Workflow Automation in the Graphic Arts Industry

CIP4
Definitions
Metadata
Process-Resource Model
(X)JDF/(X)JMF, PrintTalk

Thomas Hoffmann-Walbeck
International Cooperation for the Integration of Processes in Prepress, Press and Postpress

» Not-for-profit standards specification association

» Metadata/job tickets: They specify products and production details for automatic production and business workflows

» JDF/JMF, XJDF/XJMF, PrintTalk and PPF

» Specifications, programming libraries, tools, case studies and marketing material can be used by everybody without paying royalties

» Over 1,600 individuals from approximately 300 members from 31 countries

» Voluntary work

http://www.cip4.org/

https://confluence.cip4.org
Example for Device/Machine Automation
Automation in Graphic Arts Industry

This lecture covers only:

**Workflow Automation**

**Device Automation**

Workflow and device automation depend on each other. Often device automation needs (meta) data from outside; workflow messages provide this data.
Print Production Workflow Automation

- Which information are in the packages?
  - Green: Product intents, information for processes
  - Blue: Time & material consumed, quality report, status

Just a schematic sketch!
In reality, the communication is much more dynamic.
Metadata, Devices and Controller

- A **device** is either a software application or some machine. It initiates the execution of one or several processes. Sometimes, it refers only to the device driver for a machine.

- **Metadata** = information/data about content data (text, images, graphics, pages) or about print production.

- A **controller** routes metadata and messages in a workflow system.
  - E.g., MIS/ERP, Workflow (Management) System

- Devices and controller communicate with each other about product and production details.
Remarks to the last animated chart

- Only communication between MIS/controller and production devices are shown. Workflow, however, should cover other areas as well!
  » Artwork creation, procurement, logistics, business partners...

- The chart suggests 3 views of workflow automation:
  I. The network (arrows)
  II. The information container (green and blue container)
  III. The information inside (things inside the container)
  » Our main focus in (III). However, in order to read the information, one also needs to know about the structure of the container - not on an IP or HTTP level, but, for example, on the XML level if the information is coded in XML
Concept of Workflow Automation

- A **workflow** denotes the sequence of production and business processes, in whole or part, during which documents, information and task descriptions are passed from one participant to another for action according to a set of procedural rules.

- A **process** is an activity with a single specific objective that can be planned and executed independently.

- **Workflow Automation** means automatically supplying processes (devices) with data.

- A process can start and execute automatically, if...
  - ...the data in the “green containers” is available and relevant.
  - ...the device understands the data and can execute a process automatically using the data (device automation).
What's in the Containers?

- Green (from controller to device):
  - Job administration
    - Customer details, Job ID, production time window
  - Description of the requested product (in it's finished stage)
    - e.g., printing substrate, page formats, binding intent, type of inks
  - Production issues for pre-setting devices and apps
    - e.g., actual printing substrate, screening type, folding positions
  - Queries and commands
    - e.g. What is the status? Update job, cancel job

- Blue (from device to controller):
  - Time & material consumed (currently/in total), job tracking, ...
Data Exchange between PP, PB and 3rd Party

Print Provider

Management Information System (MIS)
*Business Workflow*

Production Controller
*Technical Workflow*

Prepress | Press | Postpress | Delivery

Print Buyer
- Purchase
- Tracking
- Approval

Services
- Supplier
- Logistic
- Other

**Data Flows:**
- Product description
- Production detail
- Operational data
- Business data
What is Metadata in the Printing Industry?

- Structured **data about data** (e.g., about images, PDF files) or product intent description, print production issues, operational data

- Metadata can be embedded in content data or defined externally to content data (i.e. stored in a file or database)

- Examples of metadata formats:
  - Embedded: Exif in images; XMP is in images, PDF, ID, AI, ...
  - Not embedded: **JDF/JMF, XJDF/XJMF, PrintTalk**, PJTF, PPF (CIP3)

- Most metadata is coded in XML or in recent times in JSON
  - Exceptions: Exif, PJTF, PPF
Metadata – Content

- Metadata often references content (print) data

```
<RunList Run="1" Status="Draft">
  <LayoutElement>
    <FileSpec MimeType="application/pdf"
      URL="file://Jobs/2010-12-13/System/uebung_01.pdf"/>
  </LayoutElement>
</RunList>
```

- Metadata in PDF: XMP, Output Intent, die cut data...
  - Receiver of single pages in PDF/VT (or PDF 2.0)
  - Product intend: ISO/CD 21812-1:2019
Metadata and Their Ecosystems

- Often several metadata formats used in a production
  
  » Format conversions during production
e.g., Exif to XMP, XMP to JDF, XJDF to JDF, CSV to JDF,...

PPF (Print Production Format), CIP3
JDF (Job Definition Format)
XJDF (Exchange JDF)
XMP (eXtensible Metadata Platform)
EXIF (EXchangeable Image File)
PJTF (Portable Job Ticket Format)
PrintTalk
Production Models

- if you want to specify a workflow - especially an automated one - you have to define the production processes and how they are connected
  - A universal print production “model” is needed!

