

Online Heritage Management

From a brand issue to a technical and organizational challenge

“Preserving brand heritage in a digital world”

Foreword

The topic on which this white paper is based is related to a personal interest.

In 2008-2009, alongside other professional duties, I started a small consulting business with an historian, in the field of the *diagnosis and enhancement of corporate cultural assets*. We established the guidelines of an analysis methodology that we entitled Bilan Patrimoine[®]: www.bilanpatrimoine.fr. The methodology has been applied to several consulting jobs, including EDF (analysis of a dam restoration project), Quechua (scenography and exhibit organization at HQ), Lyonnaise des Eaux (analysis of their new corporate civic spirit) and more recently Fenwick (heritage diagnosis).

During spring 2014, I also had the opportunity to give a Master’s course on the subject at a Parisian private business school. The course focused on corporate history and identity, and was shortly followed by an extended white paper on the subject.

These combined experiences led to the main idea I intend to develop and take up to a new challenge in the coming years, convinced that corporate heritage is an essential development issue for many organizations today, *especially in a digital world*, and more specifically from the *brand perspective*: brand awareness and global impact are indeed dramatically sensitive to the *social perception of their cultural background*.

To put this in perspective, this white paper will explore the key issues (in management, organization, and technology) drawn by brand online heritage management, after outlining general keystones and prospects.

Online Heritage Management

From a brand issue to a technical and organizational challenge

“Preserving brand heritage in a digital world”

Table

- I. Keystones and prospects
 - i. Corporate Heritage
 1. Corporate History, Identity and Heritage
 2. An asymptotic relationship
 - ii. Corporate Cultural Assets
 1. Corporate heritage issues
 2. Heritage qualification and maturity stage
 - iii. Brand Online Heritage Management
 1. Inputs
 2. Outputs
- II. Management: Building an online heritage strategy
 - i. The role of auditing
 - ii. A Unified Heritage Management System
 1. A systemic approach
 2. Involving an Heritage Committee
 3. A public-oriented example: the Web-At-Risk project
- III. Organization: Developing an archive policy
 - i. Three major actions
 - ii. Major matters involved
- IV. Technology: Writing the right plan
 - i. Structure
 - ii. Tools and software
- V. Annexes
 - i. Organization chart (Summary graph)
 - ii. Online references

I. Keystones and prospects

Section summary

- Corporate cultural heritage will take its place somehow as a bridge between history and identity, and also as an independent set of forces capable of offering the most coherent image of various identity components.
- Corporate heritage is a complex aggregate made of specific cultural assets provided over time by collective memory as well as production activity. These cultural assets can be tangible (machinery and tools, places like head-quarters or plants...) or intangible (ethical values, expertise...).
- Corporate heritage and brand heritage are necessarily correlated, corporate heritage standing as a step toward brand heritage qualifying processes, thus leading to sustainable brand preservation and identity improvements.
- For our purposes we will define Online Heritage Management (OHM) as: an integrated approach to managing a brand's or company's digital cultural heritage. OHM will involve technology to organize and automate the preservation of online content, and optimize the resulting web archives for corporate and branding needs.
- Brand preservation and identity is the OHM output likely to concern most of the organization's departments. (Shortly followed by Corporate governance and Corporate strategy)
- The board is the organization's department likely to be concerned by most of the OHM outputs. (Shortly followed by Marketing & Communication and Sales & Production)

i. Corporate Heritage

Corporate identity enhancement, logical follow-up of its global and shared awareness, has been for years one of the most promising challenges for all types of organizations involved in a more and more competitive and multicultural economic world. Understanding corporate identity, and working on it, requires enlightening the key role of one of its essential vector: corporate cultural heritage. But how then take advantage of corporate heritage, how exploit it according to a unified value-creating vision?

Board representatives, top/middle management and consultants are likely to find in corporate heritage the foundations of a forward-looking project, likely to generate development and social link.

1. Corporate History, Identity and Heritage

The corporate identity representations (sensory, ethical, cultural, social, territorial) are enriched by several sources (brand and products, knowledge and expertise, teams...), equally crossed by historical dimension. History, even though answerable to subjective perception and a certain form of corporate politicization, appears indeed as a major entry point to all cultural issues' better

understanding. But to clarify the causality link between history and identity and enhance its exploitability, we need to get through a transitional and structuring stage: there we postulate *the critical usefulness of the corporate heritage concept*.

