STANFORD UNIVERSITY LIBRARIES

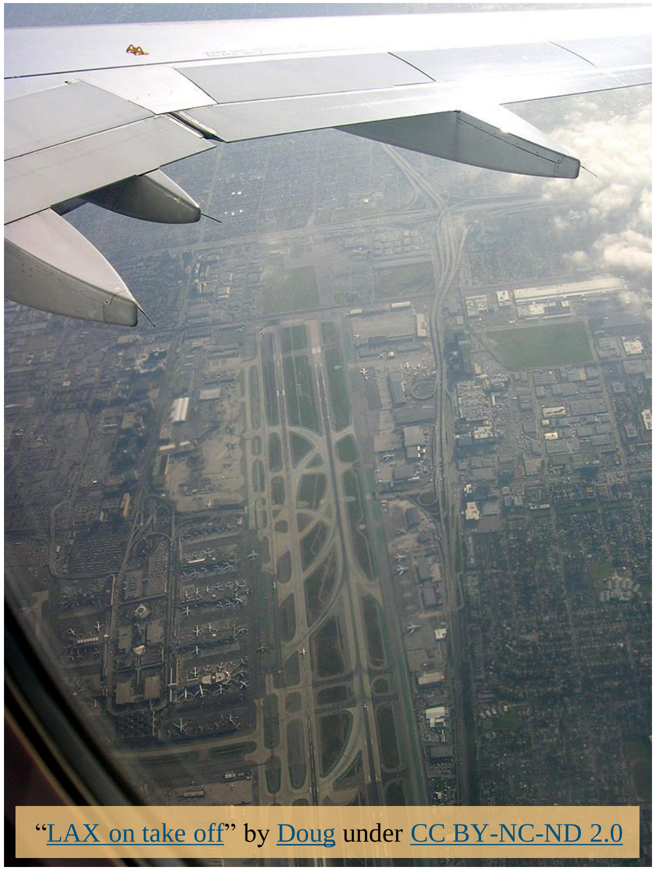
SUL Web Archiving Before and Beyond the CDL WAS Transition

Nicholas Taylor
Web Archiving Service Manager
Stanford University Libraries

IIPC General Assembly: Collaborative Initiatives and Opportunities in the Wake of the CDL WAS Transition
April 12, 2016

overview

- Stanford Web Archiving
- CDL WAS transitioning
- A more collaborative future



"LAX on take off" by Doug under CC BY-NC-ND 2.0

STANFORD WEB ARCHIVING

"Stanford Dish" by Ed Bierman under CC BY 2.0

web archiving activities

- LOCKSS

1999 – present

- WebBase

2001 – 2012

- Archive-It

2007 – present

- CDL WAS

2008 – 2015

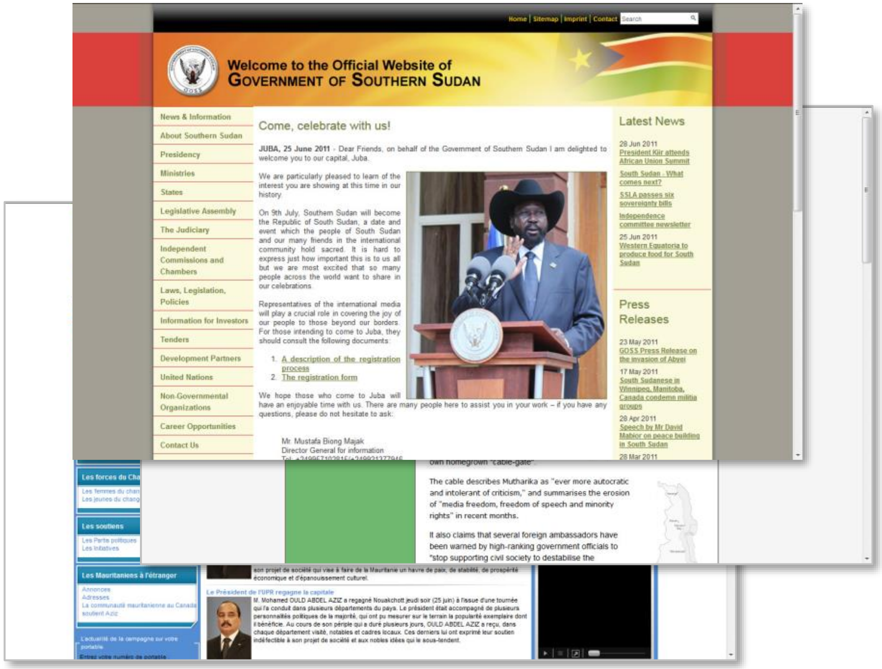
Middle East Politics collection

- duration: 2008 – 2015
- size: ~10 TB
- count: 185 websites
- contents: blogs, political orgs, NGOs