- Different models are available
  - Process list (or activity list or to-do-list): List of processes w/o execution order
  - Process chart: defines the order but does not define the interfaces between the processes
  - Flow chart
  - **Process-Resource-Model**
Flow Charts might Become Complicated
Flow Charts might Become Scary
Process-Resource Models

- Processes are described by verbs, resources are physical things, files or parameter sets
  - Processes normally have several input resources, i.e. prerequisites that are needed for running a process
  - Processes have one or more output resource(s), i.e. the result(s) that the process produces
  - A output resource of one process might be the input resource of the next process (a “transitional” resource)
Process-Resource Model (PRM)

Let's make pancakes!

Sorry, let's make a PRM recipe
Process-Resource-Model for Pancake Production
Processes and Resources Must be Specified

- If you want to replace the “chef” by a “mixing robot”, a “frying robot”, a “separating robot” and a “beating robot” a precise specification for all processes and resources is needed, which...
  - ...must be clear, unambiguous, universal (not for pancakes only!) and represents a complete “language”
  - ...should be an open standard – otherwise the pancake manufacturers cannot buy robots from different vendors

- The specification is the tricky part - and, of course, robots must understand the language and can execute the processes accordingly

- Please Note: The PRM is just a model, just a description what to do (“recipe”). It’s not the production!
Example of PRM in Print

- Production is described by a network of processes and resources
  - Each process is executable when all mandatory input resources are available and the production time window is reached

![Diagram showing the process flow from Bitmaps to Printed Sheets through Image Setting, Imaged Plates, Ink, Paper, and Conventional Printing]
Model a print production workflow

https://www.torontomu.ca/~wdp/workflow-puzzle/print/print.html
JDF/JMF, XJDF/XJDF and PrintTalk

- **Job Definition Format (JDF), Job Messaging Format (JMF)**
  - About 100 processes and 170 resources defined in JDF
  - JDF normally contains the workflow logic
  - JMF is the “SMS” for the printing industry. It allows dynamic interaction between the team players

- **Baby JDF: XJDF and XJMF**
  - X stands for “Exchange”; XJDF simplifies JDF
  - Just a protocol between controller & devices; no workflow logic

- **PrintTalk**
  - Defines business objects like RFQ, quote, confirmation, … , invoice
  - PrintTalk can embed (X)JDF
Metadata Communication Hub

PP’s Partners

- Customer
- MIS
- Job Scheduling
- Reporting

PP Management

- PP’s Partners
- PP Management

Industrial Digitalization ↔ Workflow Automation

- JDF
- XMP
- CFF2
- JMF
- DB
- private XML
- Exif
- PrintTalk
- CSV
- XJDF
- XJMF
- PPF
- PJTF
- XJDF
- PP Processes

- Paper Supplier
- Plate Supplier
- Logistic Services
- Packing & Shipping
- Die Cutting
- Cutting
- Stitching
- Collecting
- Preflight
- Plate Making
- Printing
- Folding
Information Exchange with JDF/XJDF

- **MIS**
  - Page Size
  - Sheet Size
  - Colors
  - Printing substrate
  - Run length
  - Order information (Imposition Scheme)

- **PrePress**
  - Ink zones presetting
  - Previews
  - Cutting positions
  - Folding positions
  - Operational data

- **PostPress**

**Note:**
The arrows symbolize logical addresses. However, most systems route the message through a controller, e.g. MIS.
Planning Board – The Integration Champion
Inside of JDF: Hierarchy of JDF Nodes

- Typically, MIS initializes a JDF for each print job

- A JDF-Job is defined by:
  - Product nodes
  - Process group nodes
  - Process nodes

- JDF nodes are described by XML-elements and XML-attributes.
  - Encoding: `<JDF Type="…">`
  - “Arrows” between those elements represent JDF sub-nodes
Location of Resources

- Nodes can also contain resources
- Since some resources are used by several processes the resources are not necessarily in the JDF node that refers to them as input or output
  - Copies of resources should be avoided
- Other means of defining which is an resource of a node: ResourceLink

Thus:

- The Resource can be (almost) anywhere in any JDF node
- The ResourceLinks in a JDF node define the reference (via ID) to the resources for the node.
Sequential Architecture

- A single JDF file per job passes from one JDF module to the next e.g., via hot folder
  - No database, original concept, somehow outdated
  - JDF “gains weight” during production
- In reality, processes are not sequential
  - overlaps (pipes), parallelism (data consistency)
- Quite complex if level of JDF-integration is high
  - Difficult to maintain de-central data storage
Some software is in charge for w/r of JDF, e.g., MIS

Processes may not use/update JDF data at the same time

» *Spawning* & *merging*: A JDF part rw-spawned only once at a time

» Nevertheless: Each module must be prepared to extract JDF nodes it needs from entire job ticket
(X)JDF is just a communication protocol

