

docWORKS/METAe

**Automated Conversion
of Printed Material
into METS/ALTO Objects**

What is docWORKS/METAe?

- Production tool for conversion of printed documents into fully tagged digital objects
- The METAe edition of docWORKS is the result of the EU-funded project METAe , September 2000 to August 2003
- Clients: Harvard, LC, GDZ, Royal Library Denmark

The project group

1. Leopold-Franzens-Universität Innsbruck (Co-ordinator), Austria
2. Universität Linz, Institut für Angewandte Informatik, University of Linz, Austria
3. Mitcom Neue Medien GmbH (ABBYY Europe), Germany
4. CCS Compact Computer Systeme, Germany
5. Universidad de Alicante, Spain
6. Friedrich-Ebert-Stiftung, Germany
7. Cornell University Library. Department of Preservation and Conservation, USA
8. BNF, Bibliothèque Nationale de France
9. The National Library of Norway, Rana division, Norway
10. Biblioteca Statale A. Baldini, Italy
11. Dipartimento di Sistemi e Informatica, University of Florence, Italy
12. Karl-Franzens-Universität Graz, Universitätsbibliothek, Austria
13. Scuola Normale Superiore, Centro di Ricerche Informatiche per i Beni Culturali, Italy
14. Higher Education Digitisation Service HEDS, UK

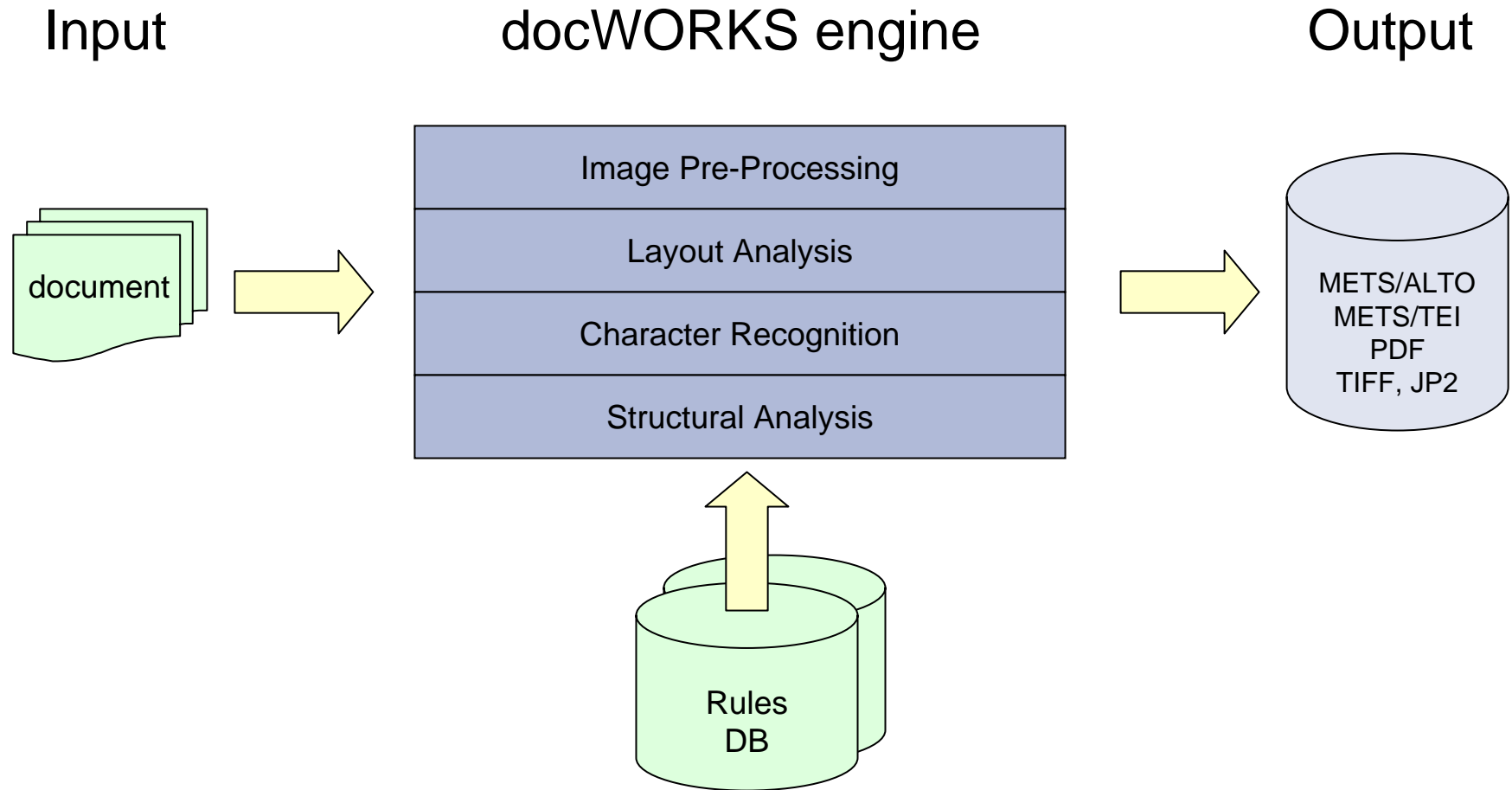
Challenges

- ☞ Digitization and retro-conversion of printed or textual material is getting more and more important:
 - Keep knowledge and cultural heritage alive
 - Preserve the origin
 - Enable quick and enhanced access by high structured documents
 - Open up new dimensions of research
 - Provide standardized output formats

Goals

- Automate the conversion process
- Make digitization more effective and safer
- Go for mass digitization
- Increase the added value of digitized collections
- Provide a standardized output format

System Overview



Matching of Image Files and Page Numbers

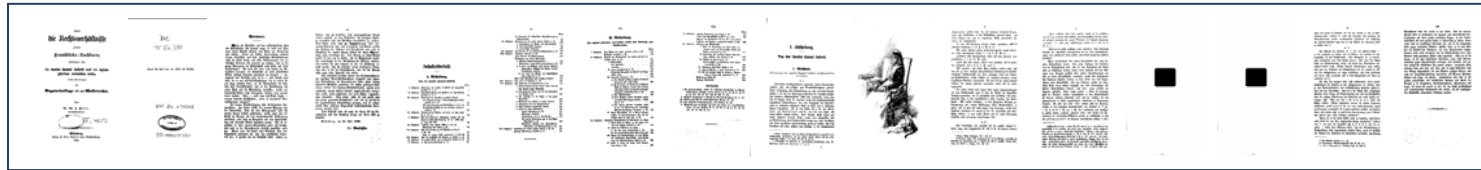
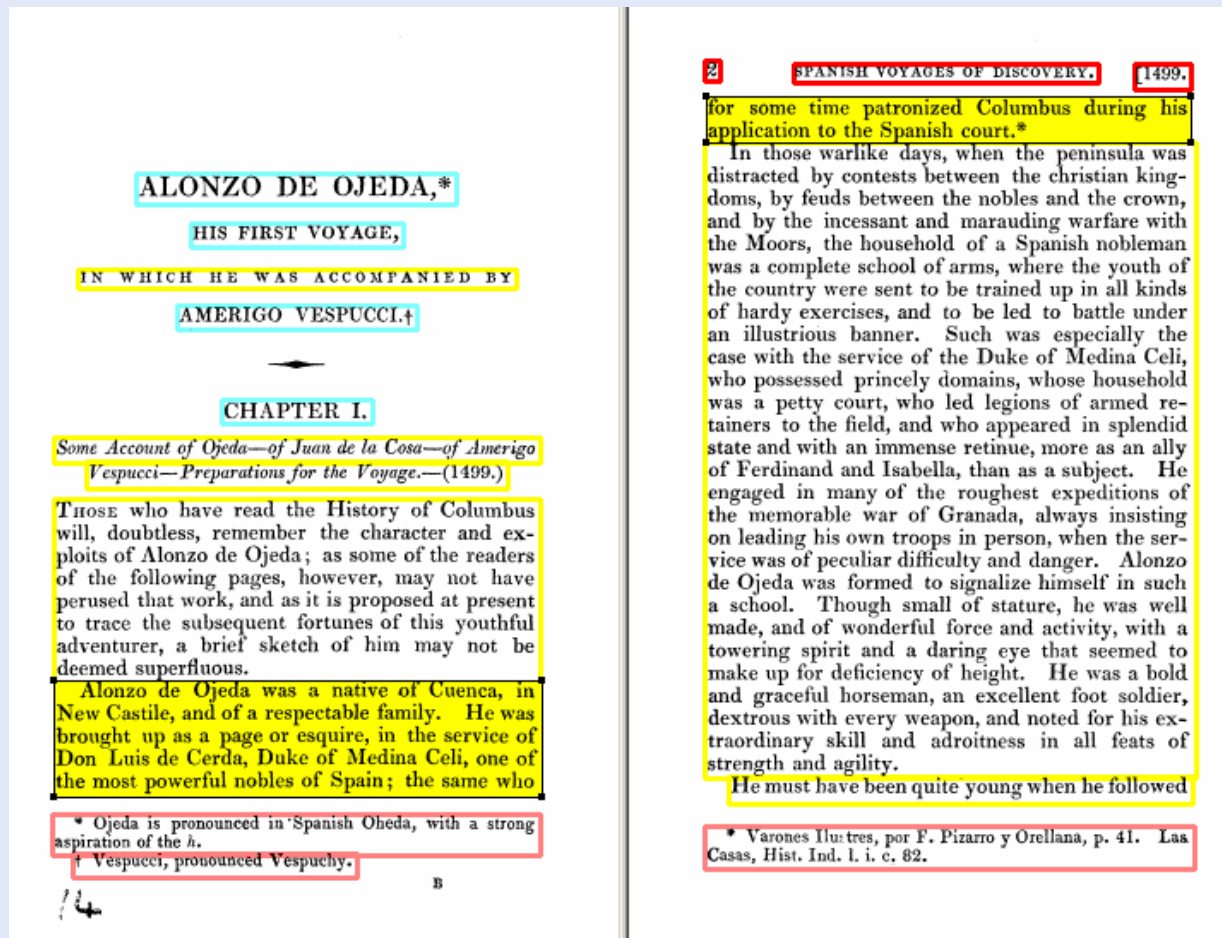


