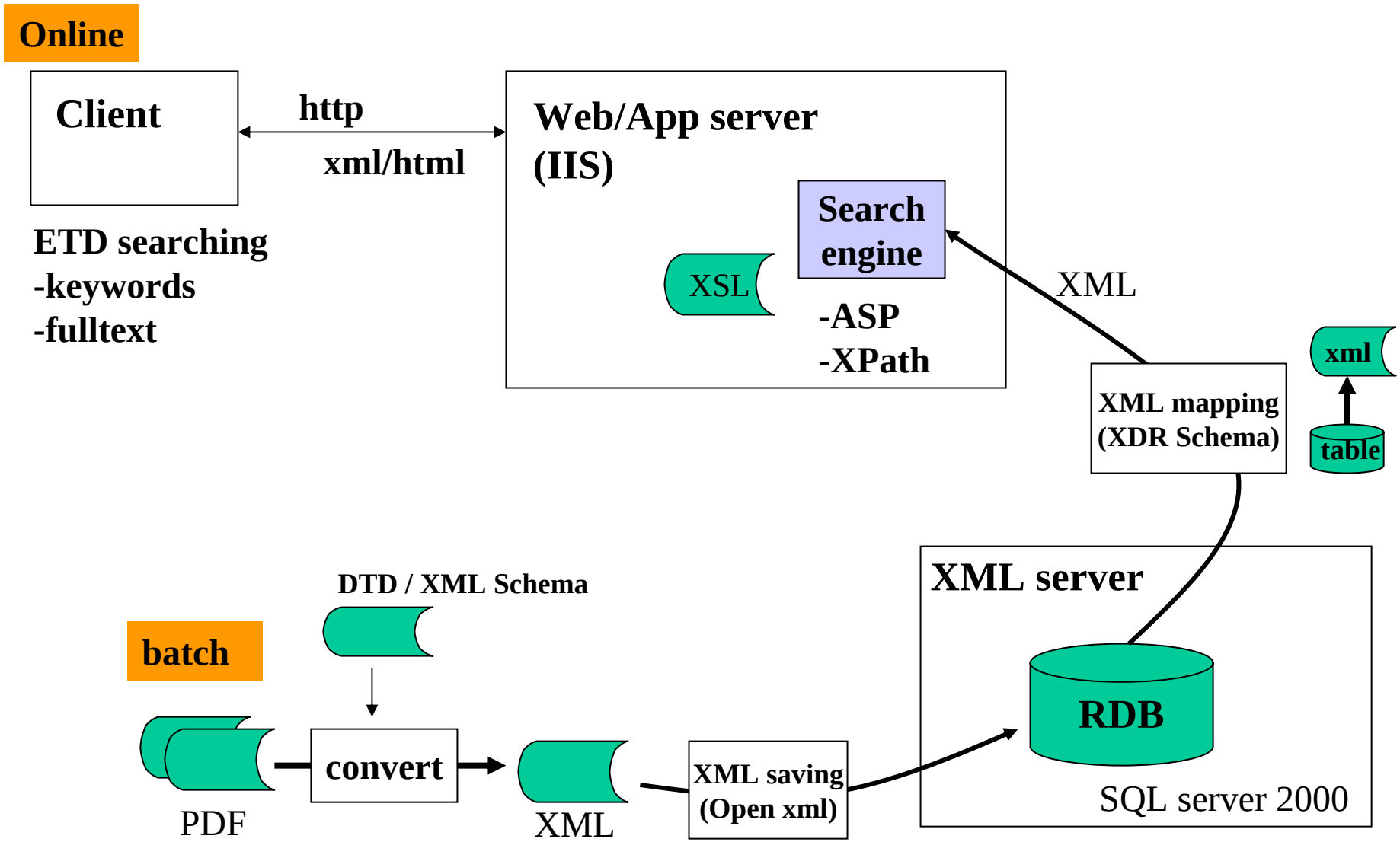