جمعية العمل الإسلامي
Islamic Action society
FREE
Al Mahfoodh
الحرية للامين العام لجمعية العمل الاسلامي
أمل - أول حزب سياسي يحاكم عسكريا بالبحرين والمنطقة
اضغط للتحميل
28-29 NİSAN'da
MUĞLA'DA

African Politics collection

- duration: 2008 – 2015
- size: ~15 TB
- count: 199 websites
- contents: campaigns, news, political parties


Home | Sitemap | Imprint | Contact
Search
GOSS
News & Information
About Southern Sudan
Presidency
Ministries
States
Legislative Assembly
The Judiciary
Independent Commissions and Chambers
Laws, Legislation, Policies
Information for Investors
Tenders
Development Partners
United Nations
Non-Governmental Organizations
Career Opportunities
Contact Us
Les forces du Cha
Les femmes du chan
Les jeunes du chang
Les soutiens
Les Partis politiques
Les Initiatives
Les Mauritaniens à l'étranger
Annonces
Adresses
La communauté mauritanienne au Canada soutient Aziz
L'actualité de la campagne sur votre portable
Entrez votre numéro de portable

Welcome to the Official Website of Government of Southern Sudan

Come, celebrate with us!

JUBA, 25 June 2011 - Dear Friends, on behalf of the Government of Southern Sudan I am delighted to welcome you to our capital, Juba.

We are particularly pleased to learn of the interest you are showing at this time in our history.

On 9th July, Southern Sudan will become the Republic of South Sudan, a date and event which the people of South Sudan and our many friends in the international community hold sacred. It is hard to express just how important this is to us all but we are most excited that so many people across the world want to share in our celebrations.

Representatives of the international media will play a crucial role in covering the joy of our people to those beyond our borders. For those intending to come to Juba, they should consult the following documents:

1. A description of the registration process
2. The registration form

We hope those who come to Juba will have an enjoyable time with us. There are many people here to assist you in your work – if you have any questions, please do not hesitate to ask:

Mr. Mustafa Biong Majak
Director General for information
Tel: +249957102815/+249921277946

Latest News

28 Jun 2011
President Kiir attends African Union Summit

South Sudan - What comes next?

SSLA passes six sovereignty bills

Independence committee newsletter

25 Jun 2011
Western Equatoria to produce food for South Sudan

Press Releases

23 May 2011
GOSS Press Release on the invasion of Abyei

17 May 2011
South Sudanese in Winnipeg, Manitoba, Canada condemn militia groups

28 Apr 2011
Speech by Mr David Mabior on peace building in South Sudan

28 Mar 2011

own homegrown "cable-gate".

The cable describes Mutharika as "ever more autocratic and intolerant of criticism," and summarises the erosion of "media freedom, freedom of speech and minority rights" in recent months.

It also claims that several foreign ambassadors have been warned by high-ranking government officials to "stop supporting civil society to destabilise the

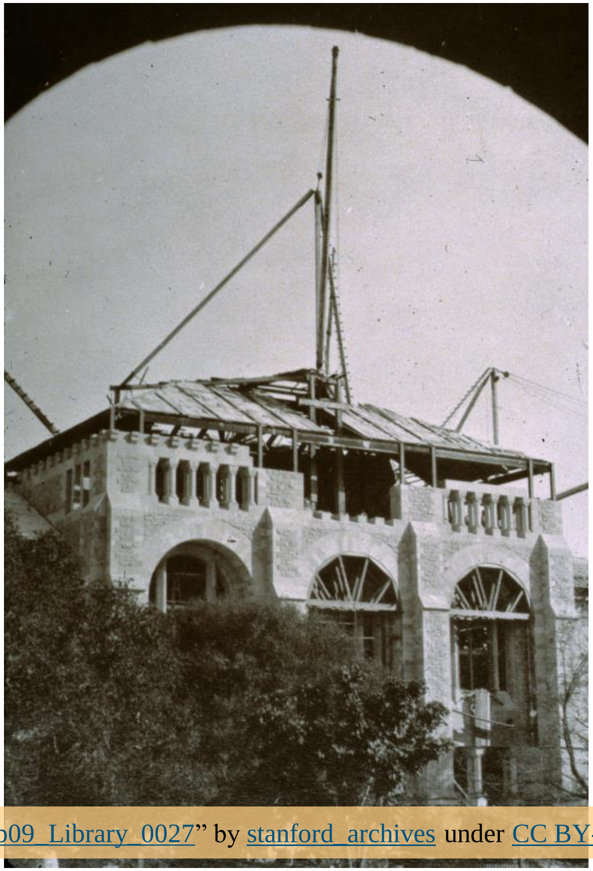
son projet de société qui vise à faire de la Mauritanie un havre de paix, de stabilité, de prospérité économique et d'épanouissement culturel.