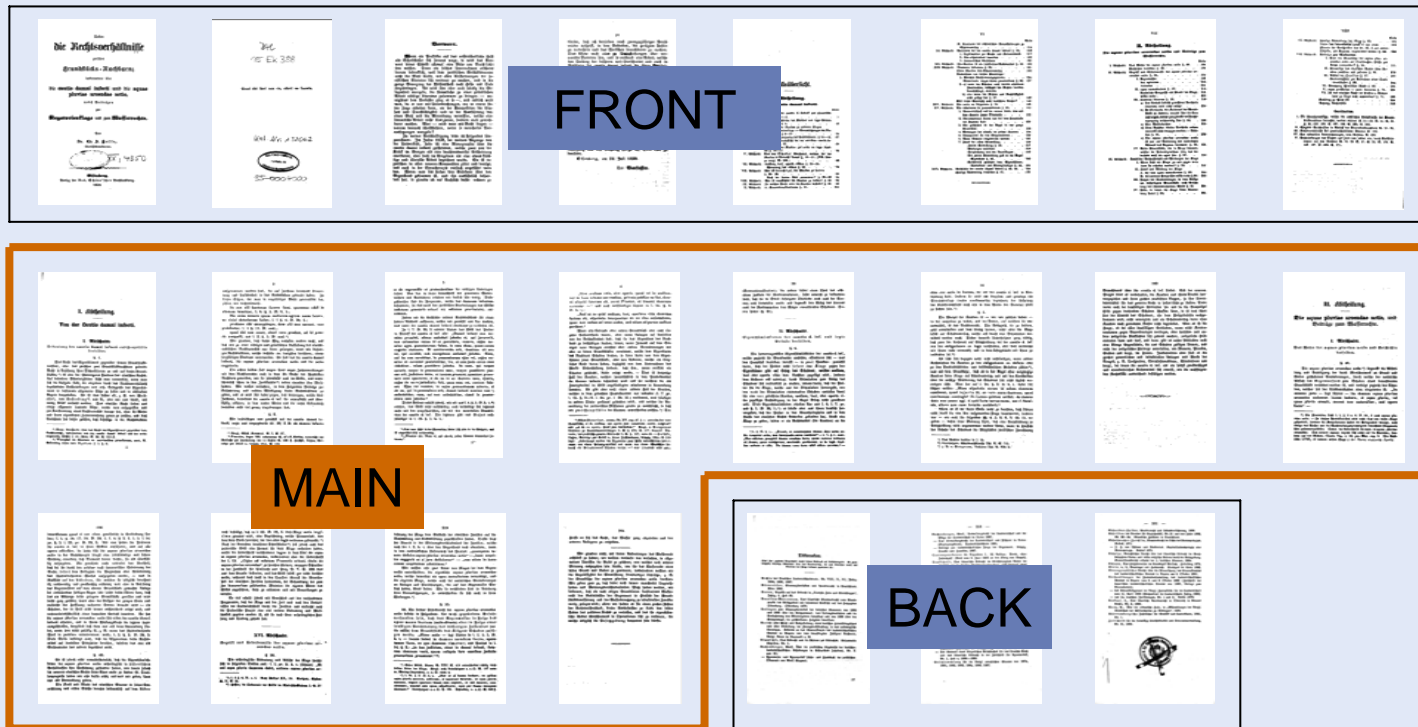
Image-file	Pagination	Page-Number
000001.tif	Not counted	Np
000002.tif	Not counted	Np
000003.tif	Counted	I
000004.tif	Counted	II
000005.tif	Counted	III
000006.tif	Counted	IV
000007.tif	Counted	V

000008.tif	Counted	VI
000009.tif	Counted	1
000010.tif	Counted, not paginated	(2)
000011.tif	Counted	3
000012.tif	Counted	4
placeholder	Missing page	5
placeholder	Missing page	6
000013.tif	Counted	7
000014.tif	Counted	8