Corporate cultural heritage will take its place somehow as a bridge between history and identity, and also as an independent set of forces capable of offering the most coherent image of various identity components.

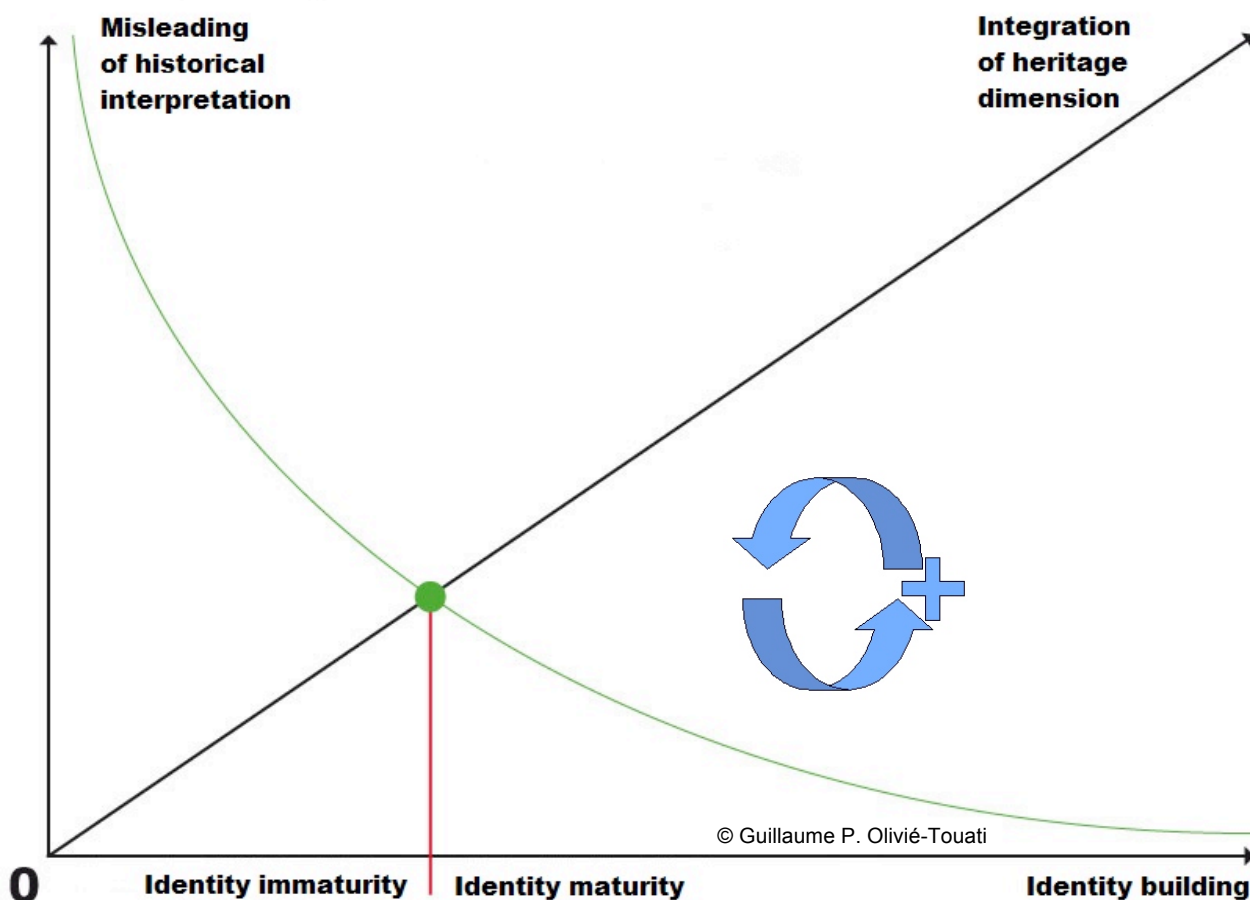
2. An asymptotic relationship

The transition phase between corporate history and identity seems very much like an asymptotic relationship: *identity building increasing indefinitely as misleading of historical interpretation decreases*. Confusion of historical interpretation cannot get to a zero point because there will always be a distortion between tangible reality of past events and their qualitative measurement.

The *integration of heritage dimension* is the theoretical leverage that will help to lower interpretative confusion, while getting through a sort of *equilibrium situation* between:

- **Identity immaturity** (zero to little brand/corporate identity consciousness);
- **Identity maturity** (effective brand/corporate identity consciousness).