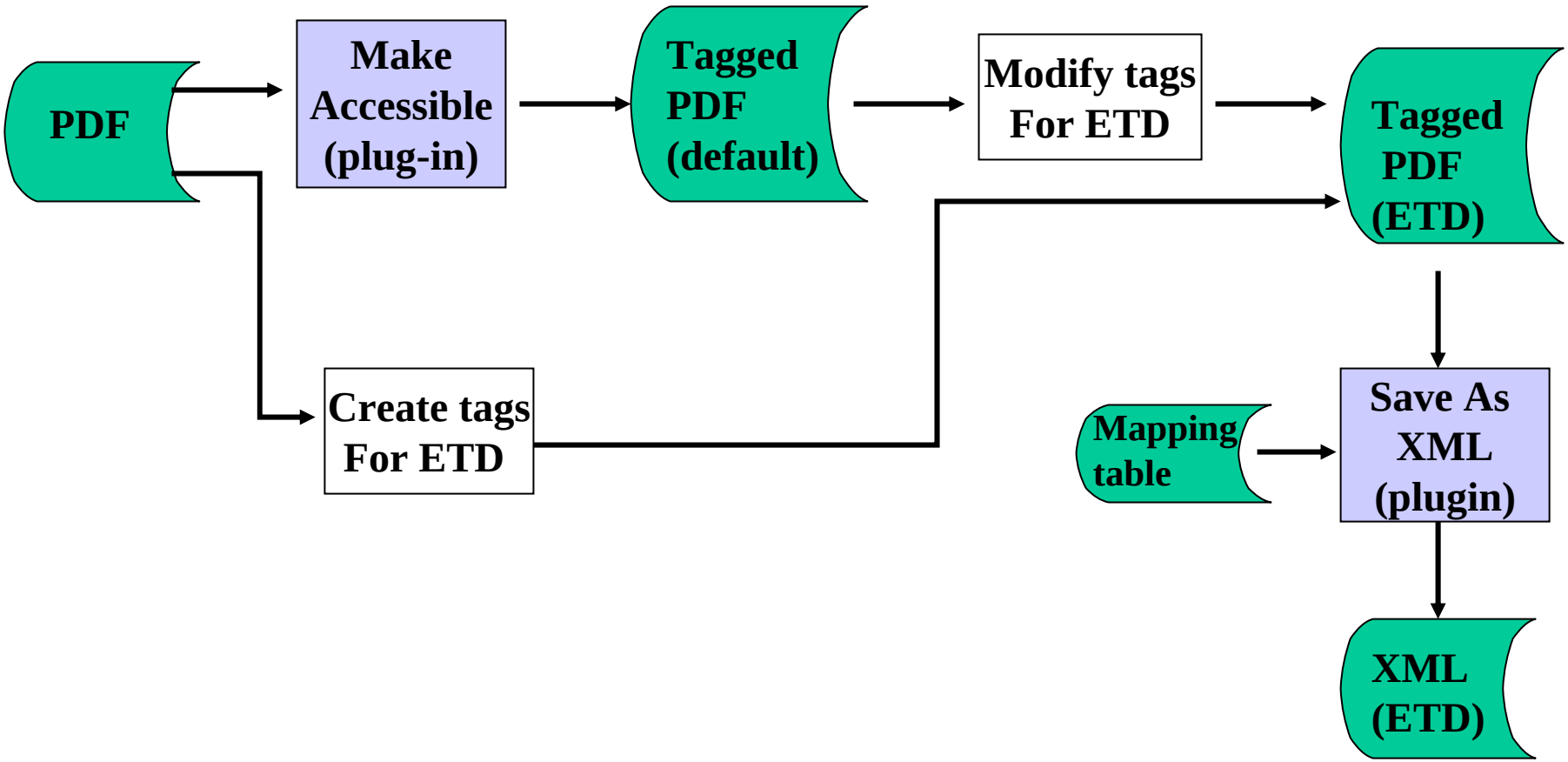
ETD DTD