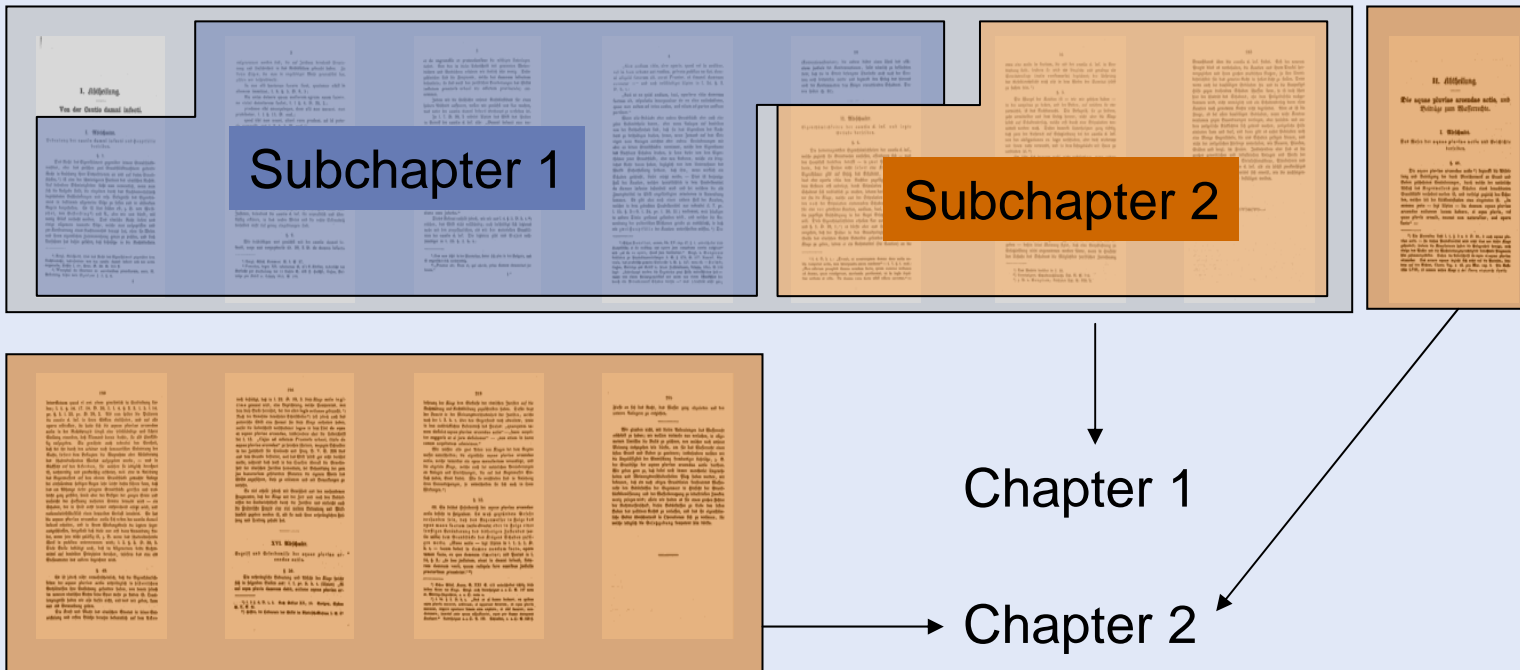
Physical and Logical Structure



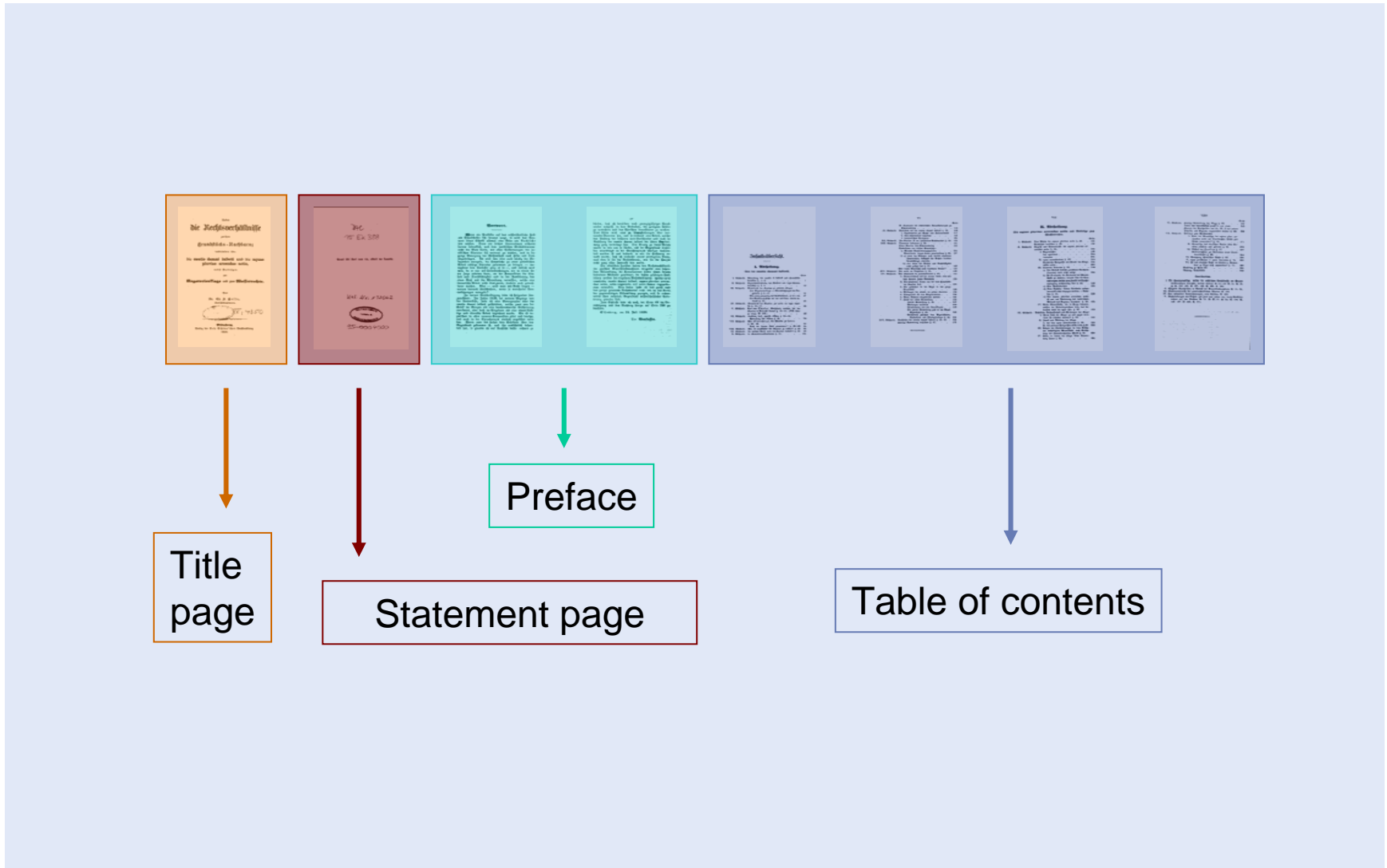
Structural Analysis



Structural Analysis



Structural Analysis



Digitization of Books and Journals (METAe)

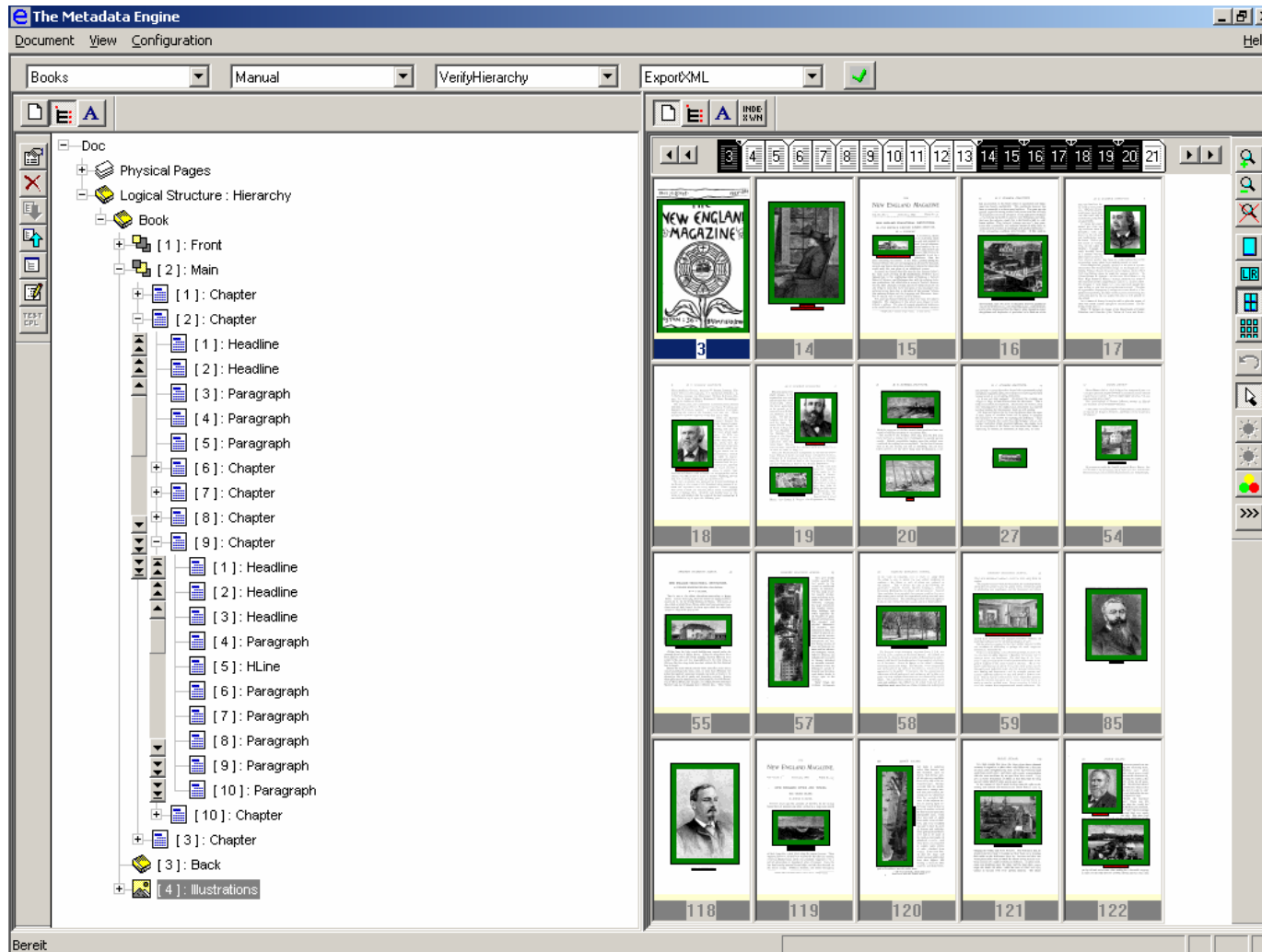
The screenshot displays the 'The Metadata Engine' software interface. The window title is 'The Metadata Engine' and it includes a menu bar with 'Document', 'View', 'Configuration', and 'Help'. Below the menu bar are several dropdown menus: 'Books', 'Manual', 'VerifyHierarchy', and 'ExportXML'. The interface is divided into three main sections:

- Left Panel (Doc):** A tree view showing the document's logical structure. It includes 'Physical Pages' (pages 150-159) and a 'Logical Structure: Hierarchy' section. The hierarchy starts with a 'Book' containing a 'Front' section, followed by a 'Main' section with multiple 'Chapter' entries (1-10), a 'Back' section, and 'Illustrations'.
- Table:** A table mapping physical page numbers to their corresponding image file names.

[162] : imagePage	150
[163] : imagePage	151
[164] : imagePage	152
[165] : imagePage	153
[166] : imagePage	154
[167] : imagePage	155
[168] : imagePage	156
[169] : imagePage	157
[170] : imagePage	158
[171] : imagePage	159
- Right Panel:** A grid of six digitized pages, labeled 162 through 167. Each page shows the original scanned content with yellow highlighting indicating detected text blocks. Page 162 features an illustration of a building. Pages 163 and 164 contain portraits of men. The text on all pages is highlighted in yellow.

The status bar at the bottom left indicates 'Ready'.

Digitization of Books and Journals (METAe)



Digitization of Scientific Documents

The Metadata Engine

Document View Configuration Help

Books Manual VerifyHierarchy ExportXML

Physical Pages

[14] : imagePage	10
[15] : imagePage	11
[16] : imagePage	12
[17] : imagePage	13
[18] : imagePage	14
[19] : imagePage	15
[20] : imagePage	16
[21] : imagePage	17
[22] : imagePage	18
[23] : imagePage	19

Logical Structure : Hierarchy

- Book
 - [1] : Front
 - [2] : Main
 - [1] : Chapter
 - [2] : Chapter
 - [1] : Headline
 - [2] : Headline
 - [3] : Paragraph
 - [4] : Chapter
 - [5] : Chapter
 - [6] : Chapter
 - [7] : Chapter
 - [3] : Chapter
 - [4] : Chapter
 - [5] : Chapter
 - [6] : Chapter
 - [7] : Chapter
 - [8] : Chapter
 - [9] : Chapter
 - [3] : Back

