



Semantic Web technologies in Unit-net IEDI

✓ **Vadim Ermolayev**

<http://eva.zsu.zp.ua/>

Zaporozhye State University,
Ukraine

<http://www.zsu.edu.ua/>

**UnIT-Net: IT in University
Management Network**

<http://www.unit-net.org.ua/>

TEMPUS/TACIS MP-JEP-2010-2003

UkrPROG'04, Kiev, 02-03.06.2004

Outline:

- What is the Semantic Web – just to remind ...
- UnIT-Net: the motivation, the domain, the project
- The State of the Art: the advances and the pitfalls
- Semantic Web technologies in UnIT-Net Infrastructure for Electronic Data Interchange

The Semantic Web*

- **W3C Initiative**
- **Aim**: to provide a comprehensible framework for **identifying, representing** and **processing** the **SEMANTICS** of **Web resources**
- The ultimate **vision**:
 - Worldwide distributed **device for computation**
 - Inhabited with **artificial** service providing **agents**

* Ermolayev, V. et al.: Towards a framework for agent-enabled semantic web service composition. *Int. J. of Web Services Research*, 1(3), 2004, p. 63-87

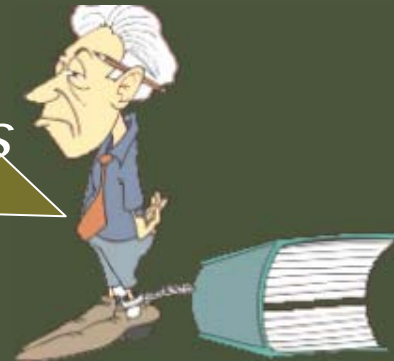
A Walkthrough Example

□ *Return the list of the 1-st year CS students who:*



■ *had received **maximal grade** in **Mathematics** at the entrance examinations*

■ *and have **failed** to pass the 1-st Term examination in **any basic course** in **Mathematics***



Why?

A Walkthrough Example

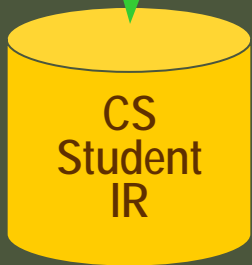
Return the list of the 1-st year CS students who:



had received **maximal grade** in **Mathematics** at the entrance examinations



and have **failed** to pass the 1-st Term examination in **any basic course in Mathematics**



Ontology

Mathematics:
— Math Analysis
— Linear Algebra
— Analytical Geometry
...

A Walkthrough Example

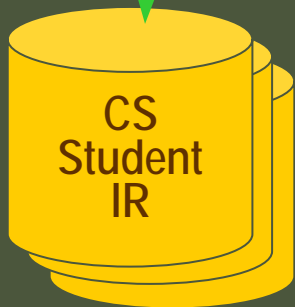
of different
Universities

Return the list of the 1-st year CS students who:

had received **maximal grade** in **Mathematics** at the entrance examinations



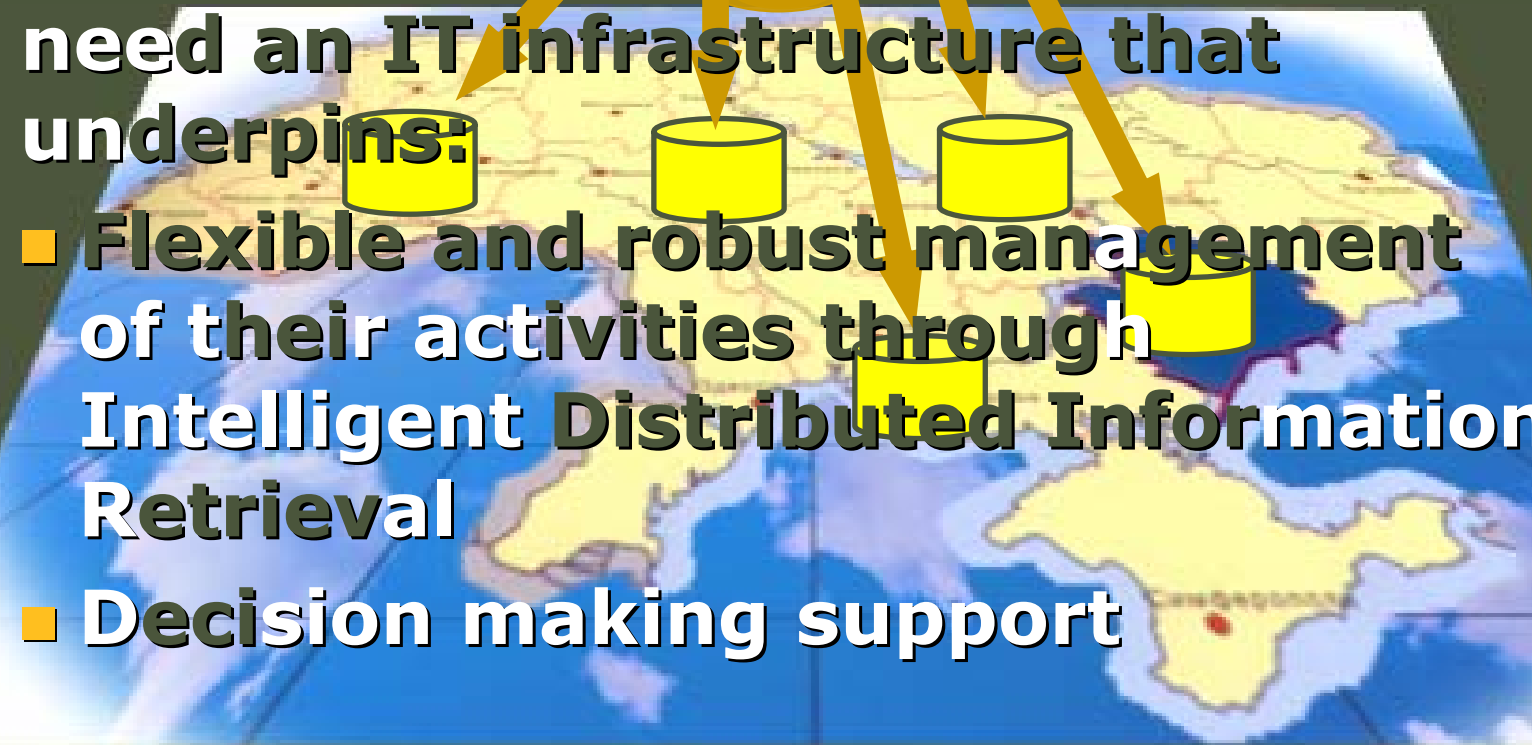
and have **failed** to pass the 1-st Term examination in **any basic course in Mathematics**



..and different
basic courses
in the 1-st
term

Mathematics:
- Math Analysis
- Linear Algebra
- Analytical Geometry
...

UnIT-Net IEDI: Motivation

- To achieve and sustain dynamic improvement service-oriented organizations, like Universities, need an IT infrastructure that underpins:
 - Flexible and robust management of their activities through Intelligent Distributed Information Retrieval
 - Decision making support
- 

UnIT-Net - TEMPUS/TACIS MP-JEP-2010-2003

□ **Objective(s):**

- Creation of the National “**Network of Excellence**”
- Dissemination of the **best practices** – IT in University Management
- Elaboration of the **Specifications** recommending the reasonable ways of using IT in University Management
- Design and implementation of the **Research Prototype** of the National **Infrastructure for Electronic Data Interchange** (motivation)

□ **Participants:**

- Kherson State University (project coordinator)
- Ministry of Education and Science of Ukraine
- Kharkiv national University
- Zaporozhye State University
- University of Nice – Sofia Antipolis, France
- Glasgow Caledonian University, UK

□ <http://www.unit-net.org.ua/>

The State of the Art ...

- ❑ **Not really a breakthrough** in the Domain!!!
- ❑ Lots of related work, e.g., **TSIMMIS**, **MOMIS**, **BUSTER**, **DOME**, **InfoSleuth**, **KRAFT**, **OBSERVER**, **Ontobroker**, **PICSEL**, **SIMS**, ... (**proves the importance**)
- ❑ **Novelties:**
 - **Ontologies** specified in W3C emerging de facto standard language (**OWL DL**)
 - **Ontology-driven** Semantic **Query Formulation, Transformation**, ... (ZSU **RACING** Project)
 - IR (RDB-structured) semantics is formalized by means of a **semi-structured Ontology Specification Language (OWL DL)**
 - Conceptually – one more layer (**Ontology**) of Semantic Specification on top of the **IR schema**
 - **Semantic Web Service** technology for **Uniform IR Wrapping**
- ❑ All these is in the **Mainstream of Semantic Web Activities**

