

# Schemas For Doms

These are the datamodels and metadata formats used in DOMS.

## General Fedora

Object

- OAI Dublin Core ([definition](#)) ([schema](#))
- RDF ([definition](#)) ([schema](#))
- XACML 1.0 ([definition](#)) ([schema](#))

Content Model

- OWL ([definition](#)) ([schema](#))
- DS-COMPOSITE-MODEL ([definition](#)) ([schema](#))

## General DOMS

Item

- PREMIS 1.1 ([definition](#)) ([schema](#))

## Radio and TV

The Radio/TV collection is the most complex in structure of all our collections. This is mainly due to three aspects:

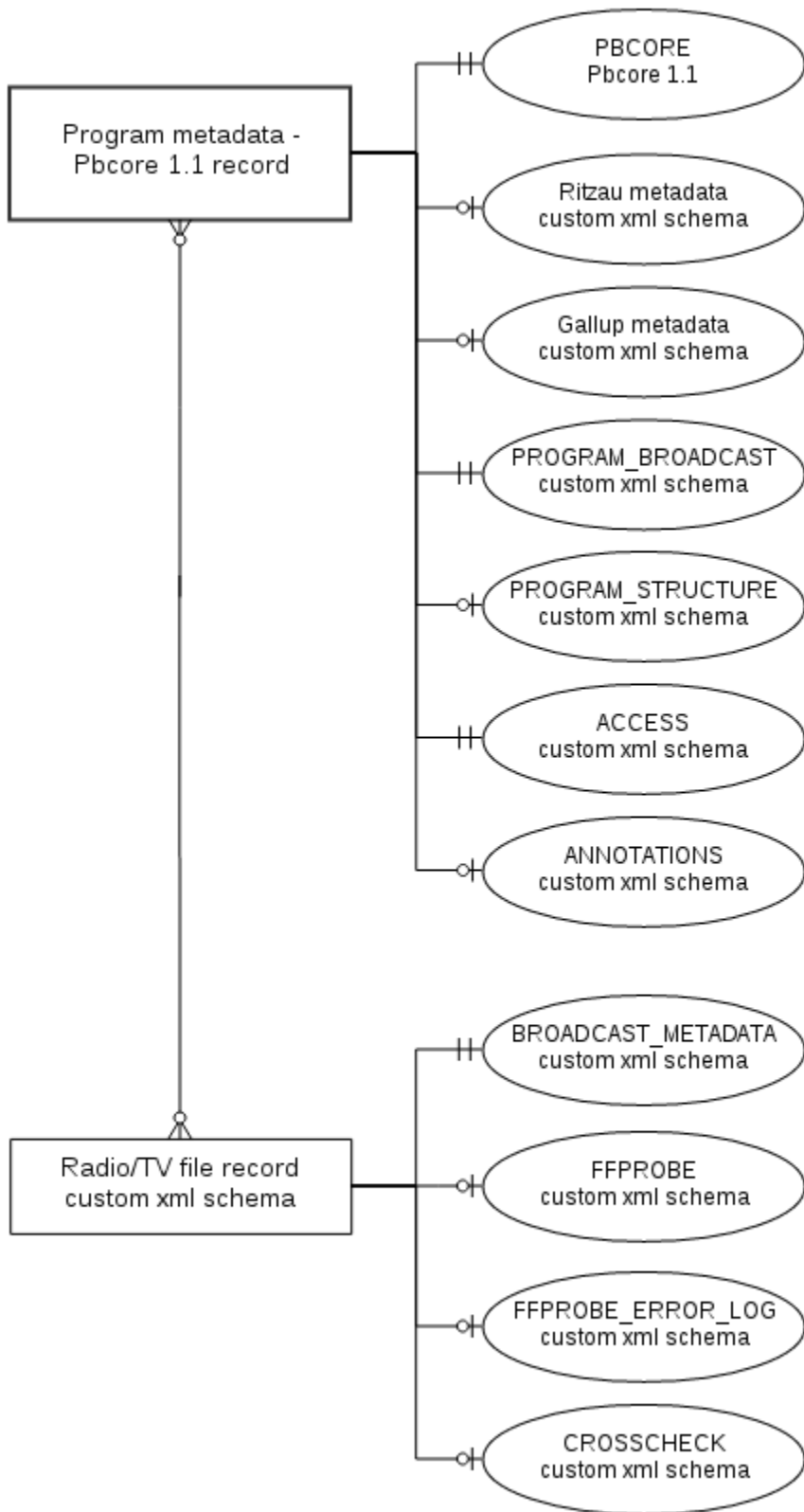
- The description of a single program has multiple sources of metadata.
- We have a significant amount of metadata we wish to preserve and ingest which is not yet linked to any files. This is an essential requirement, as it allows us to acquire new metadata, ingest and then connect it to a file at a later stage e.g. when we have digitised a VHS tape holding the program.
- The archived files contain not programs but chunks of programs, e.g. two channels for four hours in a single file. This means that a program can be contained in one or many files, with an overlap in time between the files. Therefore we need to describe which file(s) contain(s) a given program.