Chapter 2
Overview

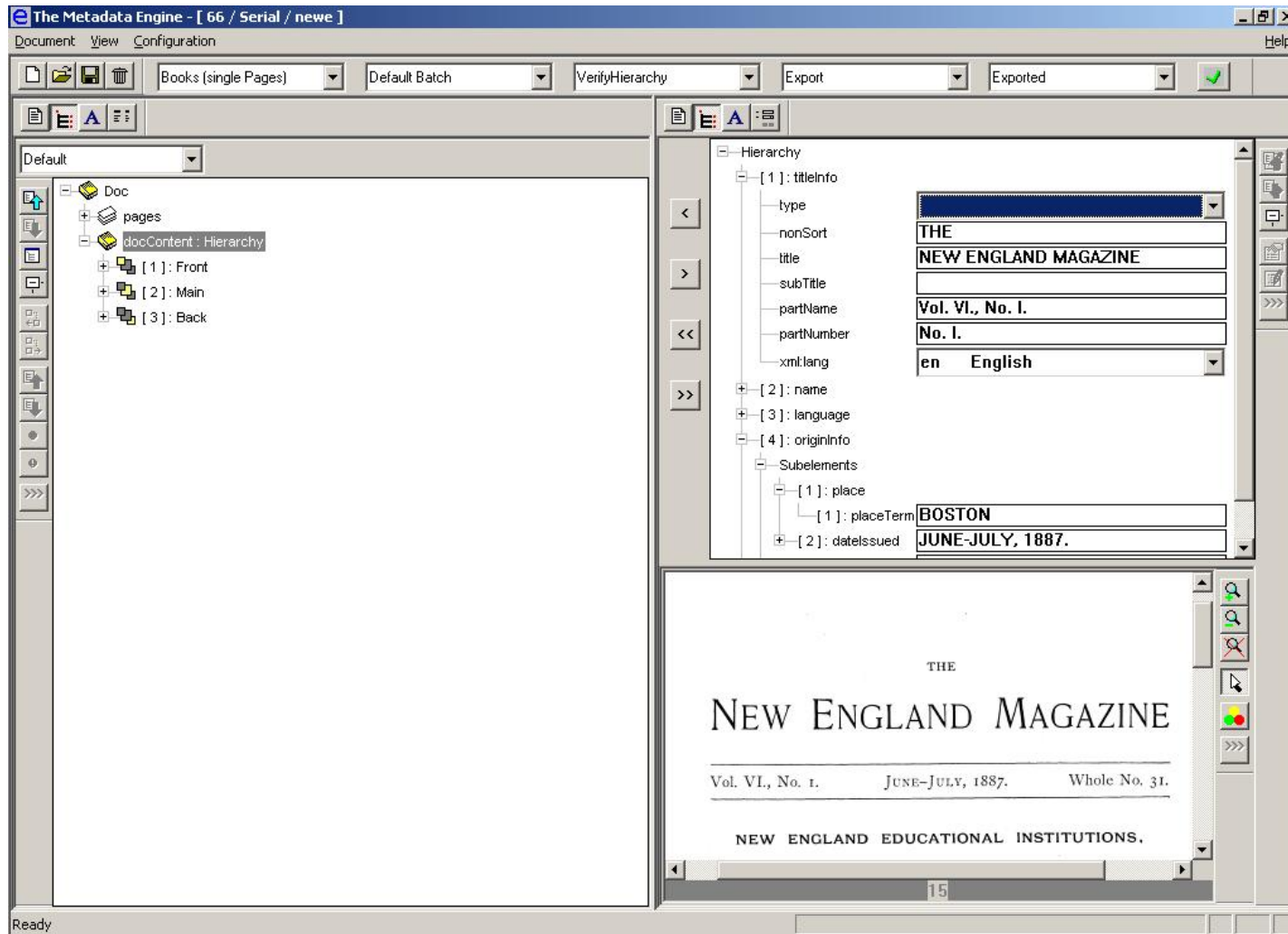
The figure presents the experimental set up of the system as well as a brief description of each component within the system. As illustrated in figure 21, the system consists of three main modules: the Character Recognition, the Measurement Instruments and the Software Module.

2.1 The Measurement Instruments

An essential module, the system uses a set of specialized measurement instruments. In particular, an HP 34401A multimeter (precision) is used to record the response magnitude of the circuit under investigation and an HP 3456A digital multimeter (precision) is used to produce the data used to drive the circuit. For circuit modelling (simulation) software under the HP 3456A digital multimeter was also developed up to the system for measurement and analysis. However, since most of the interesting information of the higher frequencies, the dynamic signal analyzer has been

25 26 27

Manual Editing of Descriptive Metadata / Volume



The screenshot shows the 'The Metadata Engine' software interface. The window title is 'The Metadata Engine - [66 / Serial / newe]'. The interface is divided into several sections:

- Top Bar:** Contains menu options (Document, View, Configuration) and a Help button. Below this is a toolbar with icons for file operations and a series of dropdown menus: 'Books (single Pages)', 'Default Batch', 'VerifyHierarchy', 'Export', and 'Exported'.
- Left Panel:** A tree view showing the document structure. The root is 'Doc', which contains 'pages' and 'docContent : Hierarchy'. Under 'docContent : Hierarchy', there are three items: '[1] : Front', '[2] : Main', and '[3] : Back'.
- Right Panel (Metadata Editor):** Displays a hierarchical tree of metadata elements. The selected element is '[1] : titleInfo', which includes fields for:
 - type (dropdown menu)
 - nonSort: THE
 - title: NEW ENGLAND MAGAZINE
 - subTitle
 - partName: Vol. VI., No. I.
 - partNumber: No. I.
 - xml:lang: en English
- Bottom Panel:** A preview of the document's title page. The text on the page reads:

THE
NEW ENGLAND MAGAZINE
Vol. VI., No. I. JUNE-JULY, 1887. Whole No. 31.
NEW ENGLAND EDUCATIONAL INSTITUTIONS.

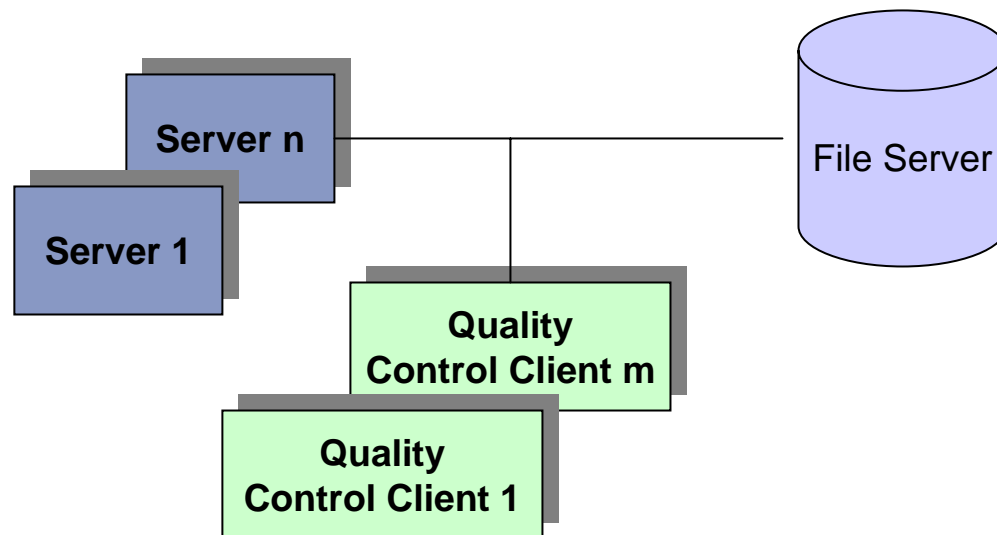
The status bar at the bottom left indicates 'Ready'.

Manual Editing of Descriptive Metadata / Illustration

The screenshot displays the 'The Metadata Engine' software interface. The window title is 'The Metadata Engine - [66 / Serial / neue]'. The interface is divided into several sections:

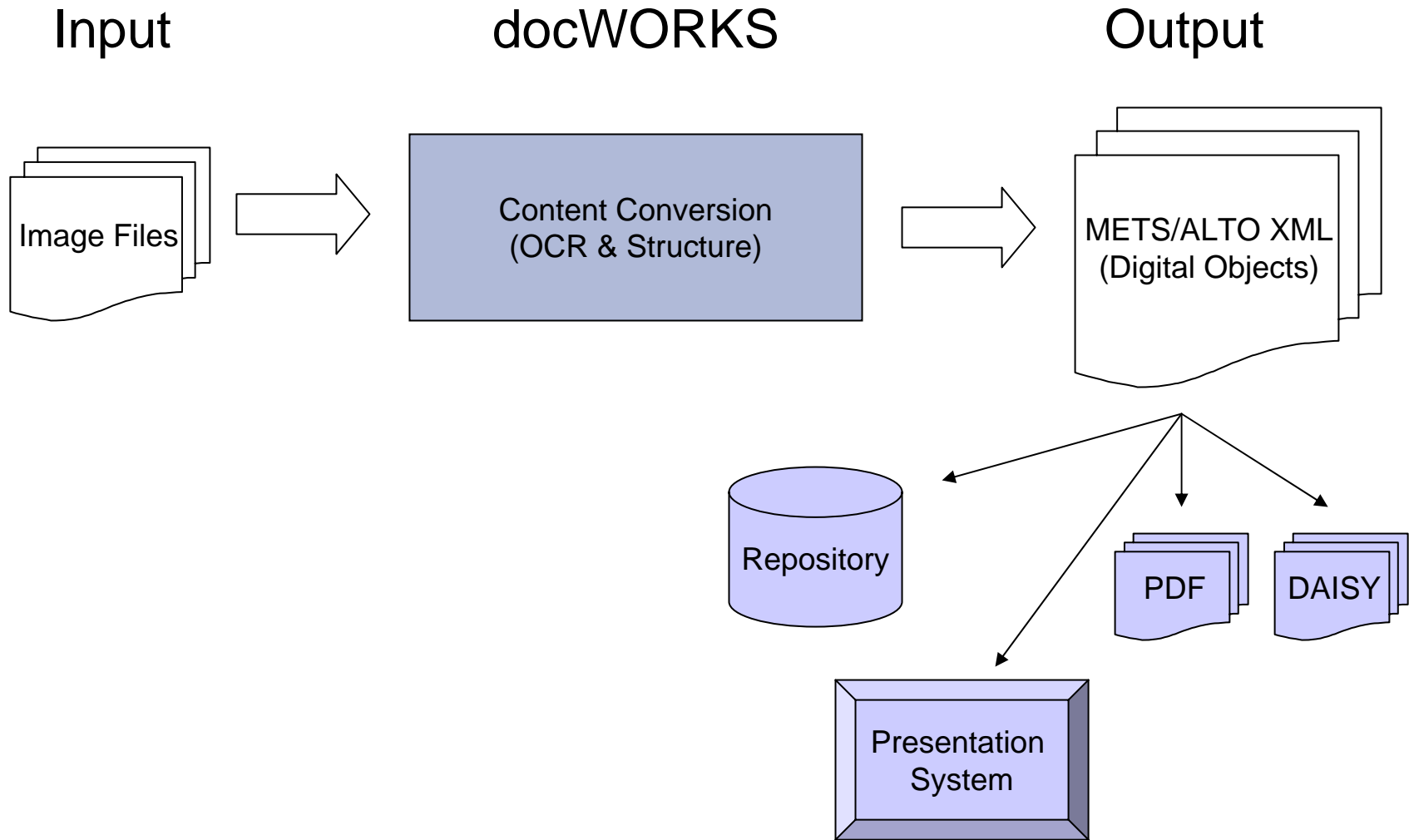
- Top Bar:** Contains menu options (Document, View, Configuration) and a Help button. Below this is a toolbar with icons for file operations and a series of dropdown menus: 'Books (single Pages)', 'Default Batch', 'VerifyHierarchy', 'Export', and 'Exported'.
- Left Panel:** A tree view showing the document structure. The selected path is: Doc > pages > docContent : Hierarchy > [1] : Front > [2] : Main > [1] : Issue > Content > [1] : Front > [2] : Main > [1] : Contribution (1) 'NEW ENGLAND EDUCATIONAL INS...'. Underneath, there is a list of illustrations, each with a dropdown menu currently set to 'Portrait'.
- Right Panel:** A metadata editing form for the selected illustration. It shows fields for:
 - type: (dropdown menu)
 - nonSort: (text field)
 - title: 'GAY HEAD, MARTHA'S VINEYARD'
 - subTitle: (text field)
 - xml:lang: 'en English' (dropdown menu)
- Bottom Right:** A preview window showing a black and white illustration of a coastal landscape with a lighthouse. Below the image is the caption 'GAY HEAD, MARTHA'S VINEYARD.' and a page indicator '20 [8]'.