- Based on ETD-ML1.1(Virginia Tech)
  - convert SGML DTD to XML DTD
- Added some tags
  -such as TableOfConetents, ListOfPicture, ListOfTables
- Added “id” attribute to some tags
  -such as bibliography, vita

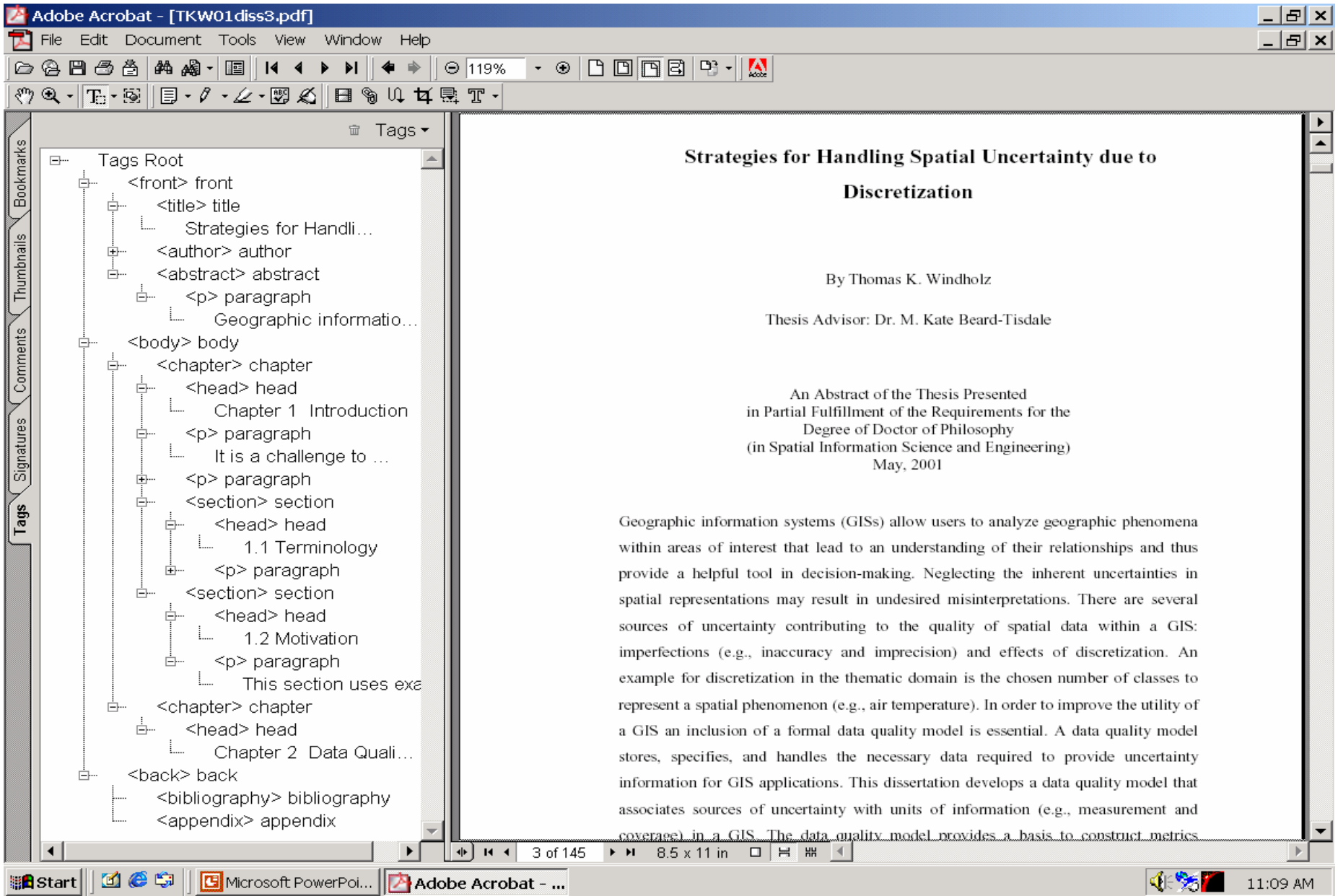
ETD System Structure



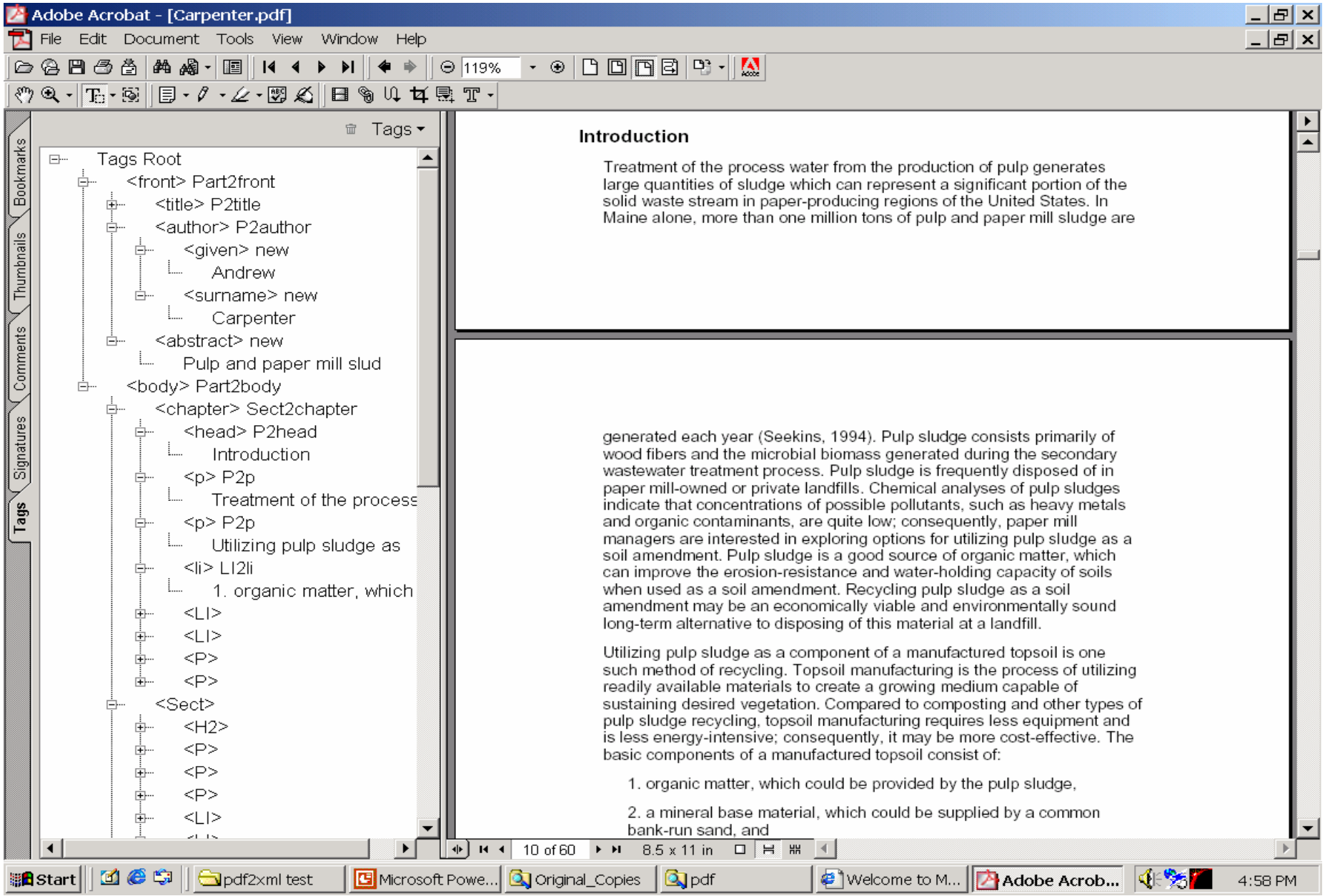
Convert PDF to XML
-- Adobe “save as xml”



Method1: Create tags for tagged PDF (ETD)



Method2: make default tagged PDF and modify it to match ETD



The XML output after Save as XML

C:\Documents and Settings\jiangma\My Documents\pdf2xml test\TKW01diss3.xml - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites History