In order to create an accurate description of a single program, we combine metadata records from two external providers of program metadata, Ritzau and TV-Meter, into one PBCore record. This record is the main descriptive record for the program. If errors are detected, the PBCore record can be manually edited .

A number of tools are run on the files, e.g. FFprobe and Crosscheck, and the output from these is also stored in the Existing Metadata Repository and linked to the file.

So for the Radio/TV collection the Existing Metadata Repository holds these types of records:

- PBCore v1.1 record. This is the main descriptive record for a program,
- Ritzau record, in a custom XML schema. This is delivered by an external supplier and ingested on a daily basis with a delay of 14 days,
- TV-Meter record, in a custom XML schema. This is delivered by an external supplier and ingested on a daily basis with a delay of 14 days,
- ACCESS record, in a custom XML schema. This contains details about special access restrictions for a single program,
- PROGRAM\_BROADCAST, in a custom xml schema. This describes the exact date, time and channel a program was broadcasted,
- Program\_structure record, in a custom xml schema. This is the output of an analysis tool and describes the exact details about which files contains the program, defects in those files and if there is an overlap between the files,
- File record, in a custom XML schema. This uniquely identifies a file in the Bit Repository,
- FFprobe record, in a custom XML schema. This contains the output from FFprobe,
- FFprobe error record, in a custom XML schema. If FFprobe reports errors in the file, this record contains the output from the tool,
- Crosscheck record, in a custom XML schema. This contains the output from Crosscheck,
- Broadcast\_metadata, in a custom XML schema. This describes from which channels and time period a file has content.

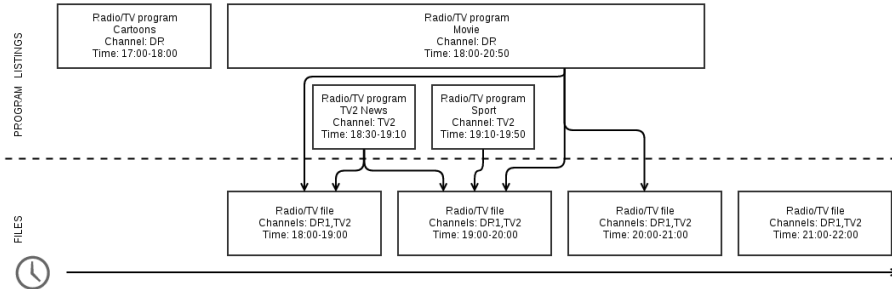


*The Radio/TV data model.*

**Relationship between Metadata and Files in the Radio/TV Collection**

The diagram below (figure 5) is a graphic representation of the relationship between metadata and files in the Radio/TV collection.

- A common situation is that we have a metadata record, describing a program, that as yet has no relations to any files in the Bit Repository.
- A program, described in a metadata record, can consist of parts of many files. Some files are recorded as chunks that each contains an hour of broadcast from several channels.
- A file can be preserved in the Bit Repository and described in the Existing Metadata Repository, even if we have no related program metadata describing its content.



