

Data Integration

1. **Course number and name:** 020IDAES4 Data Integration
2. **Credits and contact hours:** 4 ECTS credits, 35 contact hours (lectures + labs)
3. **Instructor's or course coordinator's name:** Rima Kilany
4. **Text book:**
 - a. **Other supplemental materials:**
Handouts and course material posted on the Web
5. **Specific course information**
 - a. **Catalog description:**
This course details the main XML specifications, Namespaces, DTD validation, XMLSchema validation, XPATH, XSLT, XQUery, JAXP XML parsers (SAX, DOM), XML integration pipelines, Semantic Web specifications RDF/XML, OWL, SPARQL.
 - b. **Prerequisites or co-requisites:**
 - c. **Required:** Elective for CCE students
6. **Specific goals for the course**
 - a. **The student will be able to:**
 - Validate, transform, and search an XML document.
 - Parse an XML document in order to implement a Java application (Java) using JAXP (SAX, DOM) parsers.
 - Choose a database that supports XML datatypes, and choose the most appropriate mediation architecture (datawarehouse vs virtual database).
 - Implement an XML integration pipeline that could handle heterogeneous data formats (EDI, Text, XML, SQL, etc...)
 - Define a knowledge base with RDF/XML, OWL, and use SPARQL in order to interrogate and integrate data in the context of a semantic Web.
 - b. **KPI:**

KPI	a2	e3	k2	k3
Covered	x	x	x	x
Assessed	x	x	x	x

7. Brief list of topics to be covered

Lecture	Description
1	The XML language: Evolution - Formatting – Rules

2	Namespaces XML
3-4-5	Document Validation: DTD
6-7-8	Document Validation: XML Schemas
9-10	XPATH
11-12	XSLT
13	XQuery
14-15- 16	SAX – DOM
17-18	XML support with databases and data integration
19-20	XML Pipelining
21-22- 23	RDF-OWL-SPARQL
24-28	Lab: XSLT- SAX-DOM-XML Pipeline