Scalable Client / Server Architecture



- Auto-Import
- Image Preprocessing
- Layout Analysis
- OCR
- Structural Analysis
- Export

In Practice: docWORKS in Digitization Projects



KBDK Markup Policy

148 Page 0 Cic Markup
Running title

012 Est-ce que Pierre a cessé de parler?
013 Est-ce que Marie est partie?

chapter **Les verbes implicatifs négatifs**

Les verbes implicatifs négatifs ont cette caractéristique que le compléminent est faux quand la principale est positive et vice-versa. Ils obéissent aux postulats de sens (114) et (115):

Formula → **Marked as Formula because of signs.**

014 * (S) ⇒ ~ S et (S) ont une condition suffisante pour ~ S
015 ~ * (S) ⇒ S et (S) ont une condition nécessaire pour ~ S

Ainsi, lorsqu'un locuteur énonce (116 a), il ne peut pas nier (116 b) sans se contredire. De même, quand il énonce (117 a), il ne peut pas nier (117 b) sans contradiction:

(116) (a) Pierre a cessé de parler.
 (b) Pierre ne parle plus.
(117) (a) Pierre n'a pas cessé de parler.
 (b) Pierre parle.

Comme les verbes implicatifs positifs, les verbes implicatifs négatifs peuvent d'abord se diviser en deux groupes principaux exemplifiés par *cesser de* et *discontinuer de*.

Le premier groupe peut encore se diviser en trois sous-groupes:

Table

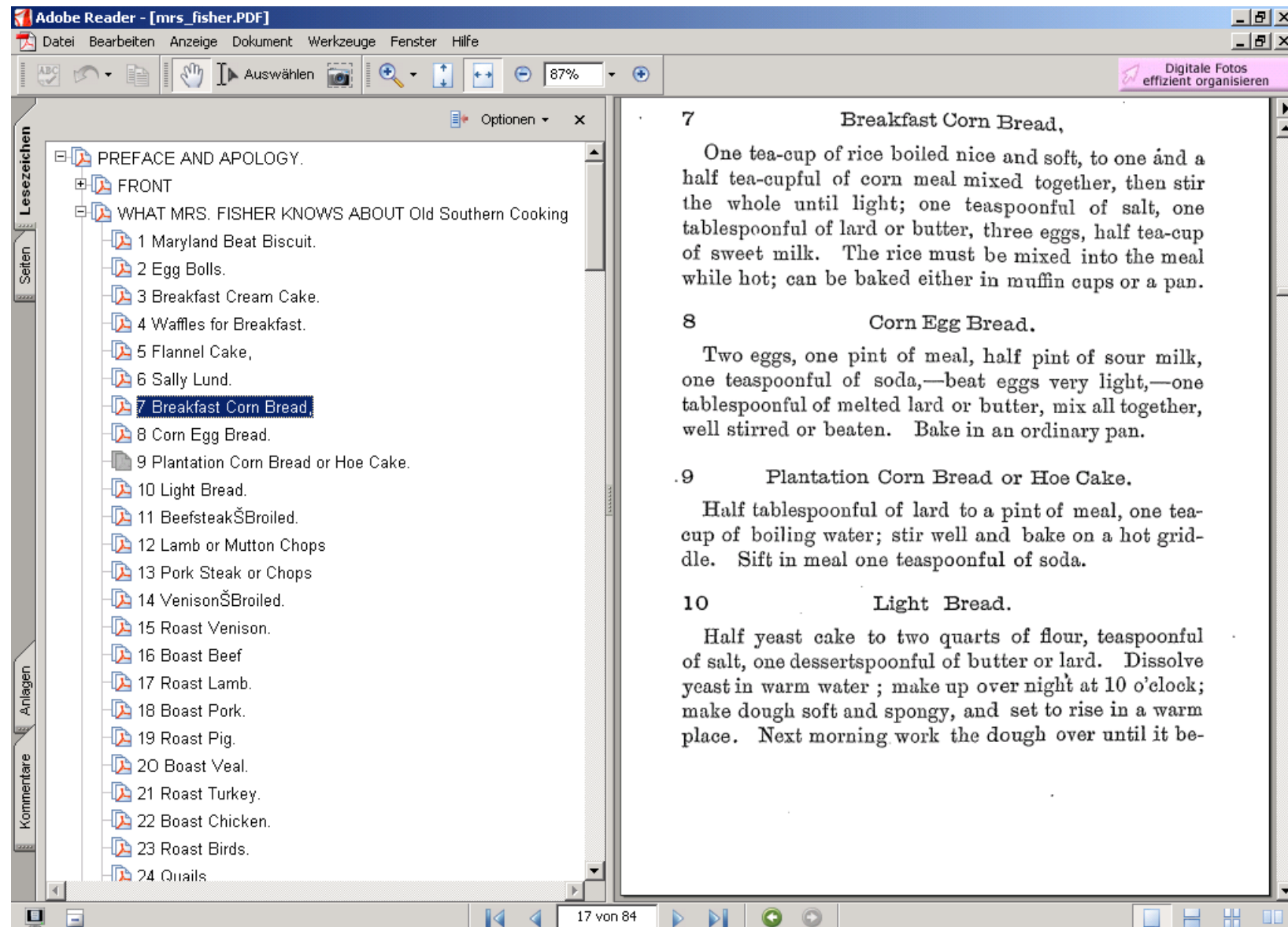
1°. Le premier sous-groupe comprend:		
adhérer de	cesser de	finir de
s'accrocher de	désordonner de	s'interrompre de
2°. Le deuxième sous-groupe comprend:		
décevoir à		
3°. Le troisième sous-groupe comprend:		
s'abstenir de	s'excepter de	se priver de
s'excuser de	faillir (à)	se priver de
dédaigner de	ne pas laisser de	refuser de
ôffenser de	manquer à/à	renoncer à
se dispenser de	négliger de	se réserver de
épargner de	construire de	résister à
éviter de/que	oublier de	

28: Ce verbe se rencontre seulement dans des constructions négatives.

Footnote

Page part of Contribution in Main

Presentation of Converted Material



The screenshot shows the Adobe Reader interface. The left sidebar contains a table of contents with the following items:

- PREFACE AND APOLOGY.
- FRONT
- WHAT MRS. FISHER KNOWS ABOUT Old Southern Cooking
 - 1 Maryland Beat Biscuit.
 - 2 Egg Bolls.
 - 3 Breakfast Cream Cake.
 - 4 Waffles for Breakfast.
 - 5 Flannel Cake,
 - 6 Sally Lund.
 - 7 Breakfast Corn Bread,**
 - 8 Corn Egg Bread.
 - 9 Plantation Corn Bread or Hoe Cake.
 - 10 Light Bread.
 - 11 BeefsteakŠBroiled.
 - 12 Lamb or Mutton Chops
 - 13 Pork Steak or Chops
 - 14 VenisonŠBroiled.
 - 15 Roast Venison.
 - 16 Boast Beef
 - 17 Roast Lamb.
 - 18 Boast Pork.
 - 19 Roast Pig.
 - 20 Boast Veal.
 - 21 Roast Turkey.
 - 22 Boast Chicken.
 - 23 Roast Birds.
 - 24 Quails

The main content area displays page 7, titled "Breakfast Corn Bread," with the following text:

7 **Breakfast Corn Bread,**

One tea-cup of rice boiled nice and soft, to one and a half tea-cupful of corn meal mixed together, then stir the whole until light; one teaspoonful of salt, one tablespoonful of lard or butter, three eggs, half tea-cup of sweet milk. The rice must be mixed into the meal while hot; can be baked either in muffin cups or a pan.

8 **Corn Egg Bread.**

Two eggs, one pint of meal, half pint of sour milk, one teaspoonful of soda,—beat eggs very light,—one tablespoonful of melted lard or butter, mix all together, well stirred or beaten. Bake in an ordinary pan.

9 **Plantation Corn Bread or Hoe Cake.**

Half tablespoonful of lard to a pint of meal, one tea-cup of boiling water; stir well and bake on a hot grid-dle. Sift in meal one teaspoonful of soda.