After reaching identity maturity stage, a synergic/virtuous circle will reflect the *necessary coherency between collective representation and collective action* from companies' operators and staff.



ii. Corporate Cultural Assets

1. Corporate heritage issues

Relationship between men and heritage is a major civilizational issue.

This relationship has been so far widely unexploited in the corporate world, in spite of an obvious interest considering the highly positive spin-off of its systematic consideration.

The corporate heritage perspective serves to look at corporate organizations not only as economic entities designed to generate positive income statements, but also as meeting points between men, values, places and memory.

Corporate heritage is a high potential resource of growth momentum and positive image for the company, *as potentially a strategic part of brand management, of human resources management and of marketing and communication management.*

It can indeed enhance productiveness and add value at each key phase of the company's life:

- Early stage and creation;
- Development based on external and organic growth;
- Transfer.

Corporate heritage is a complex aggregate made of specific cultural assets provided over time by collective memory as well as production activity. These cultural assets can be tangible (machinery and tools, places like head-quarters or plants...) or intangible (ethical values, expertise...).

Identifying corporate heritage implies both the inventory and the weighted classification of its components. We shall refer to this methodology as "cultural assets ranking". This only, can lead to the building of a rational recovery and valorization plan, which can be organized in three operating processes:

- Cultural animation;
- Cultural mediation;
- Communication.

Corporate as well as brand heritage valorization relies principally on two types of supplementary and obviously combinable actions: events, and publishing.

Corporate heritage and brand heritage are necessarily correlated, corporate heritage standing as a step toward brand heritage qualifying processes, thus leading to sustainable brand preservation and identity improvements.

2. Heritage qualification and maturity stage

Three representative curves equally proceed from the company's "cultural big bang", starting from the early creation stage and evolving synergistically:

- Historical interpretation
- Heritage qualification
- Identity building

These curves' theoretical crossing point (in a 5 to 10 years process) calls upon a maturity stage, a crucial crystallization step where:

- Corporate cultural assets get fully measurable and exploitable for brand-related purposes;
- Heritage operations can be implemented, either by in-house teams or by a third party.

Several technical skills (evenly regarding analysis, classification, protection and valorization) can therefore be put into action and produce positive effects on each type of corporate heritage.

Because of its exponential development potential we will focus on the specific needs of brand online heritage, and even more specifically heritage assets derivate from social networks and all type of user-generated digital content.

iii. Brand Online Heritage Management

Since about 10 years, the web revolution announced by Darcy Dinucci in early 1999 (...“The first glimmerings of Web 2.0 are beginning to appear”...) has clearly changed the way we see and use information and communication technologies: the top social media platforms (LinkedIn – 2003; Facebook – 2004; Twitter – 2006; Tumblr – 2007; Pinterest – 2010; Instagram – 2010; Google Plus – 2011...) have already contributed to generate more online published content (from personal as well as corporate sources) than the traditional websites taken as a whole.

No surprise when 29% of the world total population has active social media accounts! (January 2015 – source: WeAreSocial)

This online content (websites and social media) is of primary importance from a double perspective: first, it leads the global business side of digital information (today more than 70% marketers use social media to gain new customers); second, it is the most powerful resource of newly produced corporate and brand-related cultural heritage.

We already have described the weight of corporate heritage and its links to economic development and long-term stabilized identity: in a digital world, this issue must urgently be taken in account, by all organizational and technical means, the reason being that non-archived and non-protected online content is in real danger and will presumably get for ever lost.

“Bit rot” is the name for this more-or-less slow depravation of a company’s digital corporate heritage: a wave loss that will first affect the brand and its artifacts then indirectly hit the core financial assets as well as the firm’s development and recovery capabilities.

An extensive 24/7 use of IBM Social Media tools (or the like) wouldn’t even prevent a company from bit rot: web archiving has to be taken most seriously and as soon as its awareness arises.

For our purposes we will define Online Heritage Management (OHM) as: an integrated approach to managing a company’s digital cultural heritage. OHM will involve technology to organize and automate the preservation of online content, and optimize the resulting web archives for corporate and brand needs. We will use the acronym “OHM” in the following development.

1. Inputs

Proactive web archiving and more explicitly native format web archiving is the adequate response to bit rot. This response should be applied in a commercial way and according to private and specific needs: the Internet Archive, with its publicly minded approach, being of practically no help to ensure the effectiveness of *brand online heritage preservation*, even so it does provide an access to about 432 billion web pages stored.