*Program and file relations.*

## Schemas

### Program

- PBCore 1.1 ([definition](#)) ([schema](#))
- Ritzau original custom schema ([schema](#))
- Gallup original custom schema ([schema](#))
- Program structure custom schema ([schema](#))
- Program broadcast custom schema ([schema](#))
- Annotations custom schema ([schema](#))
- Access custom schema ([schema](#))

### Radio/Tv File

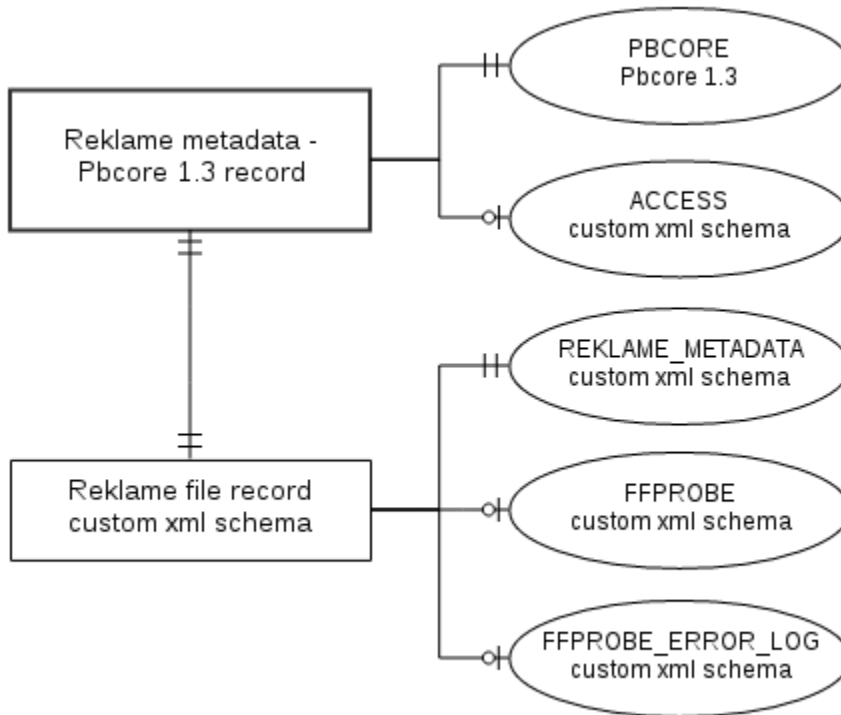
- Broadcast metadata custom schema ([schema](#))
- Ffprobe custom schema ([definition](#)) ([schema](#))
- Ffprobe error log custom schema ([definition](#)) ([schema](#))
- Crosscheck custom schema ([definition](#)) ([schema](#))

### VHS File

- VHS Metadata ([schema](#))
- Ffprobe custom schema ([definition](#)) ([schema](#))
- Ffprobe error log custom schema ([definition](#)) ([schema](#))
- Crosscheck custom schema ([definition](#)) ([schema](#))

## Commercials

The commercials collection contain both television commercials and commercials from movie theaters. The datamodel is fairly simple, with one PBCore metadata record per file.



## Schemas

### Commercial

- PBCore 1.3 ([definition](#)) ([schema](#))
- Access custom metadata ([schema](#))

