

Digital archiving in the BDIC

a working process....



BIBLIOTHÈQUE DE
DOCUMENTATION
INTERNATIONALE
CONTEMPORAINE

Observation

- More and more of traditional donators of BDIC produce born digital archives.



INTRODUCTION

- This presentation **has nothing affirmative**, it is rather about a working progress and our first conclusions after 6 months and before to begin
- In this presentation I am going to concentrate **on the first steps of our workflow**, because the first phase can have strong influence on the following process.
- It will speak essentially about practical and sometimes very concrete points.



Institutionnal and current context in the BDIC



BIBLIOTHÈQUE DE
DOCUMENTATION
INTERNATIONALE
CONTEMPORAINE

Existing systems and tools

- This project must be based on IT solutions that already exist within BDIC .
 - Production research tool in XML -EAD
 - online catalog : CALAMES
 - Digital library: L'Argonnaute
- This project must be integrated within national digital infrastructures proposed to higher education institutions.
- This project should rely on free and open source software



Legal status

Before the technical study , on legal study was conducting by a legal specialist

The principal conclusions are as follows :

- BDIC could acquire private born digital archive on a voluntary basis of archive creators
- for the licensed documents, the access to documents should be restricted and limited to computers in the reading room of BDIC



Reflexion basis



BIBLIOTHÈQUE DE
DOCUMENTATION
INTERNATIONALE
CONTEMPORAINE

- For beginning , we observed and noted some good digital management practices.
 - Huma Num recommandations
 - Participation in the consortium ArcMC
- More to the point, we are also study practices implemented by French national library and national Archives.



