing collection of scholarly and other materials produced by Trinity University faculty, students, and staff. Trinity's faculty passed an open access policy taking effect in 2010; consequently, many of their journal articles and other publications will be added to the Digital Commons and be freety available to researchers and educators throughout the world.

#### Web Archiving @ Trinity

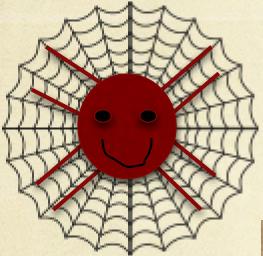
We are in the process of archiving Trinity's web presence including administrative units, academic departments, student organizations, and university publications and events. We will crawf most sites on an annual or semi-annual basis in order to capture Trinity's web presence for future scholarship. More information about our web archiving policy can be found under the Policies and Procedures tab.

# Future Outreach Ideas

- Reach out to campus through University Librarian to
  - Advertise service
  - Solicit potential sites for crawling
  - Suggest potential uses to users
    - o Classes
    - Historical Info for Clubs
  - Solicit suggestions for further outreach
- Reach out through our Special Collections blog and newsletter
- Create and share video tutorial on searching our archived sites

# Spidey made it! And you can

too...



Email me: mtoups@trinity.edu



## Photo credits

Slide 3: Kitty Scared by InertiaCreeps (CC BY-NC 2.0), https://www.flickr.com/photos/inertiacreeps/407007831/

Slide 5: Puzzle by Olga Berrios (CC BY 2.0), https://www.flickr.com/photos/ofernandezberrios/2720567252/

Slide 6, 12, 25, 31: A very tall escalator by Angelina Earley (CC BY-NC-ND 2.0), https://www.flickr.com/photos/daedrius/3517847023

Slide 18: Headache = lame[Day225]\* by Chapendra ((CC BY-NC 2.0), https://www.flickr.com/photos/chaparral/2757495001/

Slide 19: The camera and the picture it took by Theen Moy (CC BY-NC-SA 2.0), https://www.flickr.com/photos/theenmoy/8444583248/

Slide 28: Online Communities 2 by Randall Munroe (CC BY-NC 2.5), http://xkcd.com/802/

Slide 41: You Can Do It Lake Huron Beach Oscoda Trip 9-25-09 16 by Steven Depolo (CC BY 2.0), https://www.flickr.com/photos/stevendepolo/3968766889/