Le Président de l'UPR regagne la capitale

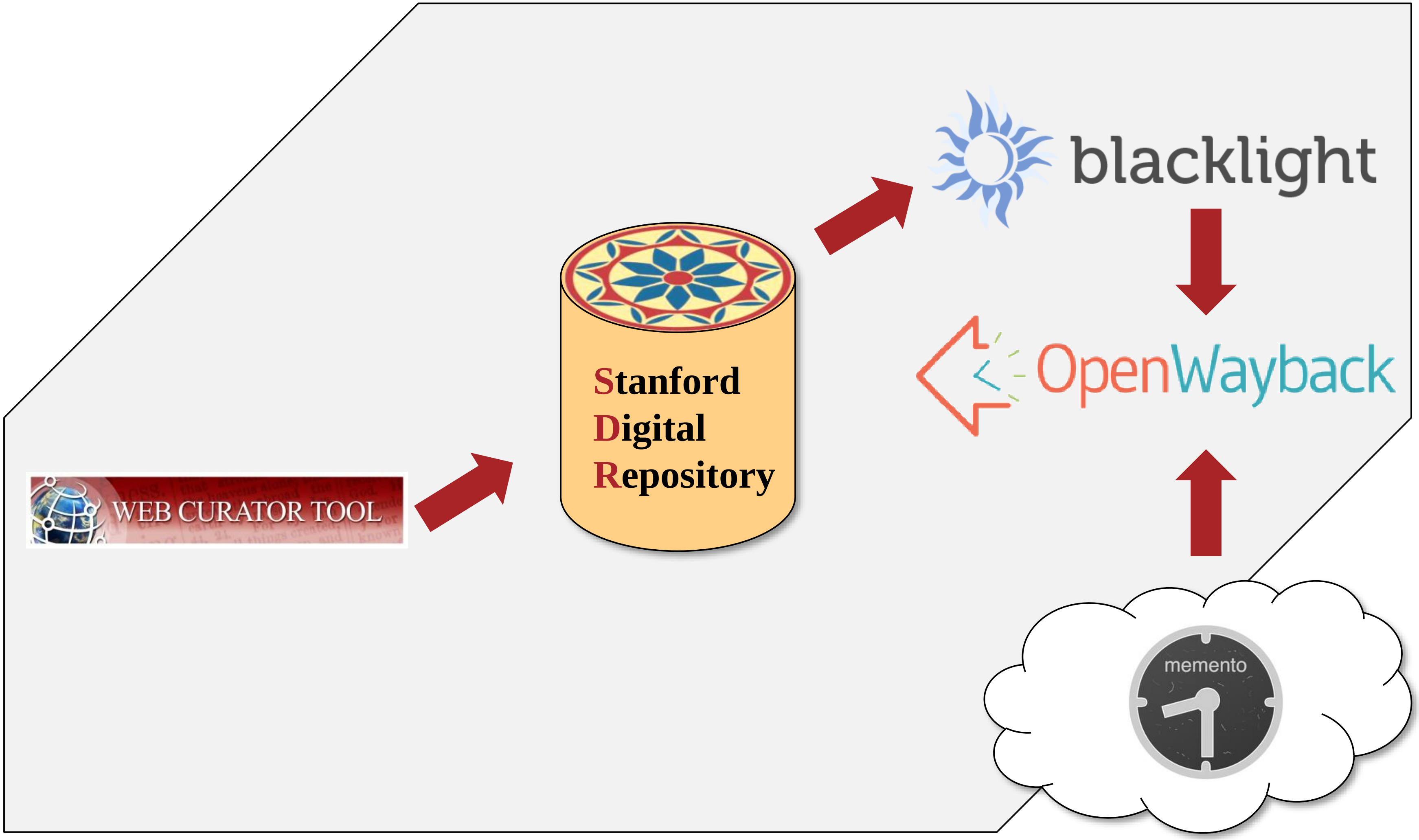
M. Mohamed OULD ABDEL AZIZ a regagné Nouakchott jeudi soir (25 juin) à l'issue d'une tournée qui l'a conduit dans plusieurs départements du pays. Le président était accompagné de plusieurs personnalités politiques de la majorité, qui ont pu mesurer sur le terrain la popularité exemplaire dont il bénéficie. Au cours de son périple qui a duré plusieurs jours, OULD ABDEL AZIZ a reçu, dans chaque département visité, notables et cadres locaux. Ces derniers lui ont exprimé leur soutien indéfectible à son projet de société et aux nobles idées qui le sous-tendent.

