Redactive Intermediate Format (RIF)

The Redactive Intermediate Format "RIF" is an XML formatted file that consists of an informational header followed by blocks of data that is a concatenation of converted records from various newsroom systems.

The first tag <REDACTIVE> is the parent tag for all stories within the document. There are a couple of informational attributes you may want to use:

SITEID   This will contain the business unit or call letters of the copyright owner of the archive material

ORIGINALPATH   This is useful only to Redactive and is the path from which the raw data was converted.

CONVERSIONDATE   This is the date the conversion was started.

Example:
```
<REDACTIVE SITEID="site id information" ORIGINALPATH="d:\!ToDo\WWMT\Delta"
CONVERSIONDATE="1/10/2015" >
```

> *Note: Although this is an XML file it is not recommended you use a standard XML parser until you first extract individual records.   There can be over a million individual records*

Each Record is enclosed in a parent tag <NRCS_RECORD> with informational attributes:

RECORDNUMBER   A simple number (beginning with zero) assigned to each converted record.

DATE   The date extracted from the original record. Typically this will be the date the record was created in the original NRCS.

ORIGINALFILE   If known, this will be the original file from the original NRCS from which the record was extracted and converted.   Often multiple files are involved in a single record – this will represent the master record.

The actual record for each story will be enclosed in a CDATA section.   Each story will be well formed XML however.

Each story record will use the following XML template

```
<inception version="1">
<story>
        <ordinal><?_ORDINAL></ordinal>
        <title><?_SLUG></title>
        <runningorder><?_SHOW></runningorder>
        <page><?_PAGE></page>
        <format><?_FORMAT></format>
        <writer><?_AUTHOR></writer>
        <videoid><?_TAPE></videoid>
        <timecode><?_TIMECODE></timecode>
        <tapelenseconds><?_TAPELENSECONDS></tapelenseconds>
        <airdate><?_SERIALAIRDATE></airdate>
        <createdate><?_SERIALDATE></createdate>
        <modifydate><?_SERIALMODDATE></modifydate>
        <modifyby><?_MODIFYBY></modifyby>
        <outcue><?_OUTCUE></outcue>
        <reporter><?_REPORTER></reporter>
        <photographer><?_PHOTOGRAPHER></photographer>
        <comments><?_COMMENTS></comments>
</story>
<rolodex>
        <contactprefix><?_CONTACTPREFIX></contactprefix>
        <contactfullname><?_CONTACTFULLNAME></contactname>
        <contactfirstname><?_CONTACTFIRSTNAME></contactfirstname>
        <contactlastname><?_CONTACTLASTNAME></contactlastname>
        <contacttitle><?_CONTACTTITLE></contacttitle>
```

```
            <contactphone><?_CONTACTPHONE></contactphone>
            <contactmobile><?_CONTACTMOBILE></contactmobile>
            <contactfax><?_CONTACTFAX></contactfax>
            <contactemail><?_CONTACTEMAIL></contactemail>
            <contactcompanyname><?_CONTACTCOMPANYNAME></contactcompanyname>
            <contactfulladdress><?_CONTACTFULLADDRESS></contactfulladdress>
            <contactaddress1><?_CONTACTADDRESS1></contactaddress1>
            <contactaddress2><?_CONTACTADDRESS2></contactaddress2>
            <contactcity><?_CONTACTCITY></contactcity>
            <contactstate><?_CONTACTSTATE></contactstate>
            <contactcountry><?_CONTACTCOUNTRY></contactcountry>
            <contactzip><?_CONTACTZIP><contactzip>
</rolodex>
<mosexport itemcount="0">
        <?_MOSOBJECTS>
</mosexport>
<body>
<?_BODY>
</body>
</inception>
```

Note: "inception" in the above example may be replaced by other target system names.

In the above template there are three parent nodes: "<story>", "<rolodex>" and "<mosexport>". <story> is used for all records that fit the category of a news story while <rolodex> is used for contacts information. <mosexport> will contain all MOS objects in their original form from the original stories.

Within the above template you see several replacement variables. This is used only during the conversion process to identify locations for standardized field information from various newsroom systems.   As a recipient of a RIF file you will not see these replacement variables.   On the following pages is the explanation for each so you can understand the nature of the data they will contain.

Tags for child nodes of <story>

<?_ORDINAL>	If available this represents the original item number for the story in a running order
<?_SLUG>	Contains the title of the story. This is typically cleaned up to remove illegal filename characters
<?_SHOW>	Contains the name of the running order from which the script came. This is not always known.
<?_PAGE>	The original page number assigned to this story in the running order.
<?_FORMAT>	Contains story format information such as VO, VOSOT, PKG, READER, etc.
<?_AUTHOR>	The person who wrote the story
<?_TAPE>	A tape ID – hopefully an immutable one, but often not. This is copied directly from the original record.
<?_TIMECODE>	The timecode position on the tape if provided
<?_TAPELENSECONDS>	The length of the tape portion of a story in seconds
<?_SERIALAIRDATE>	The date the story aired if known. Given as a Unix Epoch in GMT
<?_SERIALDATE>	The create date of the story. Given as a Unix Epoch in GMT
<?_SERIALMODDATE>	The date the story was last modified. Given as a Unix Epoch in GMT
<?_OUTCUE>	If the original NRCS provided an outcue it will be placed here
<?_REPORTER>	The reporter assigned to the story
<?_PHOTOGRAPHER>	The photographer assigned to the story
<?_COMMENTS>	A commend field from the original NRCS if provided
<?_BODY>	Contains the entirety of the script that would have appeared on both the left and right hand side of the story. This includes in-line cues and MOS IDs (MOS objects are reduced to MOS ID's only). Cues will appear inside of square brackets.

Tags for child nodes of <rolodex>

<?_CONTACTPREFIX>	Mr. Mrs. Ms. Etc.
<?_CONTACTFULLNAME>	A users concatenated first and last names. Since this may be the only field available from some newsroom systems it may be the only field with data. Also – it cannot be reliably determined if this will be firstname – lastname or lastname – firstname, or a mix.
<?_CONTACTFIRSTNAME>	If it can be reliably determined this will contain the first name
<?_CONTACTLASTNAME>	If it can be reliably determined this will contain the last name
<?_CONTACTTITLE>	Contacts proper title
<?_CONTACTPHONE>	Contacts phone number
<?_CONTACTMOBILE>	Contacts mobile number
<?_CONTACTFAX>	Contacts fax number
<?_CONTACTEMAIL>	contacts email
<?_CONTACTCOMPANYNAME>	Contacts company name
<?_CONTACTFULLADDRESS>	Contacts concatenated full address. Sometimes only concatenated full addresses are available and they cannot be reliably parsed to the following individual fields.
<?_CONTACTADDRESS1>	Contacts address1
<?_CONTACTADDRESS2>	Contacts address2
<?_CONTACTCITY>	Contacts city
<?_CONTACTSTATE>	Contacts state or province
<?_CONTACTCOUNTRY>	Contacts country
<?_CONTACTZIP>	Contacts Zip or Post code

Not all fields will have data. If a field is blank the tags will still appear, but without data.

MOS Export Node
The tag <mosexport> identifies the section that will contain all of the originally formatted MOS objects from the story. The tag has one attribute that shows the count of MOS objects within the node.

Example: <mosexport itemcount="4"> *there will be 4 <mos>…</mos> objects in this node* </mosexport>