



STANFORD UNIVERSITY LIBRARIES

# Collection Development for Selective Web Archiving

Nicholas Taylor ([@nullhandle](#))  
[Web Archiving Service Manager](#)  
[Stanford University Libraries](#)

[Archive-It Partner Meeting](#)  
August 2, 2016



# the Web is on fire



“Forest wildfire” by [Project LM](#) under [CC BY-NC-ND 2.0](#)



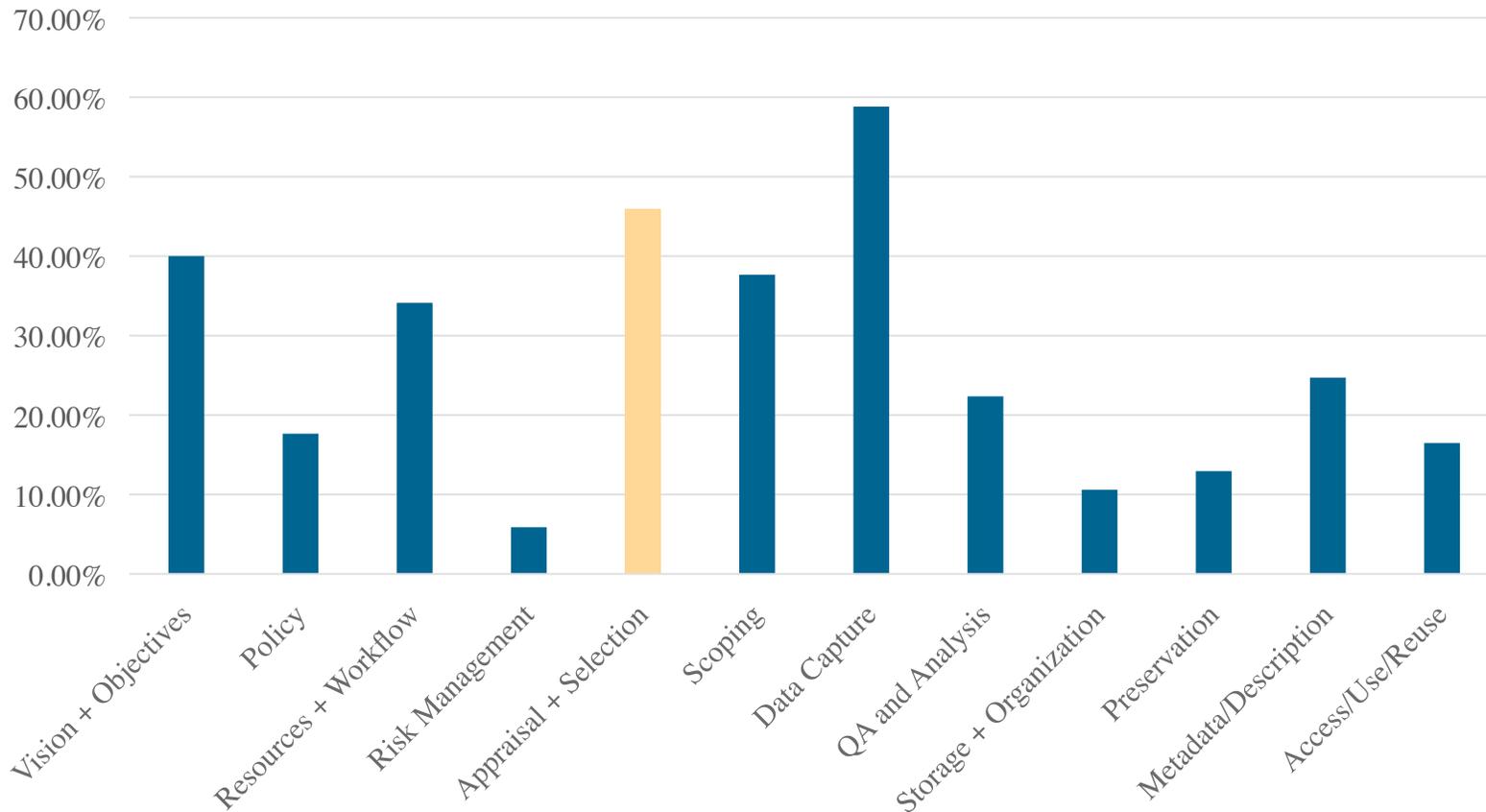
# what are we going to save?