10 **Light Bread.**

Half yeast cake to two quarts of flour, teaspoonful of salt, one dessertspoonful of butter or lard. Dissolve yeast in warm water ; make up over night at 10 o'clock; make dough soft and spongy, and set to rise in a warm place. Next morning work the dough over until it be-

The bottom status bar shows "17 von 84" and navigation icons.

Presentation of Converted Material

Adobe Reader - [virtual_landscapes.PDF]

Datei Bearbeiten Anzeige Dokument Werkzeuge Fenster Hilfe

Auswählen 100%

Adobe PDF vom Desktop aus erstellen

Leseseiten

- "TEXAS GRANITES"
- Illustrations
 - Plate 1. Texas State Capitol Building, Austin. Built with red granite
 - Plate 2. Gray granite quarry operated by Gooch & Company, 6 m
 - Plate 3. The Llano quarrymen experience considerable difficulty
 - Plate 4. Taking Mocks of red granite out of Granite Mountain
 - Plate 5. Darragh Brothers' quarry at Granite Mountain, Burnet Co
- Tables

Seiten

Anlagen

Kommentare

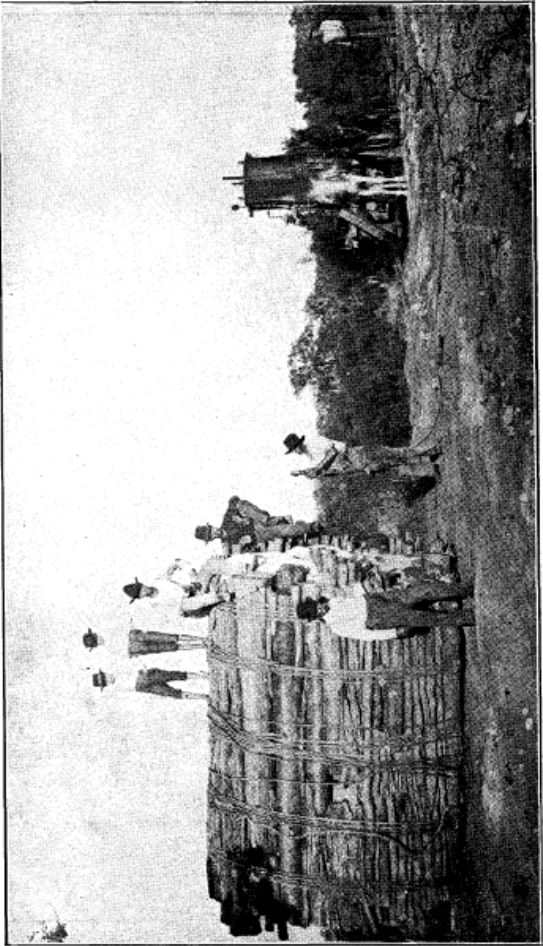


Plate 3. The Llano quarrymen experience considerable difficulty getting their granite to the market. Rolling a 35-ton block eight miles to Llano for shipment.

162,6 x 240,6 mm

12 von 14

Presentation of Converted Material

PDS Ladies' magazine. - Microsoft Internet Explorer

[Datei](#) [Bearbeiten](#) [Ansicht](#) [Favoriten](#) [Extras](#) ?

[Zurück](#) [Suchen](#) [Favoriten](#) [Wechseln zu](#) [Links](#) [SnagIt](#)

Adresse <http://pdstest.harvard.edu:8080/pdx/servlet/pds?op=f&id=410910&n=246&s=4>

Harvard University Library **Ladies' magazine.**
 Page Delivery Service

[Search](#) [View Text](#) [Printable Version](#) [Related Links](#) [Help](#) [Copyright](#)

Page Number: [Go to](#) Sequence Number: [Go to](#)

Ladies' magazine. [591 pages.]

- "THE LADIES' MAGAZINE."** [pp. unnumbered]
 - [Blank Page] [pp. unnumbered page-unnumbered]
 - [Blank Page] [pp. unnumbered page-unnumbered]
 - [Picture Page] [pp. unnumbered page-unnumbered]
 - [Title Page] [pp. unnumbered page-unnumbered]
 - [Statement Page] [pp. unnumbered page-unnumbered]
 - JANUARY. [pp. 1-49 (seq. 6-54)]
 - FEBRUARY. [pp. 49-97 (seq. 54-102)]
 - MARCH. [pp. 97-144 (seq. 102-149)]
 - APRIL. [pp. 145-193 (seq. 150-198)]
 - MAY. [pp. 193-240 (seq. 198-245)]
 - JUNE.** [pp. 241-289 (seq. 246-295)]
 - JULY. [pp. 289-337 (seq. 295-343)]
 - AUGUST. [pp. 337-384 (seq. 343-390)]
 - SEPTEMBER. [pp. 385-432 (seq. 391-438)]
 - OCTOBER. [pp. 433-480 (seq. 439-486)]
 - NOVEMBER. [pp. 481-529 (seq. 487-535)]
 - DECEMBER. [pp. 529-576 (seq. 535-582)]
 - INDEX. [pp. 3-8 (seq. 583-588)]
 - [Blank Page] [pp. unnumbered page-unnumbered]
 - [Blank Page] [pp. unnumbered page-unnumbered]
 - [Blank Page] [pp. unnumbered page-unnumbered]

LADIES' MAGAZINE.

Vol. I. JUNE. No. VI.

DRESS.

"The world is still deceived with ornament." So said William Shakspeare, and two centuries have made, in this respect, little alteration. There seems to be, in mankind, a propensity to display, to prize outward show, to look with favor on the wearer of a fine suit, rather than on the merit of him who deserves one.

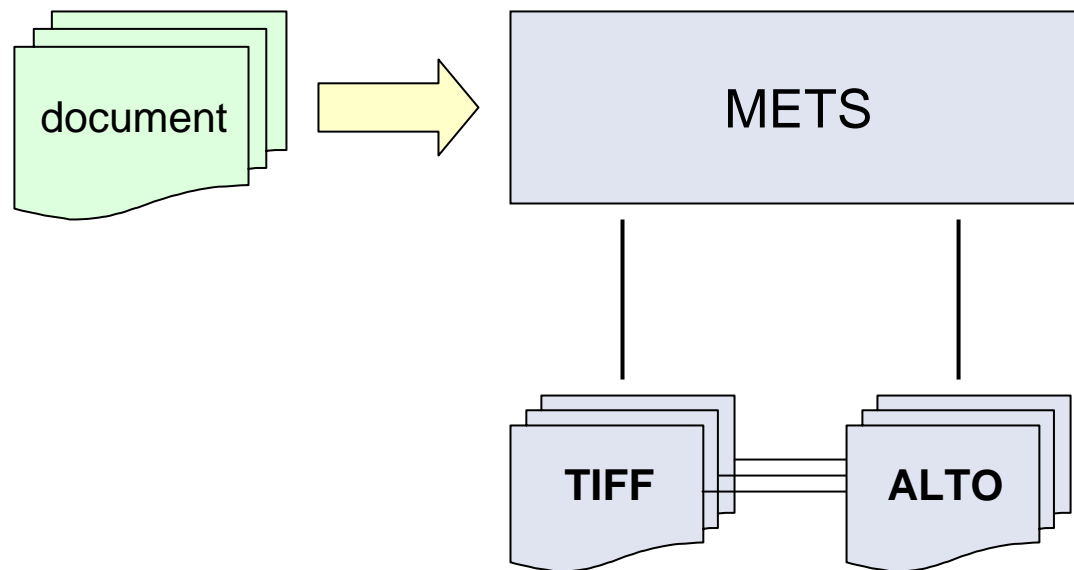
But such remarks have been made by every cynic since the days of Antisthenes, and I did not commence with the cynical intention of railing in "good set terms" against the modes and customs of the world. I believe, with the melancholy Jaques, that

—"Who cries out on fashion
 That can therein tax any private party!—
 Doth it not flow as largely as the sea,
 Till that the very, very means do ebb?"

This ebbing of the means is one of the most disagreeable drawbacks on a life of dissipation and display. Yet it is often salutary in its operation. When the tree is bowed by the fury of the storm, if, instead of sinking beneath the shock, the roots entwine and fix themselves more firmly during the agitation, then, when the tempest is over, that tree will rise again more healthy and vigorous. Just so

Internet

METS / ALTO XML



ALTO – Analyzed Layout and Text Object

METS

- Header
- DC/MODS descriptive metadata
- NISO 39.087 (mix), technical metadata
- Structural Map: Physical Structure
- Structural Map: Logical Structure

ALTO

■ Styles

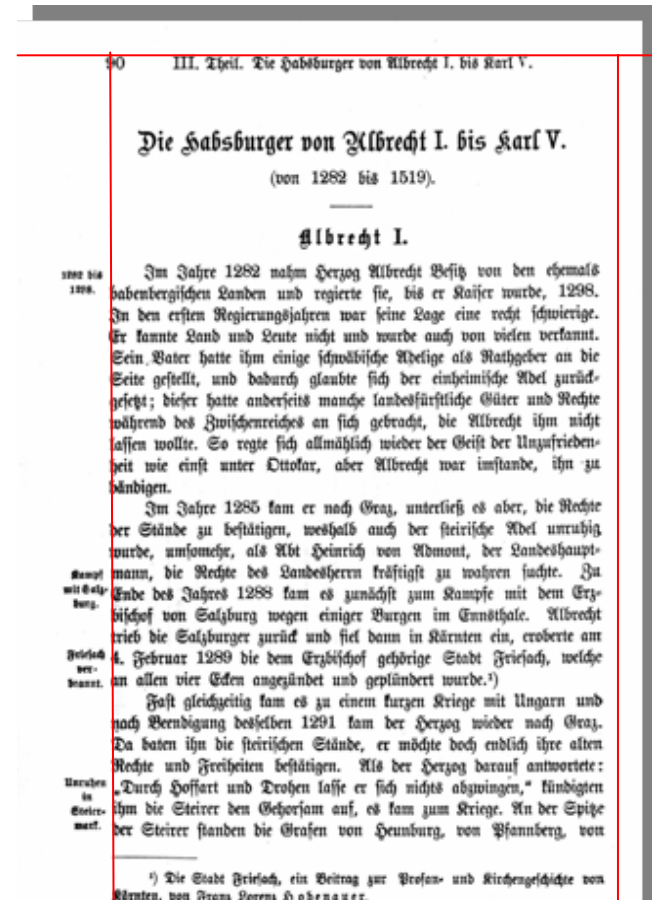
- Paragraph (alignment, linespacing, etc.)
- Font (name, size, bold, italic, etc.)