On the other hand, the International Internet Preservation Consortium, a global web archiving experts' network, can be a very useful resource for companies' staff and consultants.

In the quest for a most complete range of web-based Electronically Stored Information (ESI), we will take the following extensive inputs in consideration:

- Original non-modified native format of websites; (accounts, interactive pages, and rich media)
- Public and private social media; (Facebook, Twitter, LinkedIn, Google+, and the like)
- Collaborative systems; (SharePoint, Wikis, and the like)
- Private social networks; (Chatter, Jive, Yammer, and the like)
- Structural and descriptive metadata of content and files.

2. Outputs

OHM outputs potentially include crystal-clear forward-operating effects on:

- Strategy at the highest level
- Decision-making at every level
- Standard operating procedure in records management, sales, production...

The table below displays a digest of OHM outputs according to their scope of effectiveness (sort by decreasing public visibility) and to their level of influence on main corporate operators: (According to a 2014 Bilan Patrimoine[®] survey)

Scope	OHM output	Departments concerned*					
		Board	Marketing & Communication	Human Resource	Finance	Sales & Production	
Brand	Awareness	3	5	1	1	4	14
	Identity	5	5	3	3	4	20
	Development	4	3	2	3	4	16
	Value	5	3	2	4	3	17
	Preservation	4	5	3	4	4	20
Corporate	Identity	5	3	4	2	3	17
	Social Responsibility	3	4	5	1	3	16
	Legal	4	2	4	5	2	17
	Strategy	5	3	4	3	3	18
	Governance	5	4	3	4	2	18
	Social cohesion	4	2	5	1	3	15
	Income & Profit	5	3	1	4	4	17
	Business value	5	3	1	3	2	14
		57	45	38	38	41	

* Ranking from 1 to 5:

- 1 - Very limited to no interest
- 2 - Limited interest
- 3 - Certain interest
- 4 - High interest
- 5 - Very high interest

We can infer from this outputs digest that:

1. Brand preservation and identity is the OHM output likely to concern most of the organization's departments. (Shortly followed by Corporate governance and Corporate strategy)
2. The board is the organization's department likely to be concerned by most of the OHM outputs. (Shortly followed by Marketing & Communication and Sales & Production)

Obviously, this OHM outputs digest addresses only the one-brand corporate organization issue: multiple-brands companies and global corporations are much more complex systems requiring further detailed analysis and auditing.

II. Management: Building an online heritage strategy

After this theoretical approach, we need to apply our findings about online heritage to the area of management at the highest corporate level: that is, building an actual online heritage strategy.

Strategy never comes fully from scratch, especially for long-lasting corporate organizations. It derives from forward-thinking diagnosis, guiding policy and coherent actions.

i. The role of auditing

Auditing is crucial, in the way that it is the most trustworthy solution to ensure reliable objectivity in a cultural field with the highest misinterpretation risks: corporate operators like records managers or digital archive curators (if any) will faithfully execute their duties, but they have to be instructed according to a pre-defined online heritage strategy.

Corporate board and top management will need heritage auditing at 3 levels:

- Initial Audits
 - o To identify the collectable online heritage materials;
 - o To define the online heritage strategy;
 - o To propose the right tools for technical purposes.
- Performance Audits
 - o To verify the success of quantitative online heritage management operations;
 - o To report potential failures in technology use.
- Quality Audits
 - o To verify conformance to web archiving organization standards;
 - o To highlight areas of good practice so as to enhance continual improvement.

In the initial auditing phase, auditors will define the 5 key points of the online heritage strategy, as follows:

- Selection

Even if the broadest and largest crawls have to be performed, the selection issue refers to defining the right types of collections for the right goals.
- Harvest

Collecting data implies capturing the OHM inputs previously mentioned (Section I) with full integrity.
- Preservation

The intent of OHM is to preserve over time the strictly original form of the collected digital content.
- Access

Born-digital collections require specific software tools: extensive consultation can be processed through broad data or archived web sites page-by-page.
- Valorization

Taking advantage of online heritage is OHM ultimate goal, and as such should be extensively explained and legitimate to all corporate operators.

ii. A Unified Heritage Management System

1. A systemic approach

The concomitant understanding of the three corporate organization's cultural dimensions (history, heritage, and identity) necessitates a systemic approach: corporate organizations are indeed complex social entities obeying to the general laws of advanced cultural systems.