Complications: Natural Distribution and Heterogeneity

- ❑ **Organizations** involved in the Educational framework **are rightfully independent**
- ❑ They own and **maintain** their **data** and **knowledge sources autonomously**
- ❑ Serious complications for their integration:
 - IR-s may be **opened or closed to external access**
 - IR-s may be **provided by** different hardware and software using **various notations and protocols**
 - IR-s may be **disparately structured**
 - IR-s may have **different data models** behind them
 - IR-s are **semantically heterogeneous**

The Principles of IEDI Architecture

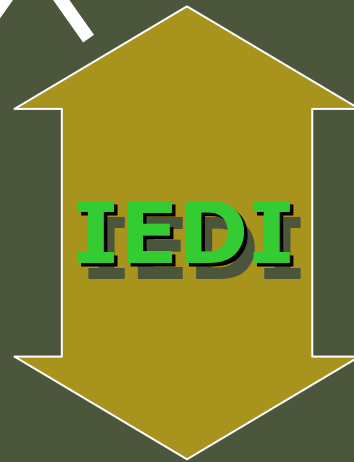
- **Mediator architecture** with the centralized mediator
- **Hybrid** approach to **knowledge representation**
 - **Centralized Mediator Domain Ontology (MDO)**
 - **De-centralized Information Resource Ontologies (IRO)**
- Use of **IR Registration** to allow the resource become available for querying
- Does **not** provide **full automation** for ontologies' mapping and alignment
- **Rewriting** technique with **mappings** and **late binding** to produce, process, and perform queries