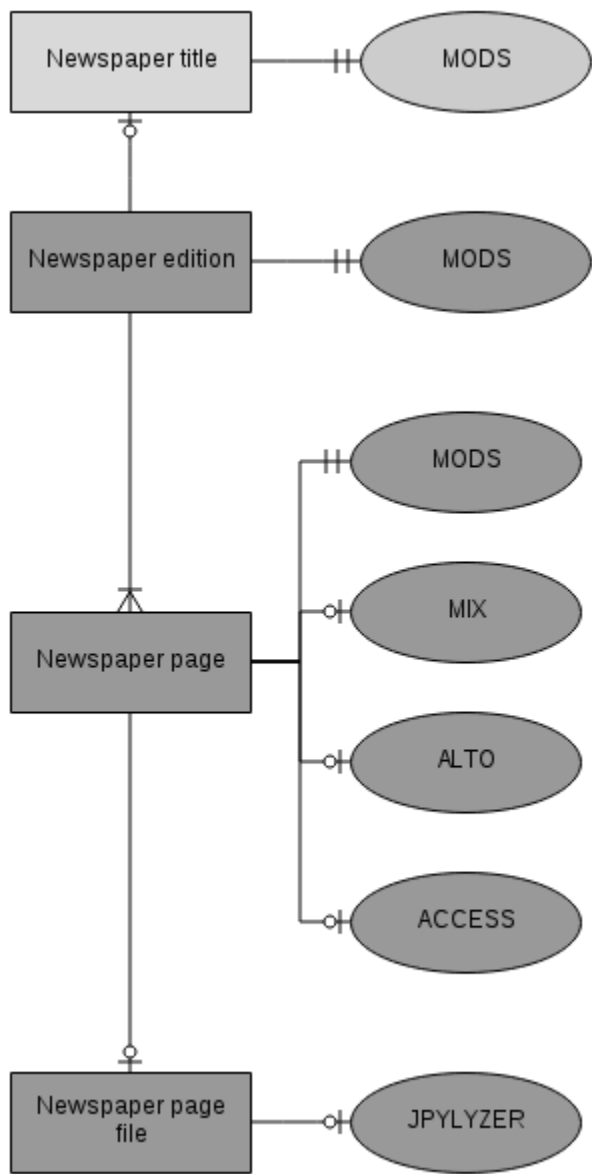
### File

- Reklame metadata custom schema ([schema](#))
- Fprobe custom schema ([definition](#)) ([schema](#))
- Fprobe error log custom schema ([definition](#)) ([schema](#))

## Newspapers

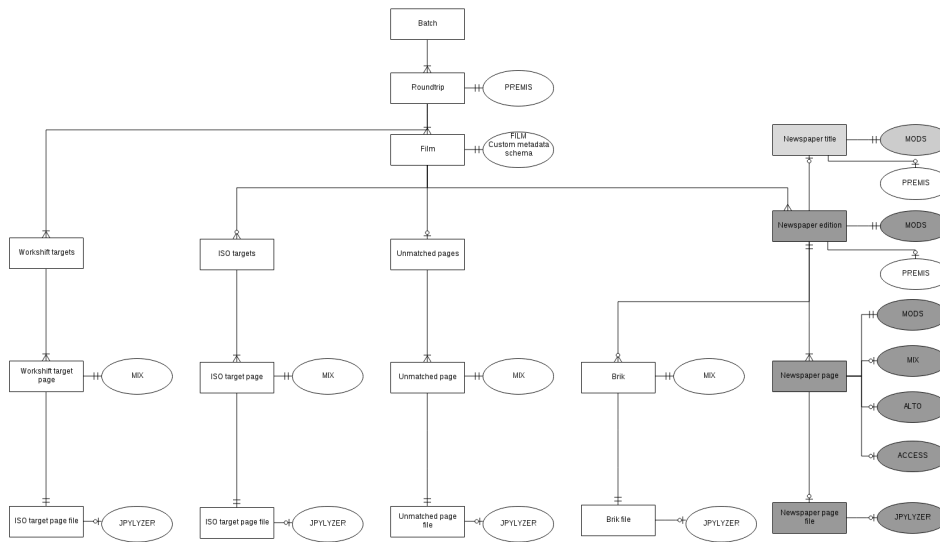
In the newspaper collection (digitised printed newspapers) we use ALTO, MIX, MODS (3 different profiles) and PREMIS. In addition we use custom schemas for characterisation information of image files, administrative data for discovery and for describing newspaper microfilm reels (administrative and technical data).

Our newspaper collection is hierarchically structured on a logical level with data on the newspaper title (MODS), on the issues (MODS) and pages (MODS, MIX, ALTO and ACCESS data) . Furthermore we have characterisation information from Jpylyzer for our JPEG2000 image files.



*The newspaper data model.*

The complete metadata for the newspaper digitisation workflow extends the model and is based on microfilm batches, each consisting of a number of microfilm reels containing issues as described above as well as extra images with test targets and a number of extra pages.



*The newspaper batch data model.*

## Schemas

### Title

- [Mods 3.5 \(definition\) \(schema\)](#)

### Film

- [Film custom schema \(definition\) \(schema\)](#)

### Edition

- [Mods 3.4 \(definition\) \(schema\)](#)
- [Access custom schema \(schema\)](#)

### Page

- [Mods 3.4 \(definition\) \(schema\)](#)
- [Mix 2.0 \(definition\) \(schema\)](#)
- [Alto 2.0 \(definition\) \(schema\)](#)
- [Access custom schema \(schema\)](#)

### Jpeg2000 File

- [Jpylyzer custom schema \(definition\) \(schema\)](#)

### Brik

- [Mix 2.0 \(definition\) \(schema\)](#)