Digital Library Buildout 2

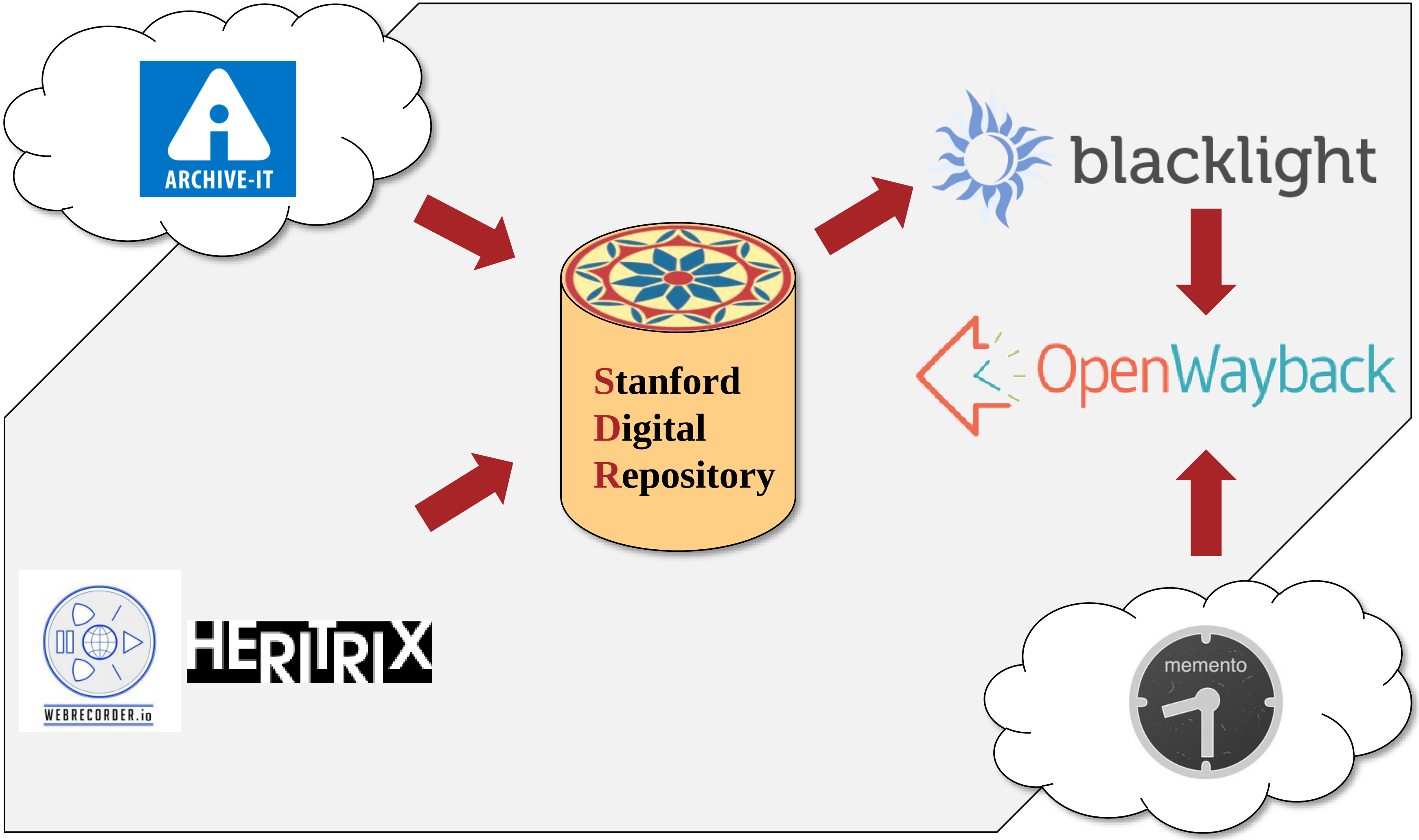
- identify needs
- secure funding
- programmatize
  - staffing
  - use cases
  - policy
  - collection development
  - service model
  - technical architecture