IEDI Architecture in a Nutshell

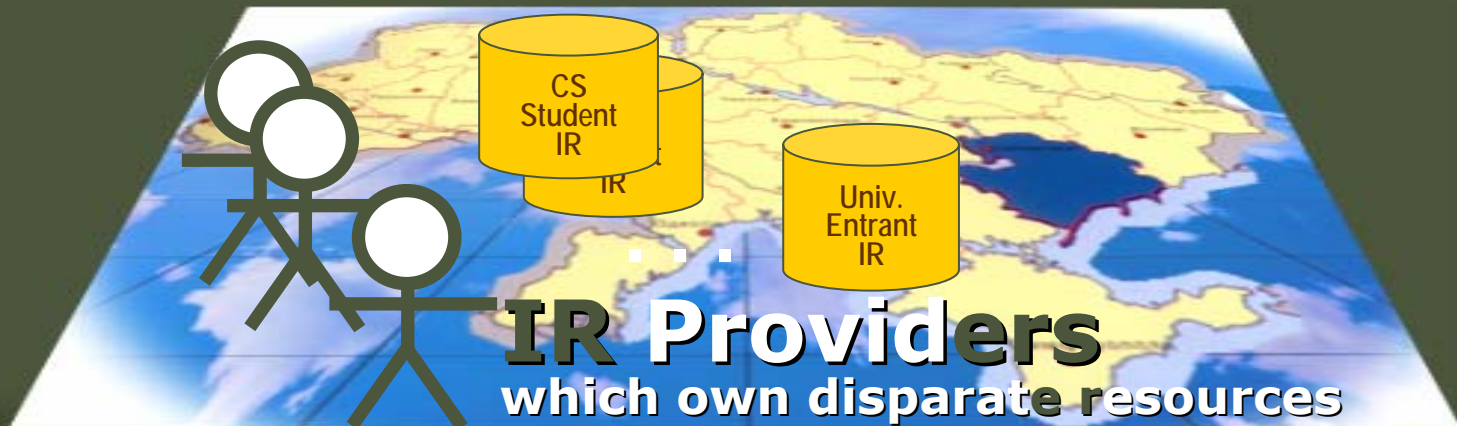


a User

having an arbitrary query



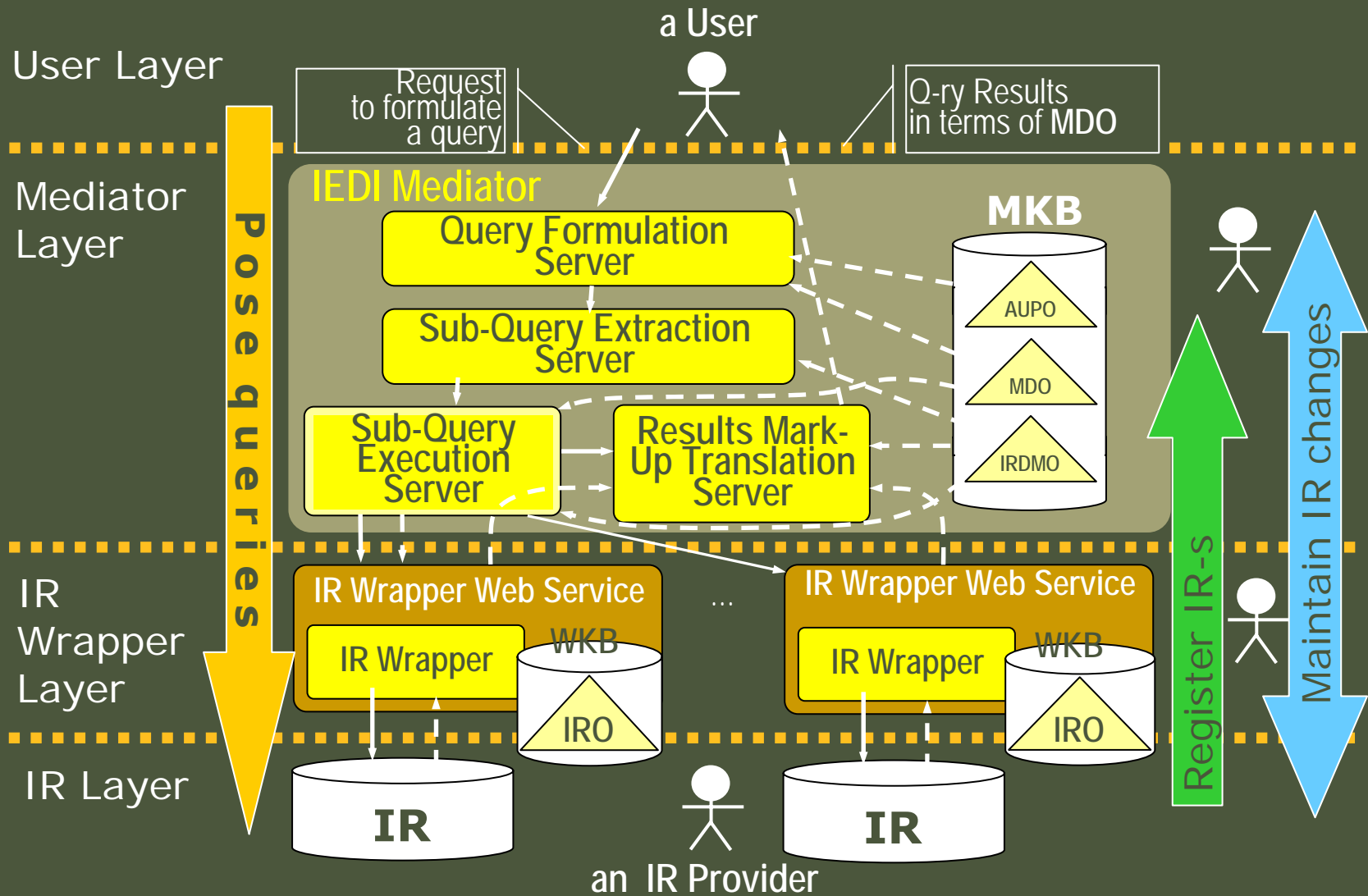
???