To satisfy to our OHM needs, it needs then to use a corporate systemic approach, the only one likely to embrace in the most satisfying way all the relational and organizational issues at work.

This systemic approach is complemented by a *holistic and federalist vision*, leading to an innovative corporate management concept: the Unified Heritage Management System.

2. Involving an Heritage Committee

The first stage of a Unified Heritage Management System is the formalization of a dedicated group of individuals, a Heritage Committee, which transversal missions will be to:

- Unify and execute the online heritage strategy;
- Coordinate in-house heritage operations;
- Respond to audit requests.

The Heritage Committee must include board representatives as well as top management of key corporate departments.

3. A public-oriented example: the Web-At-Risk project

Designed to benefiting librarians and archivists, the US-based Web-At-Risk project is a grant funded by the National Digital Information and Infrastructure Preservation Program, and initiated by the California Digital Library (University of California).

The research and assessment materials produced by Web-At-Risk strengthen and encourage the theoretical concept of a Unified Heritage Management System.

On the practical side, the institutional and technical partners of the Web-At-Risk project can be valuable knowledge sources for private and business operators.

III. Organization: Developing an archive policy

Confronted to its numerous cultural heritage challenges, the modern corporate entity needs to develop a comprehensive online heritage strategy, which will be conducted by the Heritage Committee.

The second OHM stage we are exploring now, after the management area, is in-house organization.

What sort of practical digital archive policy can be generated by online heritage strategy?

The archive policy relies on three major actions, and involves several specific matters:

i. Three major actions

- Ranking archive needs and potential risks;
- Identifying operative protection and exploitation goals;
- Identifying the most realistic mechanisms.

ii. Major matters involved

- What are the firm's most important cultural assets?
- Who generates and controls them in the firm?
- What is the proportion and structure of their online dimension?
- What existing policies are in place to store and provide efficient/productive access to these online assets?
- What level of risk is management willing to accept for each of these assets?

Providing detailed answers to these key questions will be the first step of the Heritage Committee action plan, and will directly lead to build an Acceptable Online Heritage Policy. (AOHP)

IV. Technology: Writing the right plan

Following the online heritage strategy building and the archive policy statements, writing the right technology plan closes up the initial OHM process.

i. Structure

Online Heritage Management relies on two pillars:

- Consulting/auditing; (See Sections II & III)
- Tools/software.

As for tools and software, the global technical reference is until now the WARC format.

WARC stands for “Web ARChive” and combines multiple online resources (the previously described OHM inputs) into an aggregate archival file, together with connected information. “A WARC record consists of a record header followed by a record content block and two newlines; the header has mandatory named fields that document the date, type, and length of the record and support the convenient retrieval of each harvested resource (file). There are eight types of WARC record: 'warcinfo', 'response', 'resource', 'request', 'metadata', 'revisit', 'conversion', and 'continuation'. The content blocks in a WARC file may contain resources in any format; examples include the binary image or audiovisual files that may be embedded or linked to in HTML pages.” (Source: Digital Preservation / Library of Congress – USA)

WARC file format is definitely the technology to take into account when setting up the archive policy derived from the online heritage strategy. WARC is connected to ISO 28500 certification.

For private corporate organizations, this standard ensures reliability and long-term stability of the whole OHM process.

An interesting initiative has to be mentioned here: OpenWayback is an open source project meant to develop the key technology related to the Internet Archive' Wayback Machine and used by numerous web archivists to play back archived websites. Dynamic content such as social media can also be taken in account.

Structuring the most efficient technology plan will require to weigh up all of these issues (file format, certification, open source query and access) and will be of the highest importance as a basis for the following 4 key processes: (SASA)

1. **Selection**;
2. **Acquisition**; (possibly through Heritrix archival web crawler)
3. **Storage**; (possibly through Bagit-conformant packages)
4. **Access**. (Possibly through the Wayback Machine)

ii. Tools and software

In the perspective of building and implementing a comprehensive OHM system, the Heritage Committee should consider combining some of these tools and software:

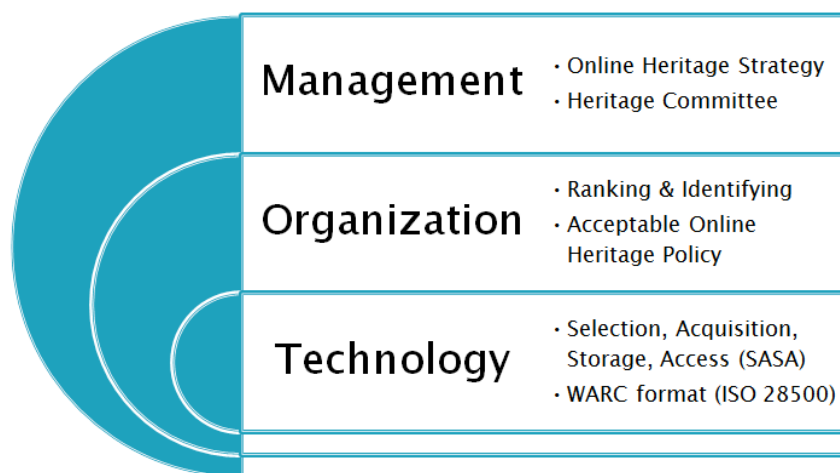
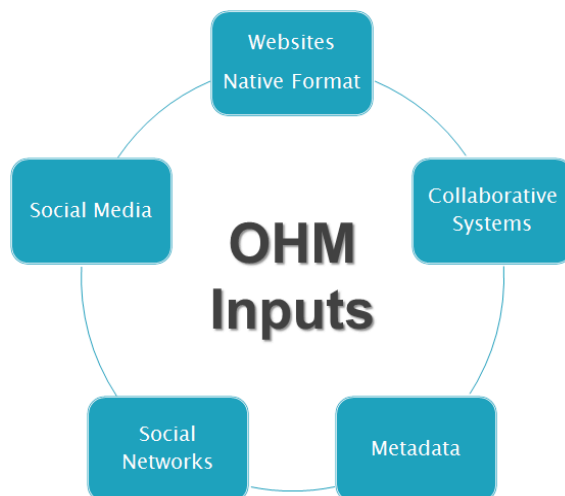
(Further development will have to be considered to patently fulfill brand/corporate needs)

Technology	OHM process	Related OHM inputs
ArchiveFacebook	Harvest (Acquisition)	Social media
Heritrix		Websites
HTTrack		Websites
SiteStory		Transactions between web clients and web servers
WARCreate		Websites
Warrick		Websites
Wget		Websites and social media
Building Collections on the Web		Harvest (Curator tools)
CINCH	Internet-accessible documents	
NetarchiveSuite	Websites and social media	
Web Curator	Websites	
HTTrack2ARC	Preservation (Collection storage and maintenance)	Websites
Java Web Arrchive Toolkit		Websites
JHOVE2		Websites
MediaWiki Memento Extension		Wikis
Web Archive Transformation Format		Websites
Web Archive Transformation Utilities		Metadata
WarcManager		Websites
WARC Tools		Websites and social media
Time Travel Portal		Access and finding aids
Time Travel APIs	Websites and social media	
Memento Time Travel	Websites and social media	
NutchWAX	Websites	
WEb aRchive Access	Websites	
Wayback Machine	Websites	
Xinq	XML database	

V. Annexes

i. Organization chart (Summary graph)

Online Heritage Management



Unified Heritage Management System



ii. Online references

WeAreSocial:

<http://wearesocial.net/blog/2015/01/digital-social-mobile-worldwide-2015/>

Hanzo Archives:

www.hanzoarchives.com/solutions/corporate-heritage/

Internet Archive:

<https://archive.org/web/>

International Internet Preservation Consortium

<http://www.netpreserve.org/>

Electronically Stored Information

<http://www.techopedia.com/definition/1013/electronically-stored-information-esi>

SharePoint

<https://products.office.com/en-us/sharepoint/collaboration>

Chatter

<http://www.salesforce.com/chatter/overview/>

Jive Software

<https://www.jivesoftware.com/>

Yammer

<https://www.yammer.com/>

Web-At-Risk

<http://www.cdlib.org/services/uc3/partners/webatrisk.html>

WARC file format

<http://www.digitalpreservation.gov/formats/fdd/fdd000236.shtmlv>

ISO 28500 certification

http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=44717

OpenWayback

<https://github.com/iipc/openwayback/wiki>

Heritrix archival web crawler

<https://webarchive.jira.com/wiki/display/Heritrix/Heritrix>

Bagit-conformant packages

<http://tools.ietf.org/html/draft-kunze-bagit-08>