“20130809-FS-LSC-0607” by U.S. Department of Agriculture under [CC BY 2.0](#)



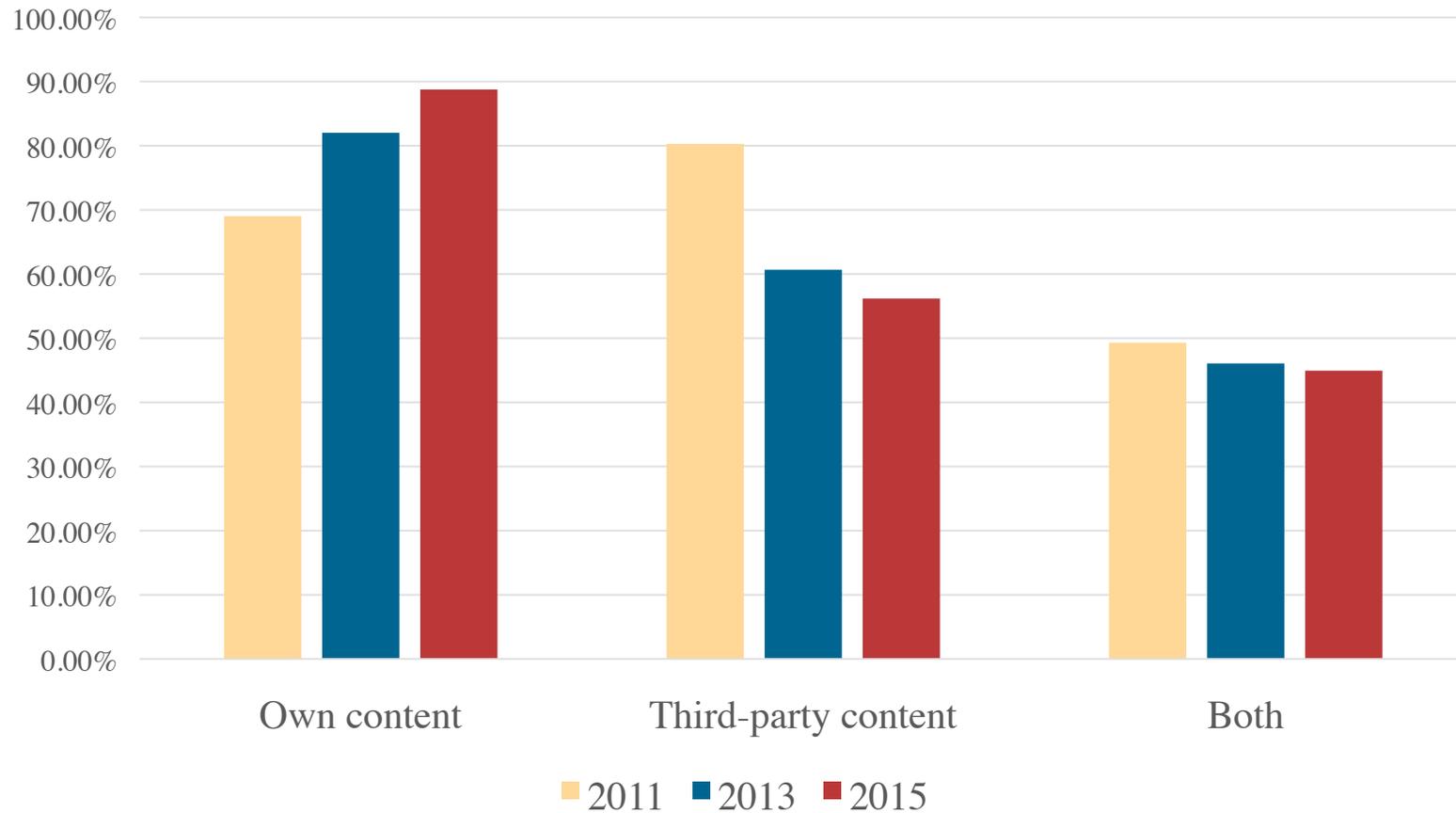
# an area of perceived progress



[NDSA](#): “2015 NDSA Web Archiving Survey”