IR Providers

which own disparate resources

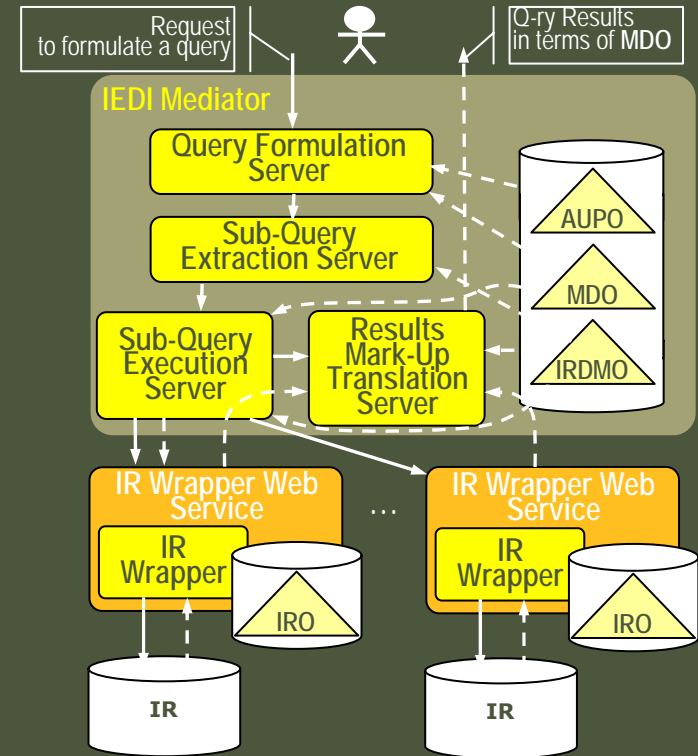
IEDI Architecture in a Nutshell



The Tasks for UNIT-NET IEDI

- ❑ **To Query Distributed Semantically Heterogeneous Information Resources**
- ❑ **To Register Information Resources**
- ❑ **To Maintain Coherent Semantic Descriptions**

- ❑ **What IEDI is NOT Supposed to Do:**
 - IR updates
 - Results Fusion



No silver bullets - ☹️

IEDI: User Categories and Roles

□ An **Authorized USER (AU)**:

- Poses queries in the terms of **University Management Domain** (a **TOOL**, a **LANGUAGE**)

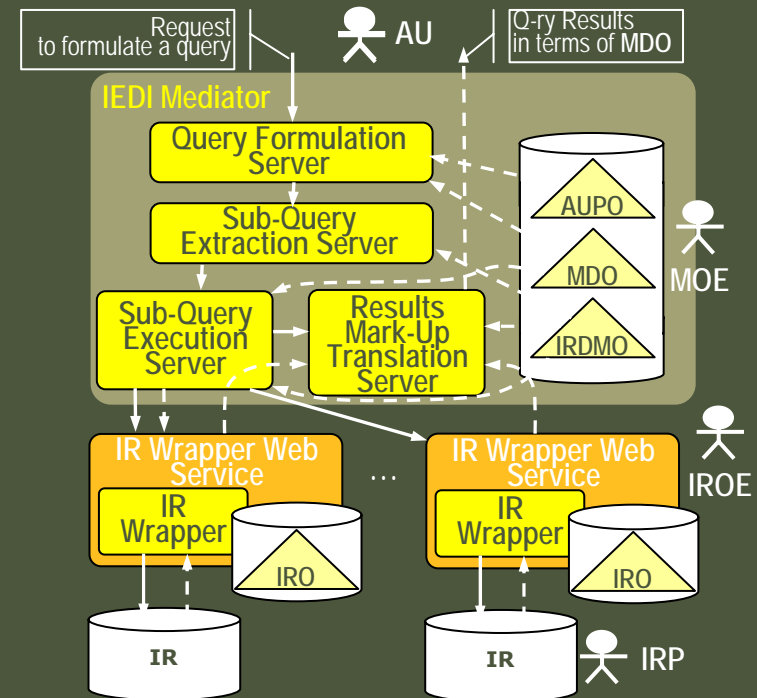
□ A **MEDIATOR ONTOLOGIES ENGINEER (MOE)**:

- Maintains **Domain Ontology KB** (a **TOOL**)
- Interacts with **RESOURCE ONTOLOGY ENGINEERS** for:

- Registering their Resources (semi-automatic, a **TOOL**)
- Aligning **Domain** and **Resource Ontologies** (Semi-automatic, a **TOOL**)

□ An **IR ONTOLOGY ENGINEER (IROE)**: ...

□ An **IR PROVIDER (IRP)**: ...