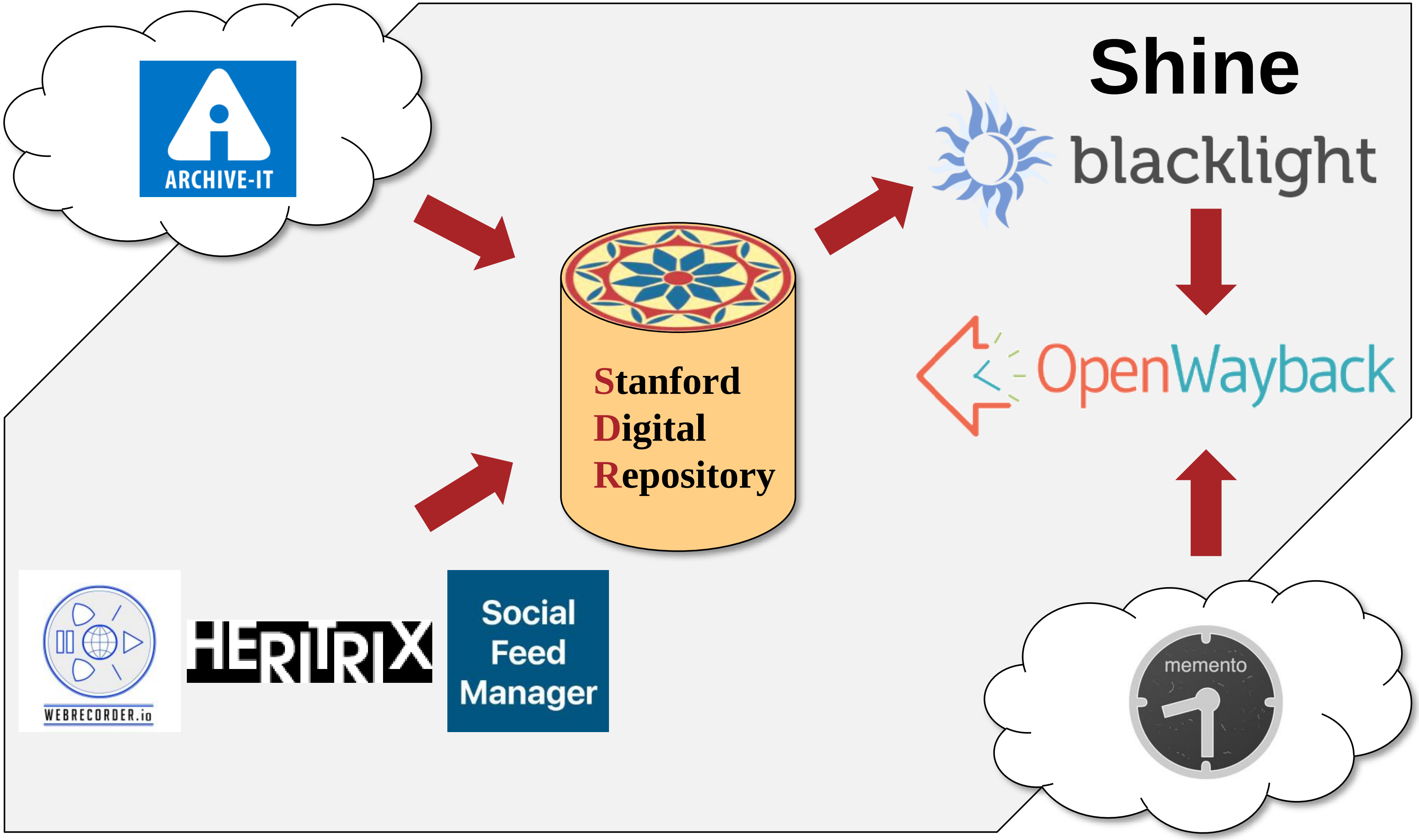
original architecture plan (local)



actual architecture (hybrid)



planned architecture (hybrid)