FIRST CONSTATATIONS



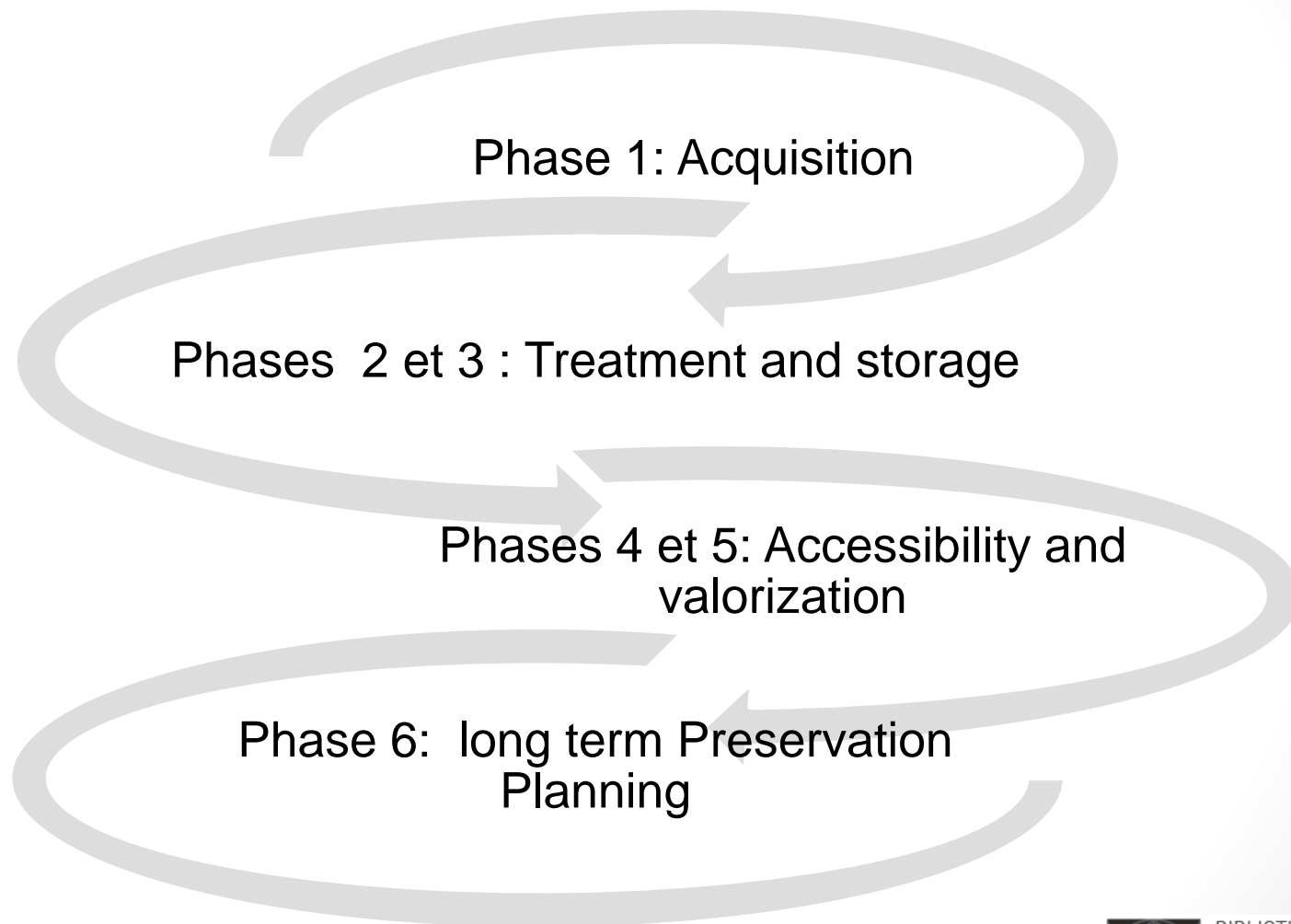
FIRSTLY

The need to quickly define (at least for ourselves)
the main steps of the archival procedure



Proposition :

Archiving process in 6 steps



Observations:

- This process is based on the OAIS standards but does not comply fully with it because BDIC is not yet equipped for a strict control of authenticity and integrity of document
- There is a need to process promptly the archiving and technical treatment for collected documents.
For several reasons;
 - storage
 - Technological obsolescence of formats
 - Relation with creator take benefit of their knowledges
- The first phases is crucial , because it have strong influence on the following process.



SECONDLY,
define

- the type of document
- the type of donation



Type of documents concerned by archiving

Take over

Born digital archives

- .Administrative paper
- . Images
- . Video documents
- . audio documents

Web archives

- .Web sites
- . Mailing list

Not take over (for the moment):

Born digital archives

- .Database
- .mapping database

Archives du web

- .Email inbox
- . Websocial network



Two possible forms of donations

Form 1

Donation is a one time delivery

The born digital archives donation should be delivered in one batch, as snapshot.

Form 2

Donation is regular deliveries

The born digital archives donations will takes the form of several and scheduled deliveries

Each delivery will be dealt with separately one after another.

But it will be described in the same research tool



Different level of collaboration

Processing for one time delivery donation

Two specific cases

1/ Producer is available :

In this case the collaboration is recommended indeed required

2/ Producer is not available

Consequently the collaboration is not possible

Processing for regular deliveries donation

THE COLLABORATION IS necessary and required

Because the treatment of open archival donation requires the implement of good practices for realise a regulal deliveries.



THIRDLY, some « requirements » for deliveries documents

- It's dangerous to accept a digital archive donation without having analyzed its content and in particular :
 - File format
 - File naming
 - Organisation of folder, subfolders and files inside the electronic repository



Answer these questions?
First practical propositions



File format

- In the BDIC, we have defined a typology of 5 file formats.
- This typology should be used to determine if we could be treat the deliveries documents.