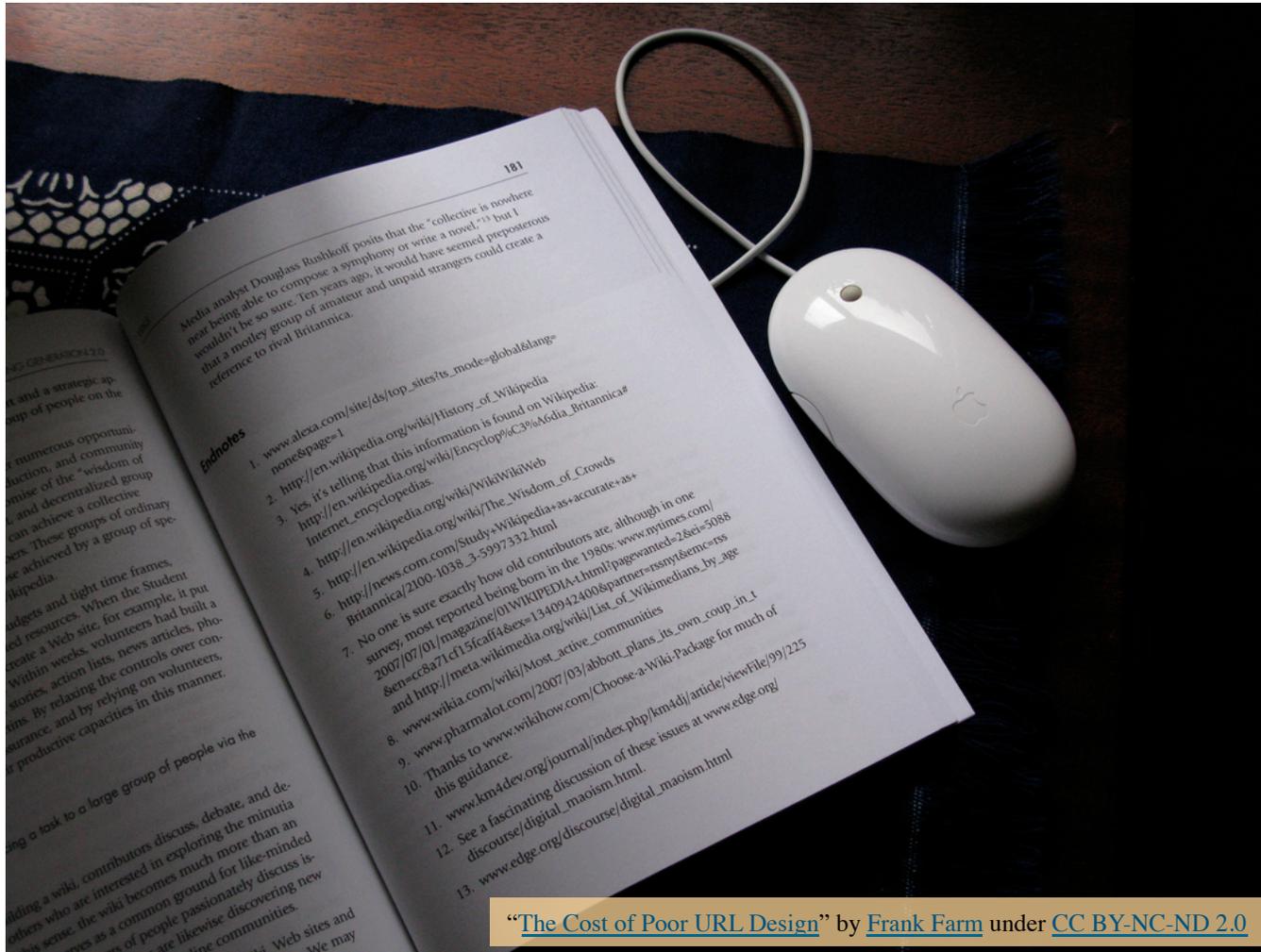
# growth in archiving own content



[NDSA](#): "2015 NDSA Web Archiving Survey"