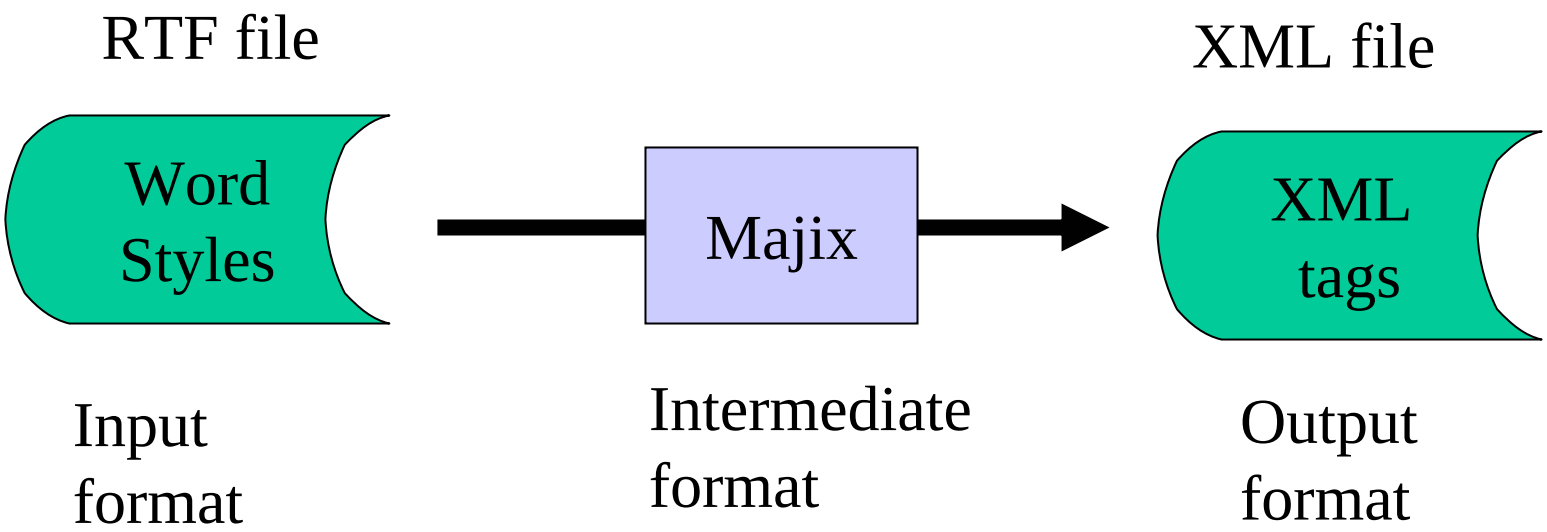
Address C:\Documents and Settings\jiangma\My Documents\pdf2xml test\TKW01diss3.xml Go Links

```
<?xml version="1.0" encoding="UTF-8" ?>
- <etd>
  - <front>
      <title>Strategies for Handling Spatial Uncertainty due to Discretization</title>
    - <author>
        <given>Thomas K.</given>
        <surname>Windholz</surname>
      </author>
    + <abstract>
    </front>
  - <body>
    - <chapter>
        <head>Chapter 1 Introduction</head>
        <p>It is a challenge to capture an infinite universe in finite systems. Geographic information
          systems (GISs) represent aspects of our world in finite computer systems.</p>
        <p>GISs model our reality in an immense variety of fields (Longley et al. 1999). GISs are used, for
          example, in utility (e.g., electric, water) management systems, models describing forest growth
          (from small regions to the entire world), and in medicine where GISs are utilized to map the
          human genome. GISs are vital to progress modeling, management, and the investigation of
          scientific as well as everyday phenomena. For all of these beneficiary applications we have to
          keep in mind that GISs are models of our world and are, therefore, constrained by our capability
          to model the complexity of processes and events.</p>
      - <section>
          <head>1.1 Terminology</head>
          <p>This section gives a brief overview of some key terms used within the thesis. This
            explanation of terminology should be seen as a condensed and simplified clarification.
            Throughout the thesis each of the terms will be discussed in more detail.</p>
        </section>
      - <section>
          <head>1.2 Motivation</head>
          <p>This section uses examples to give an overview of the effects of spatial discretization on the
            uncertainty of a representation. The motivation for the research presented in this dissertation
```

My Computer

Start Microsoft PowerPoi... Adobe Acrobat - [T... C:\Documents a... 11:18 AM

Convert RTF to XML
-- Majix



-can modify input style name
-can modify output xml tag name

-can not rearrange xml tree structure