Categorization of formats

Catégories	Définition	Actions
Inconnu	Format reçu dont on ignore les caractéristiques techniques, c'est-à-dire non identifié par les outils FITS. Il ne peut donc pas être évalué, ni traiter.	Non pris en charge Les fichiers ne sont pas conservés. A éliminer avant le traitement.
Identifié	Format identifié mais pour lequel la BDIC ne dispose pas d'outils pour accéder aux données qu'il contient. Leur traitement est possible mais par défaut, la BDIC fait le choix de ne pas le traiter.	Non pris en charge Au moment de l'évaluation d'entrée, la liste des fichiers aux formats identifiés a été dressée, 3 options - Soit le producteur peut effectuer une conversion de ces fichiers vers un format connu ou maîtrisé - soit la conversion est réalisée par un tiers et cofinancé avec le déposant - Soit ces fichiers sont détruits au moment du traitement technique.
Connu	Format identifié pour lequel la BDIC dispose d'outils pour accéder aux données qu'il contient et pour les conserver. (outils de lecture et outils de conversion vers un format maîtrisé)	Pris en charge Le contenu du fichier est accessible Il doit être converti vers un format pérenne pour être conservé
Maîtrisé(-)	Format identifié et qui est directement compatible avec les préconisations du CINES, mais dont la BDIC ne peut assurer l'accès.	Pris en charge Contenu du fichier non accessible Il peut être conservé. Pour que le contenu du fichier soit rendu accessible, il doit être traité.
Maîtrisé (+)	Format identifié et qui est directement compatible avec les préconisations du CINES, dont la BDIC peut assurer l'accès.	Pris en charge Contenu du fichier accessible Directement conservable car format pérenne

Define the future use and consultation of documents

- Exemple
 - Conservation for patrimonial purposes may prefer conserving the original form of document
 - Conservation for patrimonial purposes may prefer conserving data content and are less interested by the original form of the document (considered like a support)
- In the BDIC we have chosen to favorise the first option and fix/freeze the entry's form document. (word /pdf).



The choice of technical treatment of file format

- We plan to set up batch conversion process, which allow us to work on many files without having to open them



Collaborate with archives creator : crucial point for facilitate the technical treatment

- We have to inform creators about these different procedures in order to:
 - not to disappoint their expectations
 - to encourage them to collaborate with us for preparation of delivering documents
- We would like to encourage them to check their documents before the delivery.



Another important expectation: The organisation of the directories and the names of the folders and files

- Why ?
 - Directory too poorly organised is a complicating factor for filing plan realisation
 - Files named without signification (e.g. o00001, 00002, etc) could complicate the process (of treatment)



The key to success lies with the
degree of collaboration creators
and BDIC



Depending on the degree of cooperation the BDIC will be able to qualitatively, quantitatively and effectively treat the deliveries documents



FIRST PRATICAL SOLUTIONS



BIBLIOTHÈQUE DE
DOCUMENTATION
INTERNATIONALE
CONTEMPORAINE

Find good practices of creation and organisation of documents are essential

- We are working on different “good practices guidance” for archives creators
 - For file format
 - For file naming
 - For organisation of their digital document
- We propose them different meetings, regular visits for working with creator to the preparation of their delivery directory



Define an entry protocol



Entry protocol in 4 steps

1

First contacts
with creator

2

Delivery of
archives
documents

3

Technical
Evaluation of
documents

4

Preparation of
donation
agreements



Phase 1: First contacts with creators

1. Born digital donation proposition of archives creators (from organisation or individual)

2. Organisation a meeting for discover the content and the format of donation

3. Define a first collections of “eligibility documents”

4. “Archives values “of the donation proposition is evaluated

5. If the director of BDIC formules an positive approval, the second step could begin.

Phase 2: Delivery of documents

2.1 Preparation to documents Transfer

- After having met archives creator a first time, we come back and proposes first recommendations for **preparation of their documents for the transfer** .
- BDIC ask also to complete an **information request form** concerning the content of delivery directory