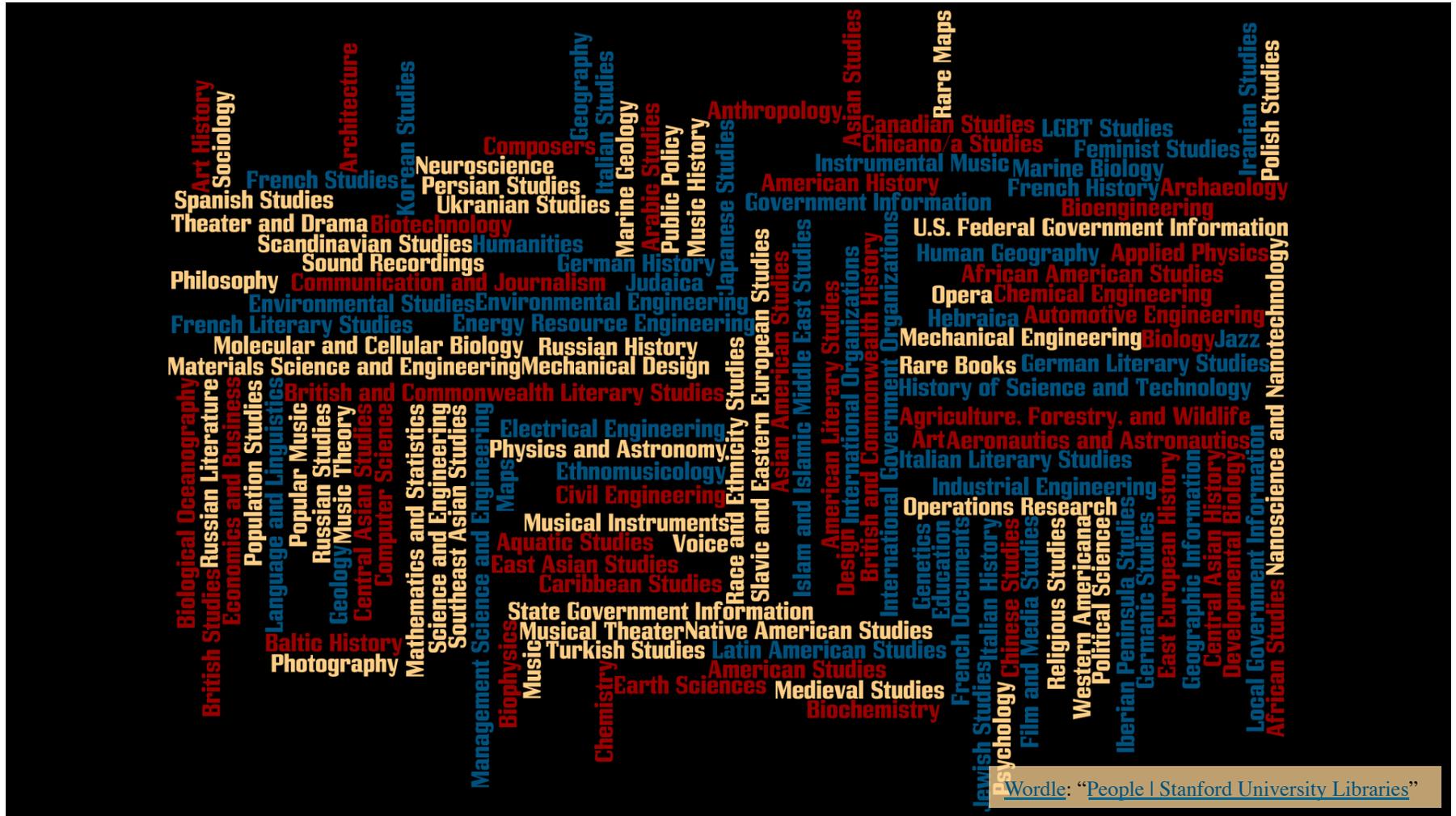
# traditional + web content collecting



“The Cost of Poor URL Design” by Frank Farm under CC BY-NC-ND 2.0



# subject expertise





# necessary but not sufficient



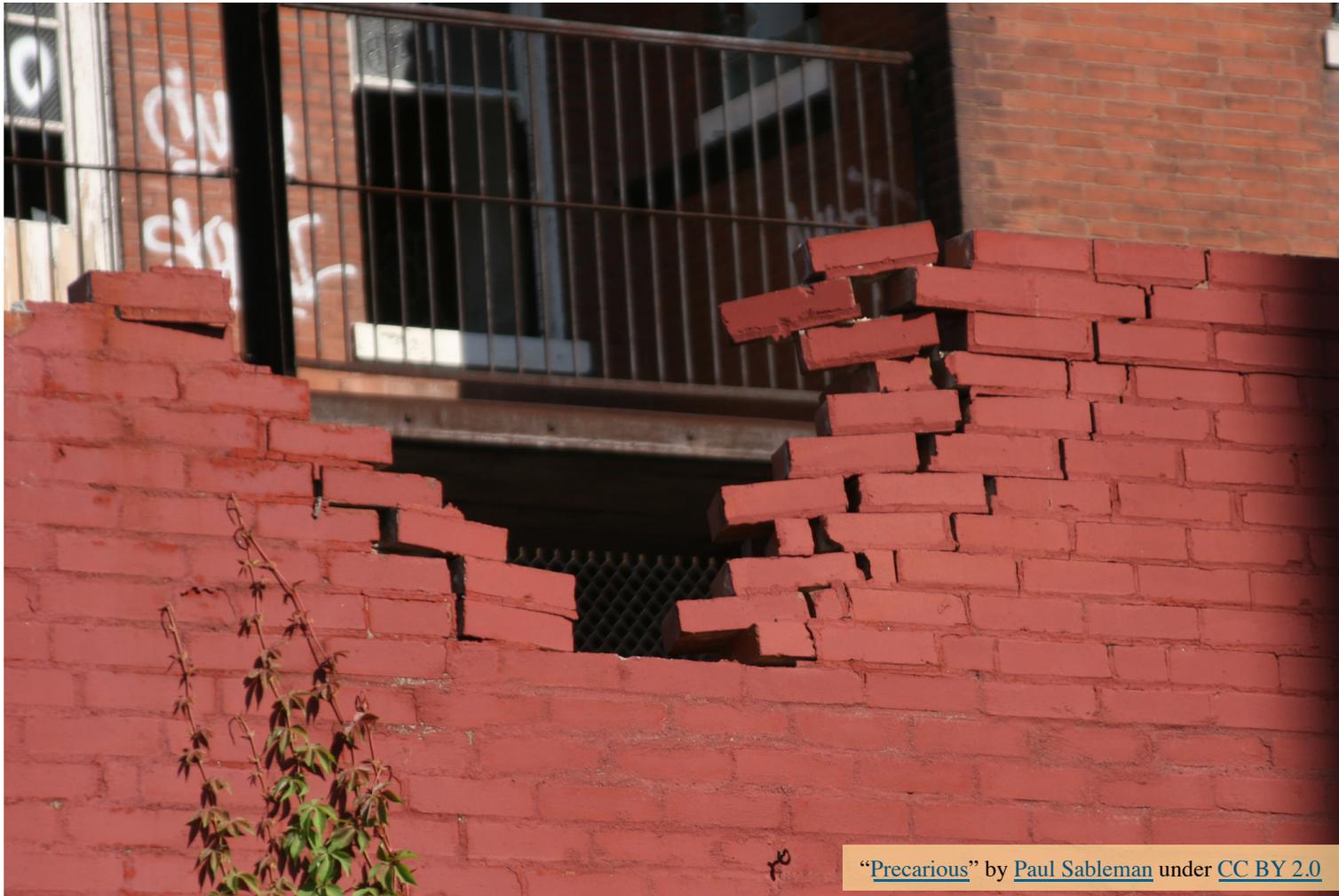
THE TAMIMENT LIBRARY &  
ROBERT F. WAGNER LABOR ARCHIVES

*“In principle, the **collection development policy** for the Tamiment Library’s Web Archive **parallels** that of the Tamiment Library as a whole (labor and radicalism)”*

*In practice, this is complicated by (a) the **enormous size and variety of born digital materials** within Tamiment’s collecting scope...and (c) **resource restraints**. Thus the Library will not only have to **carefully appraise** materials, but to **set priorities and limitations**.”*