■ Layout

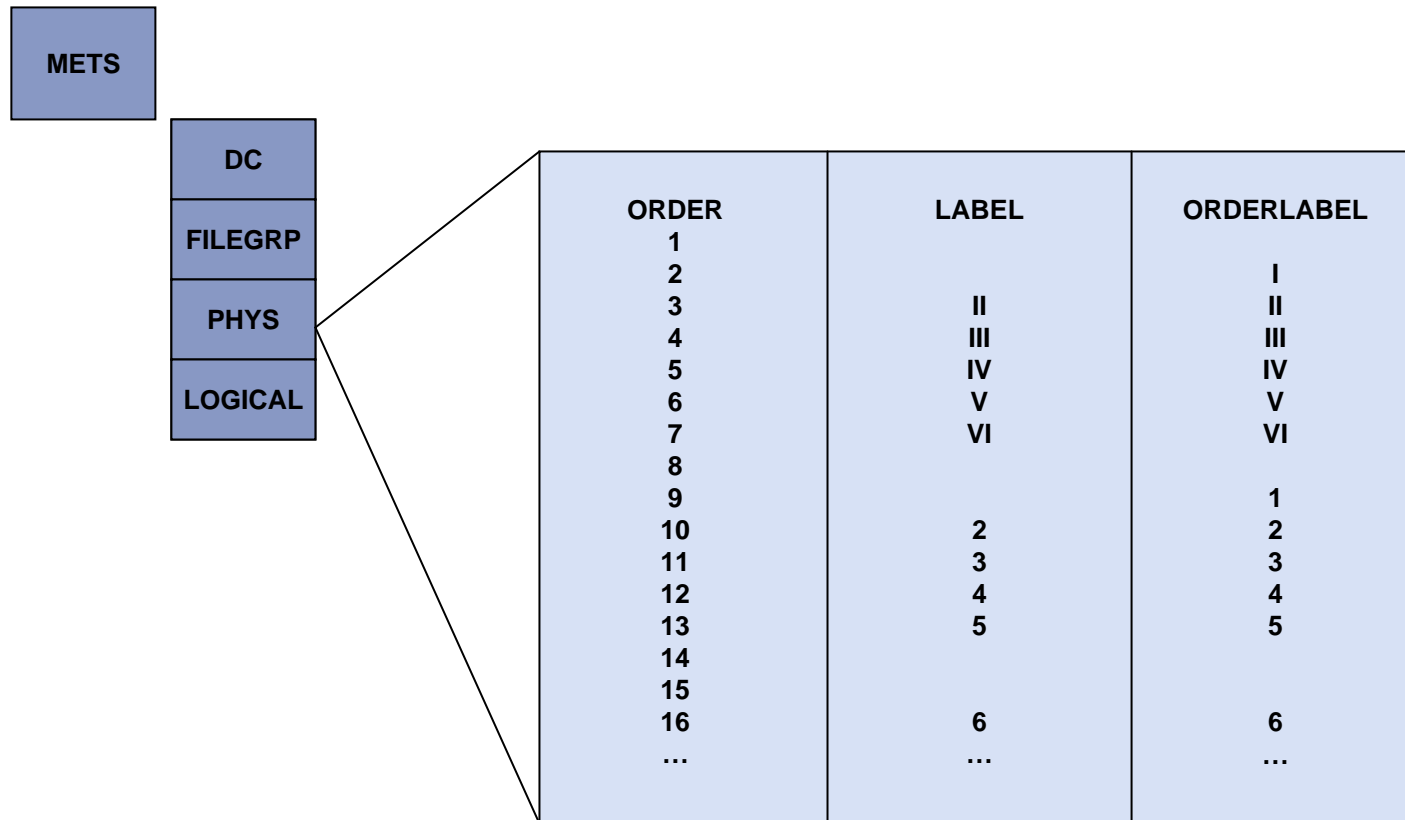
- Printspace
- TopMargin
- InnerMargin
- OuterMargin
- BottomMargin

■ Objects in 5 areas above:

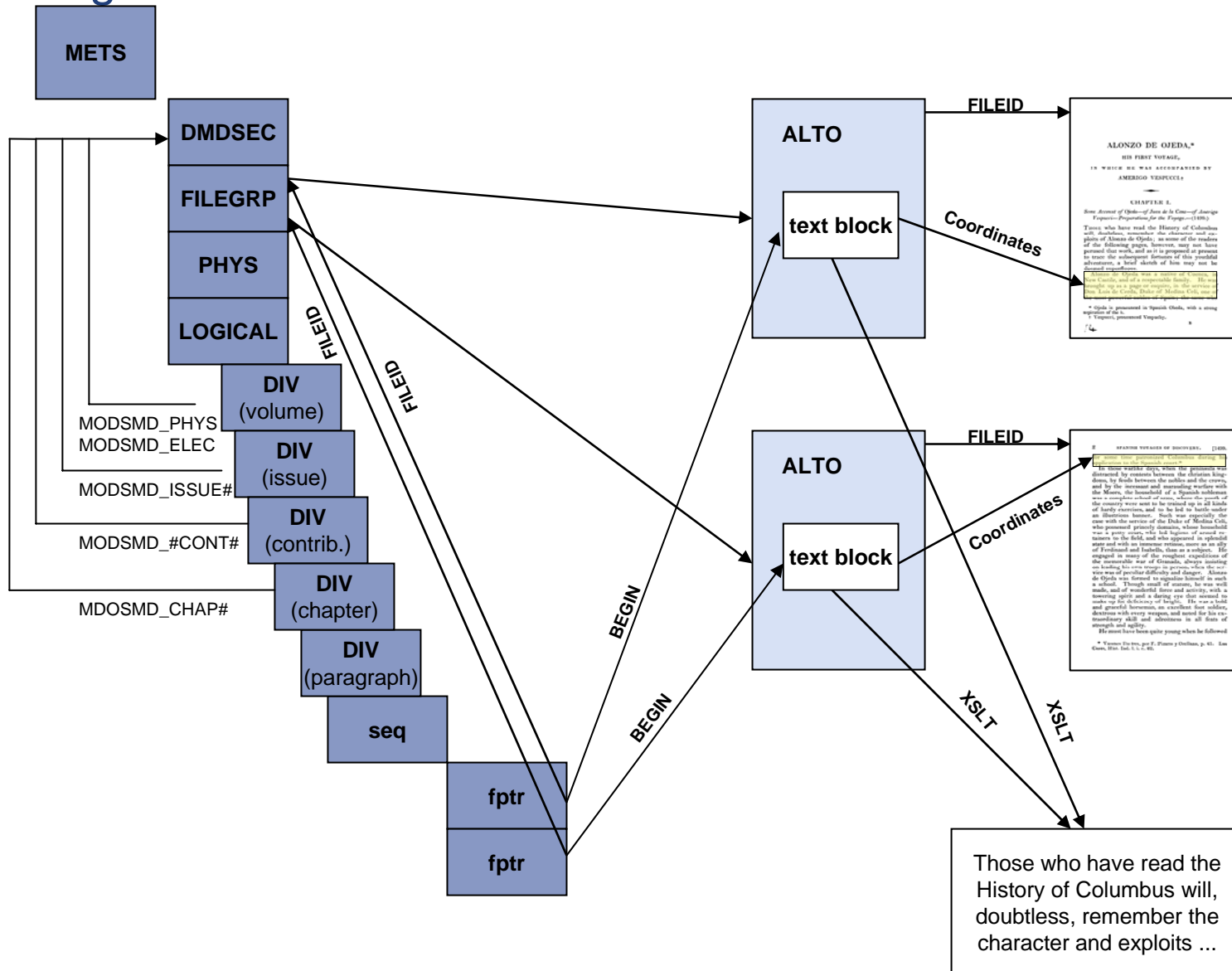
- Text block
 - Text lines
 - Strings [coordinates, string (as printed), substitution (hyphenation)]
 - Spaces
- Composed block
- Picture
- Table
- Formula