(X)JDF, (X)JMF
Private DB access
XJDF – The simplified JDF

- JDF is designed to contain information for the entire production
  - JDF node tree, resources, ResourceLinks, ...
  - This leads to quite complex data structures
  - Moreover, such description is often stored additionally in private data bases of MIS or production controller

- XJDF does not contain the entire workflow logic
  - It is merely an interface protocol between a production controller and specific devices
  - Easier to implement, less costly – especially for the devices
  - Workflow logic is stored in private data bases of controller
Comparison JDF - XJDF

Main technical advantages of XJDF to JDF

» Devices need not to extract relevant data from the job ticket

» Data bases are handling data update – no need for extra mechanism (*spawn* & *merge*) as with JDF

» One XJDF node only with XJDF, no node tree, no *ResourceLinks*
PrintTalk

- PrintTalk® is a CIP4 protocol that handles business aspects in the graphic arts industry
- Main application area is the interface PB - PP
  » PrintTalk is used in the background for example by
  » Web-to-Print systems
  » Customer portals
- Moreover, PrintTalk is deployed by MIS – MIS and by MIS – WMS communication
  » E.g., between subsidiaries of a print shop
- PrintTalk is XML based
  » The root element of a PrintTalk document is <PrintTalk ...>
Lab Work: EasyXJDF Hands on I

- Download EasyXJDF from CIP4.org
  - See Tools & Libraries
  - For Mac take dmg file
- Open the dmg file on Macintosh
  - Control & Open
- Fill entries in GUI with whatever you like
  - Choose PDF as content
- “Save as …”
- Unzip and open the ptk file with XML editor or any dumb text editor and…
Lab Work: EasyXJDF Hands on II

... answer the following questions:

» What is the type of PrintTalk Business Object?
» What is the DescriptiveName of the XJDF job?
» How many Products are in the ProductList?
» How many Intent elements do you find?
» What are the names of the Intent elements?
» How many ResourceSets do you find in the file?
» What are the names of the ResourceSets?
Outcome of EasyXJDF - 1

```
<ptk:PrintTalk version="2.0" payloadID="..." Timestamp="20190327T16:44:40Z">
  <ptk:Request>
    <ptk:PurchaseOrder BusinessID="4711">
      <xjdf:XJDF Category="Web2Print" DescriptiveName="Industry 4.0 and JDF"
                 JobID="4711" Version="2.0">
        <xjdf:GeneralID IDValue="42" IDUsage="CatalogID"/>
        <xjdf:ProductList>
          <xjdf:Product Amount="2222">
            <xjdf:Intent Name="MediaIntent">
              <xjdf:MediaIntent MediaQuality="Lumisilk 120"/>
            </xjdf:Intent>
            <xjdf:Intent Name="ColorIntent">
              <xjdf:ColorIntent>
                <xjdf:SurfaceColor NumColors="4" Surface="Front"/>
                <xjdf:SurfaceColor NumColors="0" Surface="Back"/>
              </xjdf:ColorIntent>
            </xjdf:Intent>
            <xjdf:Intent Name="LayoutIntent">
              <xjdf:LayoutIntent FinishedDimensions="566... 651... 0.0"
                               PrintedPages="1" Sides="OneSided"/>
            </xjdf:Intent>
          </xjdf:Product>
        </xjdf:ProductList>
      </xjdf:XJDF>
    </ptk:PurchaseOrder>
  </ptk:Request>
</ptk:PrintTalk>
```
Outcome of EasyXJDF - 2

```xml
<ptk:PrintTalk>
  <ptk:Request>
    <ptk:PurchaseOrder>
      <xjdf:XJDF>
        <xjdf:ResourceSet Name="CustomerInfo">
          <xjdf:Resource ID="CustomerInfo_c799d2eb-1377-45da-8679 ...">
            <xjdf:CustomerInfo CustomerID="MickeyMouse"/>
          </xjdf:Resource>
        </xjdf:ResourceSet>

        <xjdf:ResourceSet Name="RunList">
          <xjdf:Resource ID="RunList_0ba58ec1-3093-4c0e-ae44-e38 ...">
            <xjdf:RunList>
              <xjdf:FileSpec URL="asset/JDF Specification 1.7.pdf"/>
            </xjdf:RunList>
          </xjdf:Resource>
        </xjdf:ResourceSet>
      </xjdf:XJDF>
    </ptk:PurchaseOrder>
  </ptk:Request>
</ptk:PrintTalk>
```
(X)JDF most important metadata in print production

- They are based on the Process Resource Model
- They define the product intent and the production processes via their input and output resources
- They contain information only, not the specifics how the information is used and distributed. They fuel the automation, but they don’t execute the automation.

XJDF is a simplified version of JDF. It does not contain the entire workflow logic

- Both formats are used in parallel

PrintTalk defines the business workflow of a print provider and print buyer
Thank you for your attention!

hoffmann@hdm-stuttgart.de