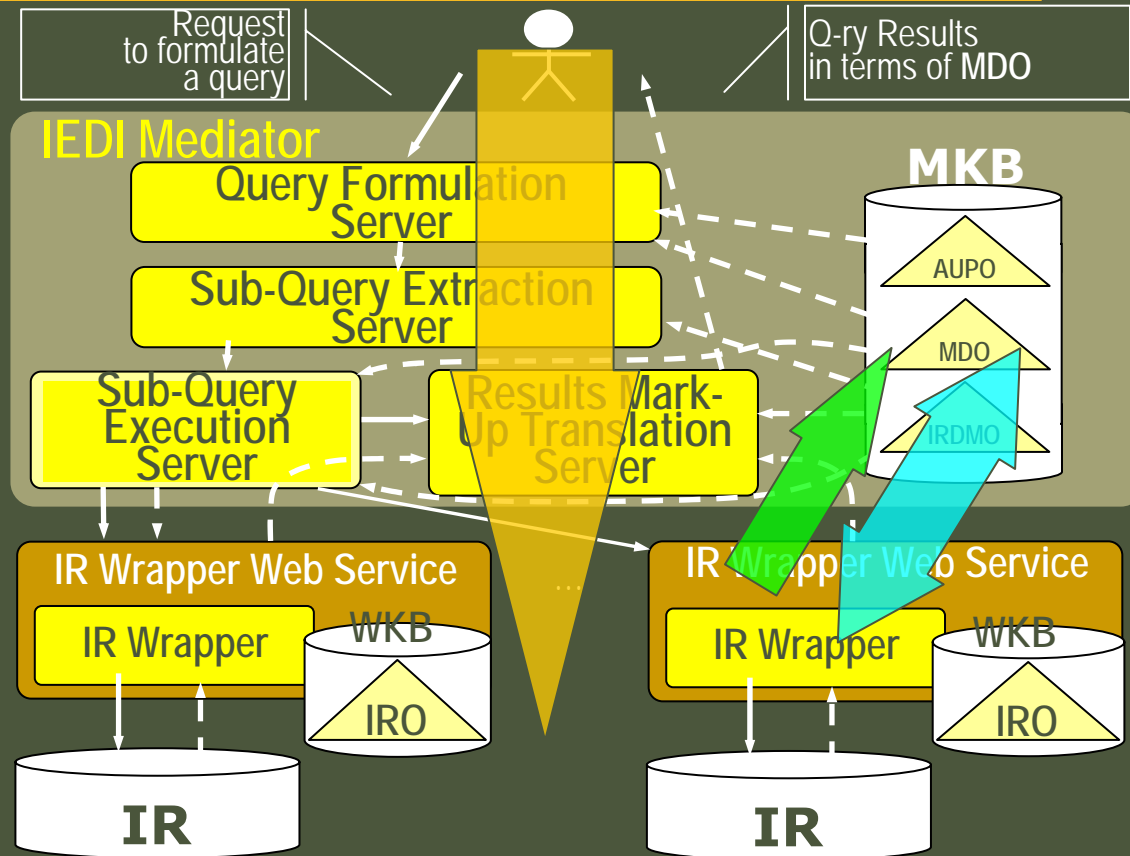
IEDI: Functionalities

- Query (automatic)
Distributed
Semantically
Heterogeneous
Information
Resources

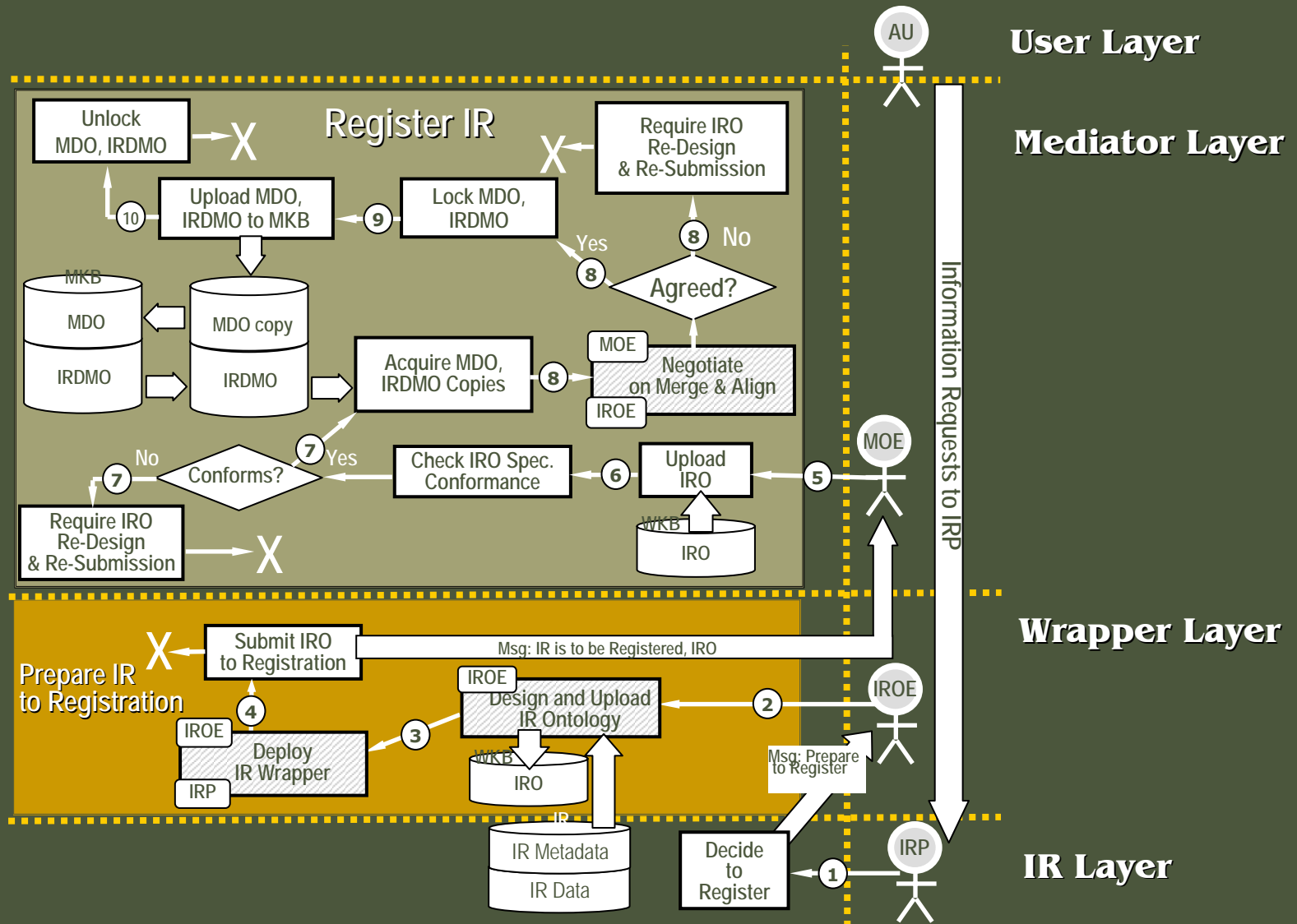
- Register (semi-auto)
Information
Resources
(ontology merge)