# focus on at-risk content



“Precarious” by [Paul Sableman](#) under [CC BY 2.0](#)



# complement collecting strengths



“Symbiosis” by [John Spaderuiz](#) under [CC BY 2.0](#)



# observe resource constraints



“Abandoned” by [Daniel D'Auria](#) under [CC BY-SA 2.0](#)



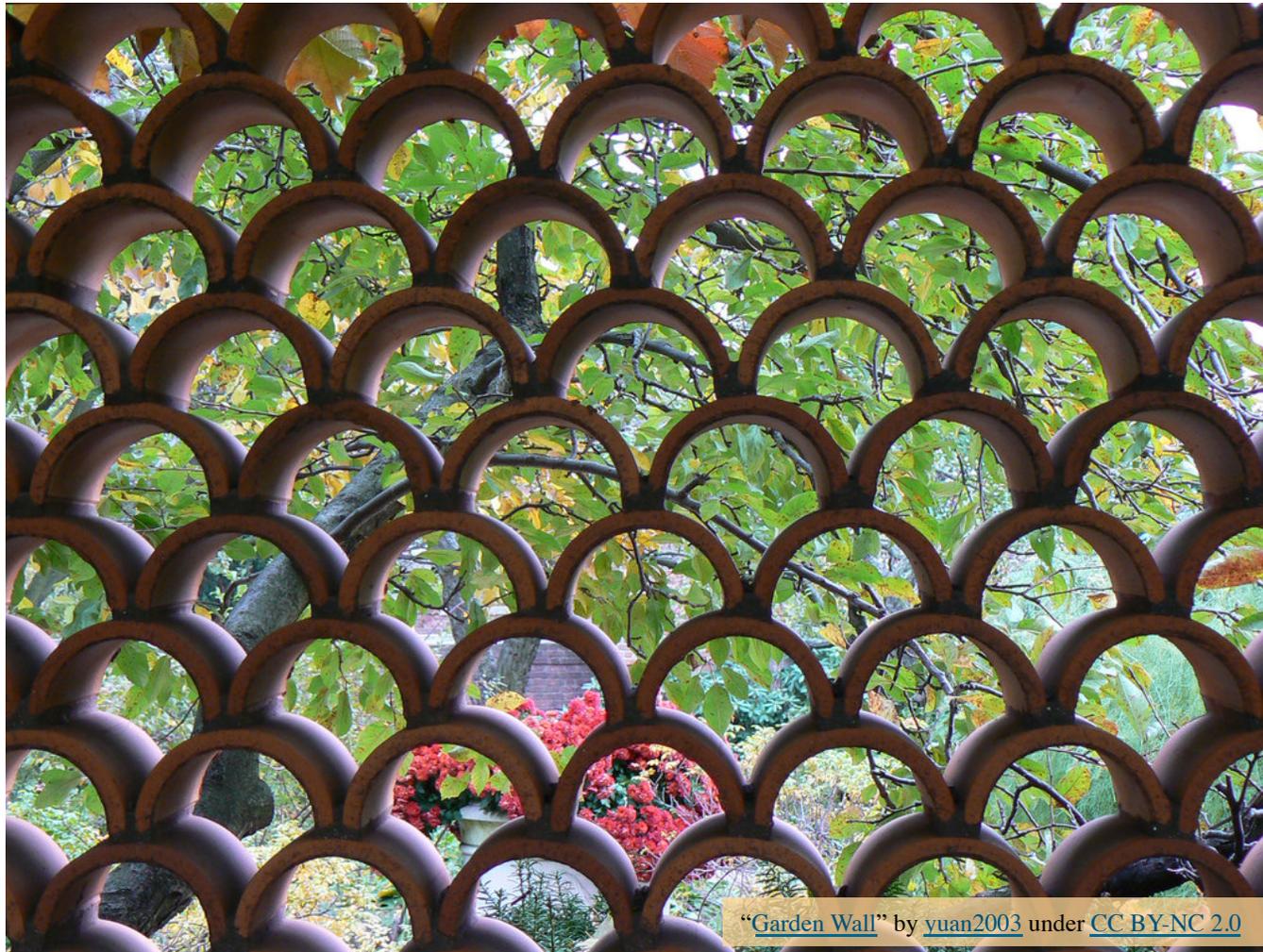
# consider what others are collecting



[“2009 san diego comic-con: comics, still an elemental part of the con”](#) by [george ruiz](#) under [CC BY 2.0](#)