Phase 2: Delivery of documents

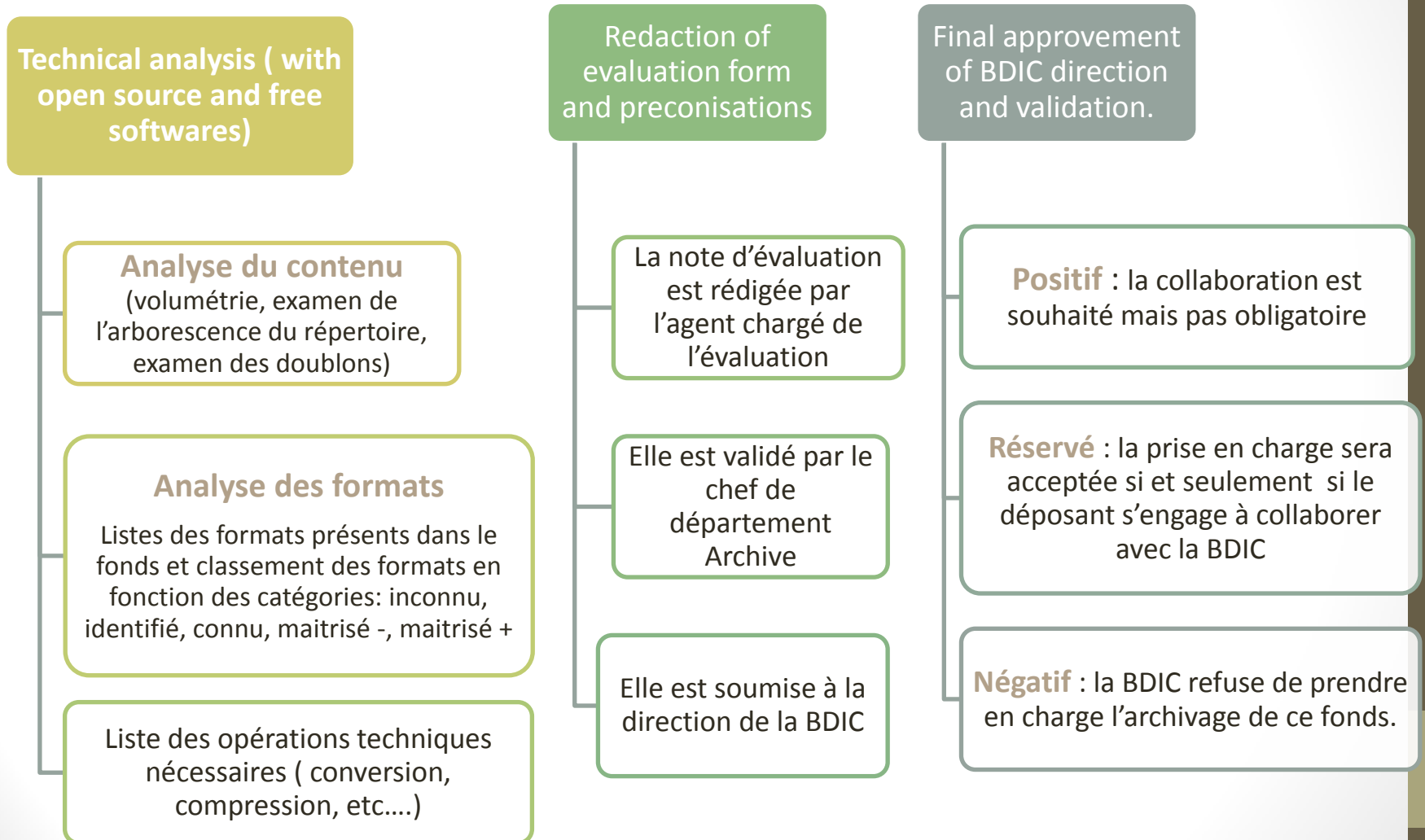
2.2 Documents Transfer

When the archives creator have finished the preparation, we come for manage the transfert of documents.

- BDIC archivist comes with an external drive
- He operates the transfert of the source directory to the external driver in a new directory named with provisionnal access number
- He attaches to this directory, the information request form already completed by archives creator, or completed together the day of transfert.