- Maintain (semi-auto)
Coherent
Semantic
Descriptions (ontology alignment)

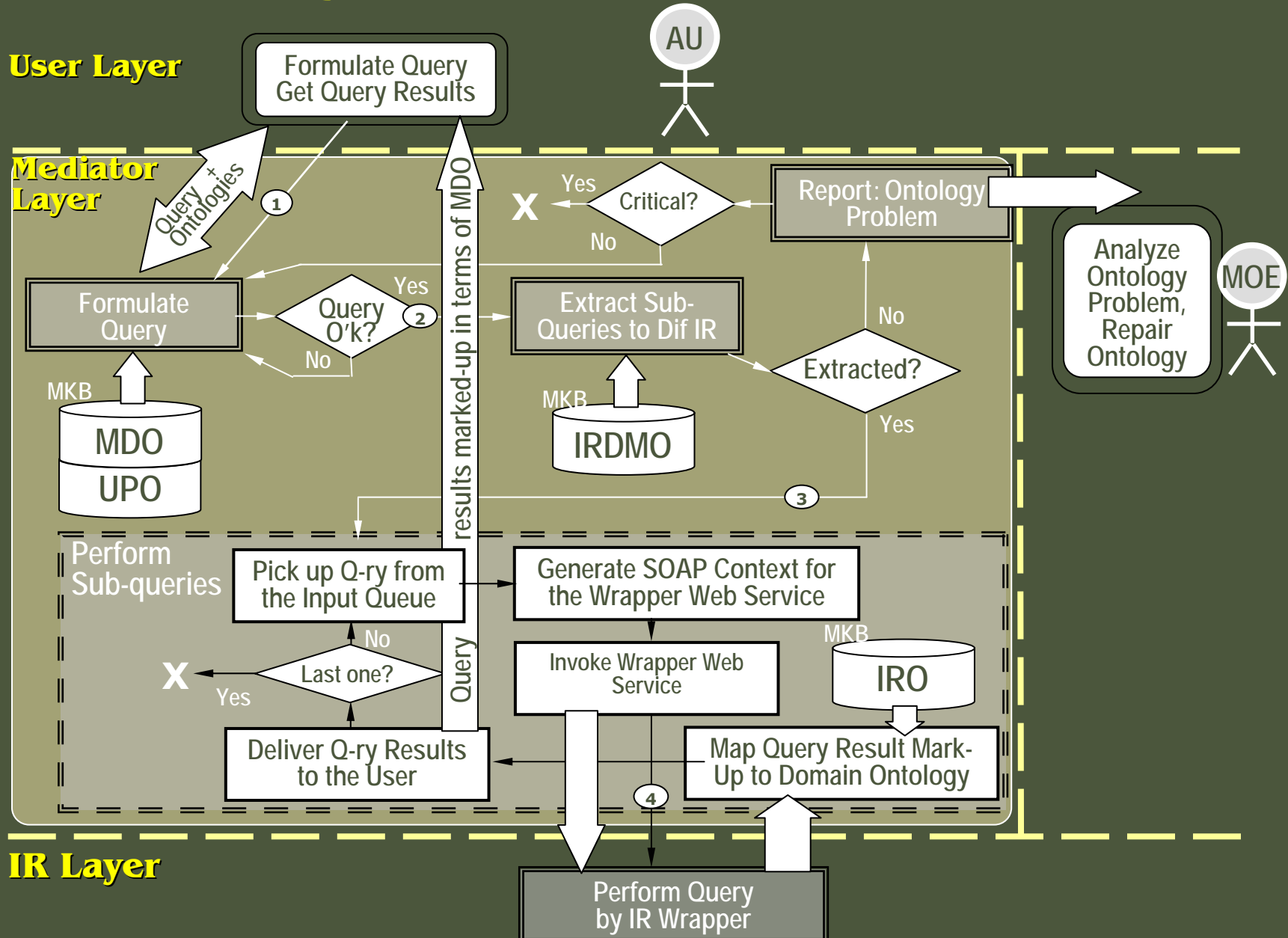
- !!! Semi-automatic, authorized, secure ...