# consider others' access restrictions



“Garden Wall” by [yuan2003](#) under [CC BY-NC 2.0](#)



# assess value to researchers

October 2014 entities – Columbia U.’s Human Rights collection (~4M docs; 330GB)

## TOP 6 People Entities (of 1,887,523)



Barack Obama

Ilham Tohti

Julian Assange

George W. Bush

Pu Zhiqiang

Bashar al-Assad

## TOP 6 Organizations Entities (of 1,335,232)



## TOP 9 Locations Entities (of 373,155)





# enable specific research

**Le blog de Jean-Luc Mélenchon**  
**L'ÈRE DU PEUPLE**

**La France insoumise jlm2017.fr**  
Cliquez ici pour appuyer la proposition de candidature de Jean-Luc Mélenchon

**À CRIER**  
Macron abolit, la caravane passe  
Pour l'égalité des droits!  
La révolution citoyenne a déjà commencé!

**VIDEOS**  
Pour un bilan constructif, les responsables de l'opposition sociale et environnementale

**NOUVEAUTES**  
Non à l'aéroport  
Au Sénat, la France s'engage  
Du candidat à candidat

**À FOND**  
MET INTERNATIONALE  
VILLE DE NICE  
La France insoumise jlm2017.fr

**LA FRANCE INSOUMISE**  
Appuyez la candidature de Jean-Luc Mélenchon pour 2017

**MA CHAÎNE VIDEO**  
Discours de 3 juin  
Discours de Jean-Luc Mélenchon

**AGENDA**  
Discours "Place Marie de la France" (Paris) du 20 au 21 juillet 2017

**FACEBOOK**

**Et s'abonner**  
La France insoumise  
Abonnez vous au blog  
Dossiers industriels

**«Jean-Luc Mélenchon | Le blog»**

**ALAIN JUPPÉ 2017**

**« FACE AU TERRORISME, MA LIGNE EST CLAIRÉ : INTRANSGÉANCE ET SANG FROID. »**

**Ensemble, construisons la France de demain**

**Agir contre le terrorisme avec les citoyens d'Alain Juppé**

**ACTUALITÉS**  
L'ambassadeur d'Alain Juppé à Berlin le 18 juillet  
Alain Juppé à Berlin le 18 juillet

**En 2016, nous avons encore le choix.**  
Le 20 et le 21 mai 2016, c'est Alain Juppé!

**JE DOIS**

**«Alain Juppé pour la France»**

**LA FRANCE POUR LA VIE**

**ON REPREND LA ROUTE REJOIGNEZ-NOUS !**

Adresse e-mail  
Cliquez ici pour **JE REJOINS !**

**REJOIGNEZ NOS ÉQUIPES ENGAGEZ-VOUS**

**NOS CÔTÉS** → **AUTOUR NICOLAS SARKOZY** →

**ARTICLE**  
Nicolas Sarkozy dénonce les dérives du gouvernement

**LA FRANCE POUR LA VIE !**

**LIRE LA VIDEO**

**NS | LA FRANCE POUR LA VIE**  
ASSOCIATION DE SOUTIEN À L'ACTION DE NICOLAS SARKOZY

SUIVRE NICOLAS SARKOZY

<https://www.sarkozy.fr/>

**MARINE 2017**

Bienvenue sur Marine2017.fr

**ARTICLE**  
Marine Le Pen dénonce les dérives du gouvernement

**LA FRANCE POUR LA VIE !**

**SUIVRE MARINE LE PEN**

**«Marine Le Pen 2017»**



# consider appropriate archiving tool



[“Fruit Picker”](#) by [Naoto Sato](#) under [CC BY-NC-SA 2.0](#)



# save content, not links



“Signs!” by [Brian Rawson-Ketchum](#) under [CC BY-SA 2.0](#)