Phase 3: Analysis and Evaluation

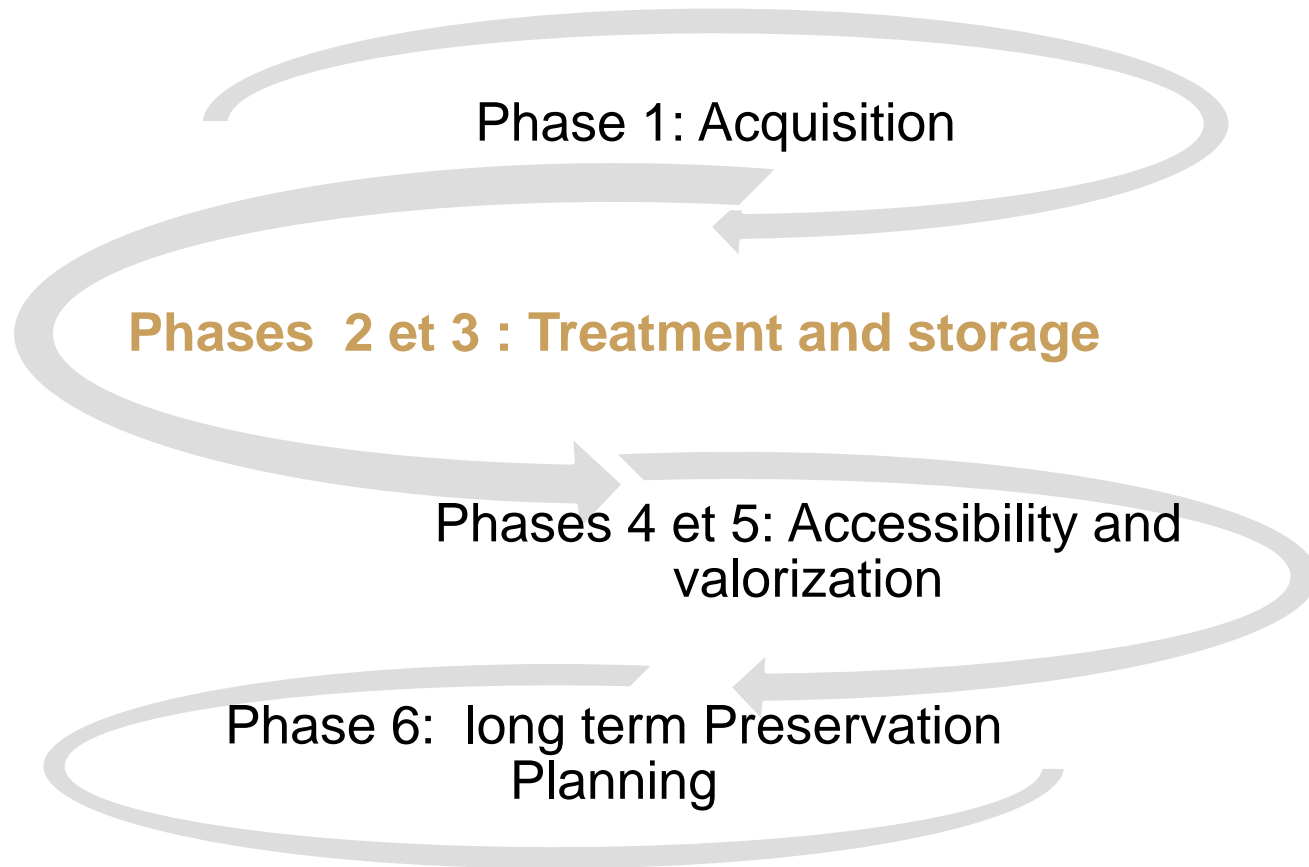


Phase 4 : Preparation of donation agreements and finalisation of acquisition

- The engagements of the 2 parts (the creator and the BDIC) are defined and especially
 - the condition in which the archive will be made accessible to researchers
 - The level of collaboration between creator and BDIC for the treatment.
- If two parts approve the agreements, the donation becomes effective.
- The provisional accession number becomes the definitive
- The content of external harddrive repository is copied to repository of BDIC, in a “buffer zone”.



The technical treatment and documentation treatment could have now begin



This means : Start the phase 2 on our workflow... ;) !