CDL WAS TRANSITIONING

“Transition” by RicardoRQ under CC BY-NC-SA 2.0

challenges

quality assurance

- backlog
- purge soft 404s

data accessioning

- ingest congestion
- non-working workflows

data transfer

- data volume
- retrieved everything?
- checksums match?

description + discovery

- crosswalk metadata
- improve metadata

challenges

quality assurance

- backlog
- purge soft 404s

data transfer

- data volume
- retrieved everything?
- checksums match?

data accessioning

- ingest congestion
- non-working workflows

description + discovery

- crosswalk metadata
- improve metadata

challenges

quality assurance

- backlog
- purge soft 404s

data accessioning

- ingest congestion
- non-working workflows

data transfer

- data volume
- retrieved everything?
- checksums match?

description + discovery

- crosswalk metadata
- improve metadata

challenges

quality assurance

- backlog
- purge soft 404s

data accessioning

- ingest congestion
- non-working workflows

data transfer

- data volume
- retrieved everything?
- checksums match?

description + discovery

- crosswalk metadata
- improve metadata

A MORE COLLABORATIVE FUTURE

"There's No Place Like The Death Star" by JD Hancock under CC BY 2.0

share collection content

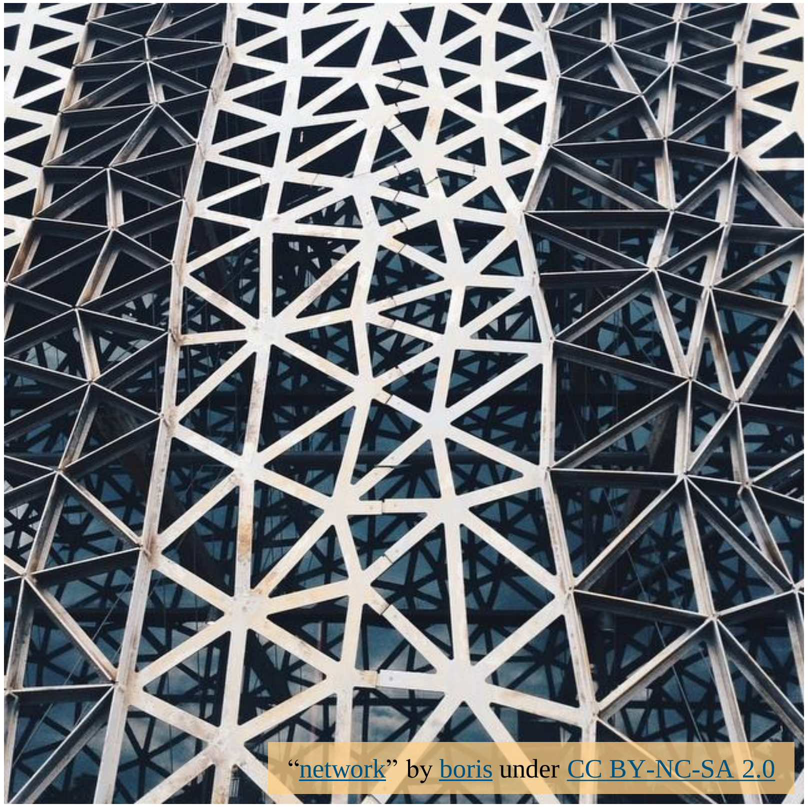
- advantages
  - larger, unified collection(s)
  - distributed preservation
- challenges
  - missing/mixed provenance
  - institutional ownership
  - ad hoc data transfer
  - redundant effort
- opportunity: data transfer APIs (WASAPI)

collaborative collecting

- advantages
  - distribute curation costs
  - more comprehensive collection
- challenges
  - curatorial roles
  - cost sharing
  - institutional ownership
- opportunity: collaborative collecting interface (Cobweb, UNT Nomination Tool)

distributed services

- changing landscape
  - CDL transition
  - Archive-It predominance
  - Harvard environmental scan
- community interest in APIs
- SUL (web archiving + LOCKSS) needs



"network" by boris under CC BY-NC-SA 2.0

let’s combine forces

“Stages of flow” by Peter Thoeny under CC BY-NC-SA 2.0