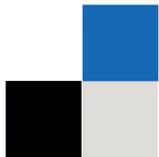
# prefer current, esoteric content



79%



16%



del.icio.us

68%



19%



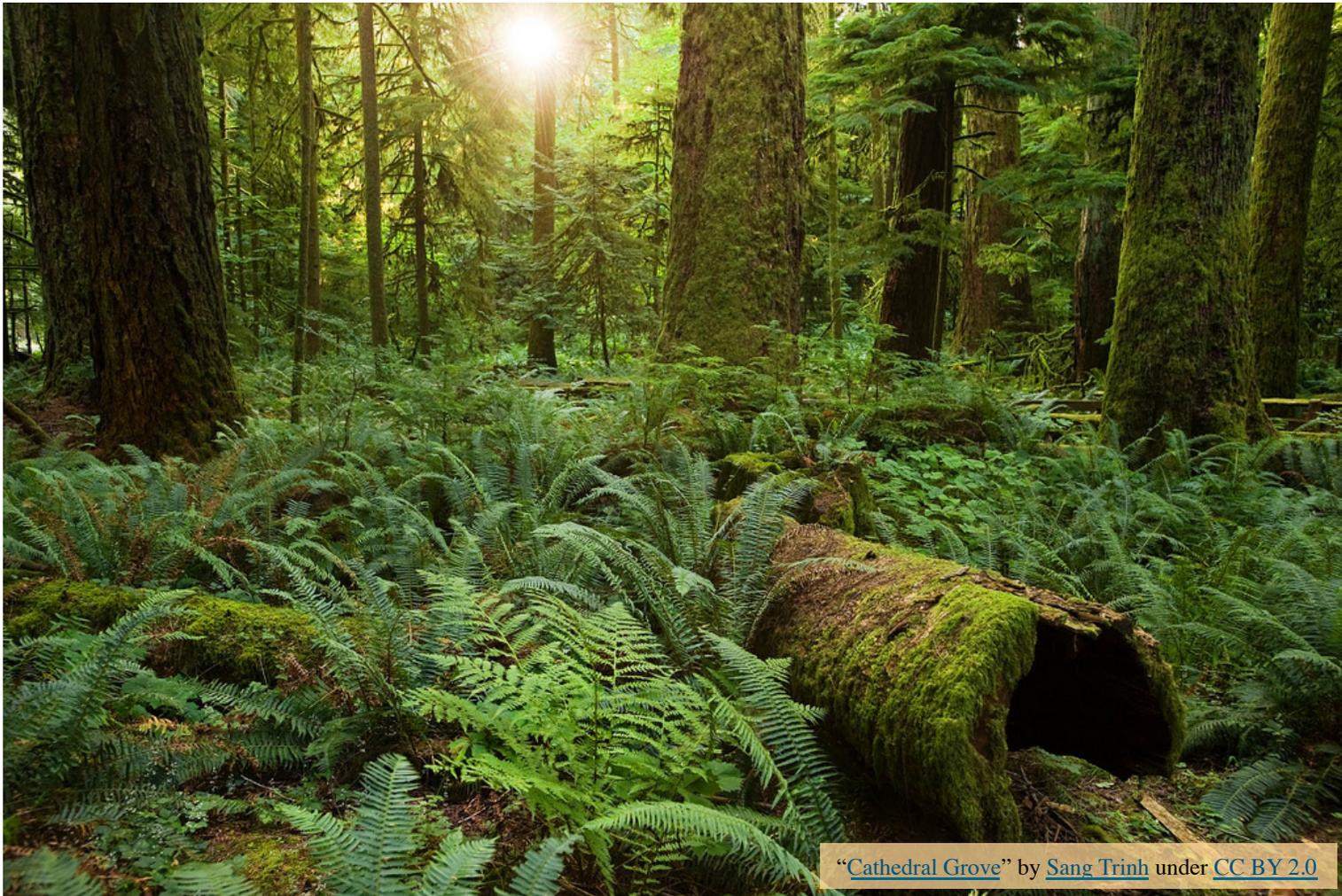
# support community self-archiving



“traveling Pantry community workshop” by [Dan Thompson](#) under [CC BY-NC-ND 2.0](#)



# together we can preserve the Web



“[Cathedral Grove](#)” by [Sang Trinh](#) under [CC BY 2.0](#)