METS / physical structure



METS / logical structure



ALTO / Page Layout and Text Content



```

D:\CCS\METAe\Metae\Export\EXP00099\EXP00099-ALTO00023.xml - Microsoft Internet Explorer
Datei Bearbeiten Ansicht Favoriten Extras ? Links »

<?xml version="1.0" encoding="UTF-8" ?>
- <alto xmlns:xsi="http://www.w3.org/2000/10/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://meta-
e.uibk.ac.at/schemas/alto.xsd" xmlns:xlink="http://www.w3.org/TR/xlink">
- <Styles>
  <TextStyle ID="TXT_0" FONTSIZE="12" FONTFAMILY="Times New Roman" />
  <TextStyle ID="TXT_1" FONTSIZE="8" FONTFAMILY="Times New Roman" />
  <TextStyle ID="TXT_2" FONTSIZE="6" FONTFAMILY="Times New Roman" />
  <TextStyle ID="TXT_3" FONTSIZE="9" FONTFAMILY="Times New Roman" />
  <TextStyle ID="TXT_4" FONTSIZE="7" FONTFAMILY="Times New Roman" />
  <ParagraphStyle ID="PAR_LEFT" ALIGN="Left" />
  <ParagraphStyle ID="PAR_BLOCK" ALIGN="Block" />
</Styles>
- <Layout>
- <Page ID="PAGE23" PHYSICAL_IMG_NR="23" HEIGHT="3225" WIDTH="2104">
  <TopMargin ID="P23_TM00001" HPOS="0" VPOS="0" WIDTH="2104" HEIGHT="197" />
  <InnerMargin ID="P23_IM00001" HPOS="0" VPOS="197" WIDTH="163" HEIGHT="2790" />
  <OuterMargin ID="P23_OM00001" HPOS="1828" VPOS="197" WIDTH="276" HEIGHT="2790" />
  <BottomMargin ID="P23_BM00001" HPOS="0" VPOS="2987" WIDTH="2104" HEIGHT="238" />
- <PrintSpace ID="P23_PS00001" HPOS="163" VPOS="197" WIDTH="1665" HEIGHT="2790">
- <TextBlock ID="P23_TB00001" HPOS="455" VPOS="588" WIDTH="1076" HEIGHT="85" STYLEREFs="TXT_0 PAR_LEFT">
  - <TextLine ID="P23_TL00001" HPOS="193" VPOS="249" WIDTH="456" HEIGHT="36">
    <String ID="P23_ST00001" HPOS="193" VPOS="249" WIDTH="178" HEIGHT="30" CONTENT="ALONZO" />
    <SP ID="P23_SP00001" HPOS="371" VPOS="279" WIDTH="24" />
    <String ID="P23_ST00002" HPOS="395" VPOS="249" WIDTH="58" HEIGHT="30" CONTENT="DE" />
    <SP ID="P23_SP00002" HPOS="453" VPOS="279" WIDTH="25" />
    <String ID="P23_ST00003" HPOS="478" VPOS="249" WIDTH="171" HEIGHT="36" CONTENT="OJEDA,*" />
  </TextLine>
  </TextBlock>
- <TextBlock ID="P23_TB00002" HPOS="643" VPOS="759" WIDTH="707" HEIGHT="61" STYLEREFs="TXT_1 PAR_LEFT">
  - <TextLine ID="P23_TL00002" HPOS="272" VPOS="321" WIDTH="300" HEIGHT="27">
    <String ID="P23_ST00004" HPOS="272" VPOS="322" WIDTH="49" HEIGHT="20" CONTENT="HIS" />
    <SP ID="P23_SP00003" HPOS="321" VPOS="342" WIDTH="17" />
    <String ID="P23_ST00005" HPOS="338" VPOS="322" WIDTH="86" HEIGHT="20" CONTENT="FIRST" />
    <SP ID="P23_SP00004" HPOS="424" VPOS="342" WIDTH="16" />
    <String ID="P23_ST00006" HPOS="440" VPOS="321" WIDTH="132" HEIGHT="27" CONTENT="VOYAGE," />
  </TextLine>
  </TextBlock>
- <TextBlock ID="P23_TB00003" HPOS="255" VPOS="918" WIDTH="1484" HEIGHT="40" STYLEREFs="TXT_2 PAR_LEFT">
  - <TextLine ID="P23_TL00003" HPOS="108" VPOS="980" WIDTH="620" HEIGHT="17">

```

ALTO / Hyphenated Word

```

D:\CCS\METAe\Metae\Export\EXP00099\EXP00099-ALTO00023.xml - Microsoft Internet Explorer
Datei Bearbeiten Ansicht Favoriten Extras ? Links >>

</TextBlock>
- <TextBlock ID="P23_TB00007" HPOS="168" VPOS="1709" WIDTH="1651" HEIGHT="619" STYLEREFs="TXT_3 PAR_BLOCK">
- <TextLine ID="P23_TL00008" HPOS="72" VPOS="723" WIDTH="698" HEIGHT="32">
  <String ID="P23_ST00038" HPOS="72" VPOS="724" WIDTH="88" HEIGHT="24" CONTENT="THOSE" />
  <SP ID="P23_SP00023" HPOS="160" VPOS="748" WIDTH="19" />
  <String ID="P23_ST00039" HPOS="179" VPOS="724" WIDTH="55" HEIGHT="23" CONTENT="who" />
  <SP ID="P23_SP00024" HPOS="234" VPOS="748" WIDTH="17" />
  <String ID="P23_ST00040" HPOS="251" VPOS="724" WIDTH="62" HEIGHT="23" CONTENT="have" />
  <SP ID="P23_SP00025" HPOS="313" VPOS="748" WIDTH="18" />
  <String ID="P23_ST00041" HPOS="331" VPOS="723" WIDTH="60" HEIGHT="24" CONTENT="read" />
  <SP ID="P23_SP00026" HPOS="391" VPOS="748" WIDTH="17" />
  <String ID="P23_ST00042" HPOS="408" VPOS="724" WIDTH="40" HEIGHT="23" CONTENT="the" />
  <SP ID="P23_SP00027" HPOS="448" VPOS="748" WIDTH="19" />
  <String ID="P23_ST00043" HPOS="467" VPOS="723" WIDTH="105" HEIGHT="32" CONTENT="History" />
  <SP ID="P23_SP00028" HPOS="572" VPOS="755" WIDTH="18" />
  <String ID="P23_ST00044" HPOS="590" VPOS="724" WIDTH="30" HEIGHT="23" CONTENT="of" />
  <SP ID="P23_SP00029" HPOS="620" VPOS="755" WIDTH="13" />
  <String ID="P23_ST00045" HPOS="633" VPOS="724" WIDTH="137" HEIGHT="25" CONTENT="Columbus" />
</TextLine>
- <TextLine ID="P23_TL00009" HPOS="72" VPOS="756" WIDTH="697" HEIGHT="29">
  <String ID="P23_ST00046" HPOS="72" VPOS="757" WIDTH="55" HEIGHT="28" CONTENT="will," />
  <SP ID="P23_SP00030" HPOS="127" VPOS="785" WIDTH="20" />
  <String ID="P23_ST00047" HPOS="147" VPOS="757" WIDTH="134" HEIGHT="28" CONTENT="doubtless," />
  <SP ID="P23_SP00031" HPOS="281" VPOS="785" WIDTH="19" />
  <String ID="P23_ST00048" HPOS="300" VPOS="757" WIDTH="136" HEIGHT="23" CONTENT="remember" />
  <SP ID="P23_SP00032" HPOS="436" VPOS="785" WIDTH="22" />
  <String ID="P23_ST00049" HPOS="458" VPOS="756" WIDTH="41" HEIGHT="24" CONTENT="the" />
  <SP ID="P23_SP00033" HPOS="499" VPOS="785" WIDTH="18" />
  <String ID="P23_ST00050" HPOS="517" VPOS="757" WIDTH="124" HEIGHT="24" CONTENT="character" />
  <SP ID="P23_SP00034" HPOS="641" VPOS="785" WIDTH="18" />
  <String ID="P23_ST00051" HPOS="659" VPOS="758" WIDTH="48" HEIGHT="23" CONTENT="and" />
  <SP ID="P23_SP00035" HPOS="707" VPOS="785" WIDTH="18" />
  <String ID="P23_ST00052" HPOS="725" VPOS="767" WIDTH="44" HEIGHT="15" CONTENT="lex"
    SUBS_TYPE="HypPart1" SUBS_CONTENT="exploits" />
  <HYP CONTENT="-" />
</TextLine>
- <TextLine ID="P23_TL00010" HPOS="72" VPOS="789" WIDTH="697" HEIGHT="33">
  <String ID="P23_ST00053" HPOS="72" VPOS="790" WIDTH="73" HEIGHT="32" CONTENT="ploits"
    SUBS_TYPE="HypPart2" SUBS_CONTENT="exploits" />
  <SP ID="P23_SP00036" HPOS="145" VPOS="822" WIDTH="12" />
  <String ID="P23_ST00054" HPOS="157" VPOS="790" WIDTH="29" HEIGHT="23" CONTENT="of" />
    
```



lex

ploits

ALTO / Hyphenated Word

AMERIGO VESPUCCI.†



CHAPTER I.

Some Account of Ojeda—of Juan de la Cosa—of Amerigo Vespucci—Preparations for the Voyage.—(1499.)

THOSE who have read the History of Columbus will, doubtless, remember the character and exploits of Alonzo de Ojeda; as some of the readers of the following pages, however, may not have perused that work, and as it is proposed at present to trace the subsequent fortunes of this youthful adventurer, a brief sketch of him may not be deemed superfluous.

Alonzo de Ojeda was a native of Cuenca. in

Summary: Milestones

- ➔ **METAe Project Start** - September 2000
- ➔ **Product Launch** - March 2003
- ➔ **Royal Danish Library** - July 2003
- ➔ **University Library of Lower Saxony (Göttingen)** - March 2004
- ➔ **Stanford University Library Evaluates docWORKS** - July 2004
- ➔ **Harvard University Library** - September 2004
- ➔ **LC contracts CCS for research project (NDNP)** - April 2005
- ➔ **LC evaluates docWORKS** - May 2005

Thank you!

Claus Gravenhorst
Daniel Lanz

claus.gravenhorst@ccs-gmbh.de
daniel.lanz@ccs-gmbh.de

Content Conversion Specialists
www.ccs-gmbh.de

<http://meta-e.uibk.ac.at/>