Register IR

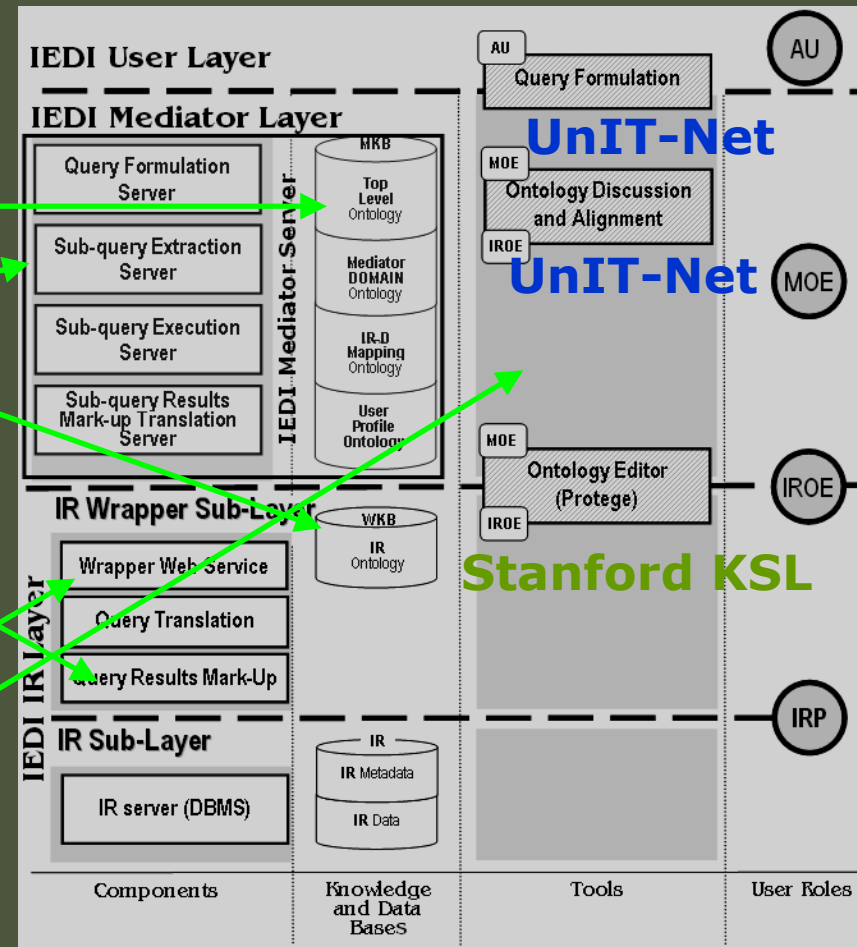


Perform Queries to IR Collection



Semantic Web components in IEDI

- ❑ Ontology Language: **OWL** (W3C recommendation)
- ❑ Ontologies at Mediator and IR layers
- ❑ Mediator Query Language: **RDQL** (W3C recommendation)
- ❑ Mark-up Language: **XML** (W3C standard recommendation)
- ❑ Ontology processing tools
- ❑ **Semantically reinforced Web Services** for IR wrapping

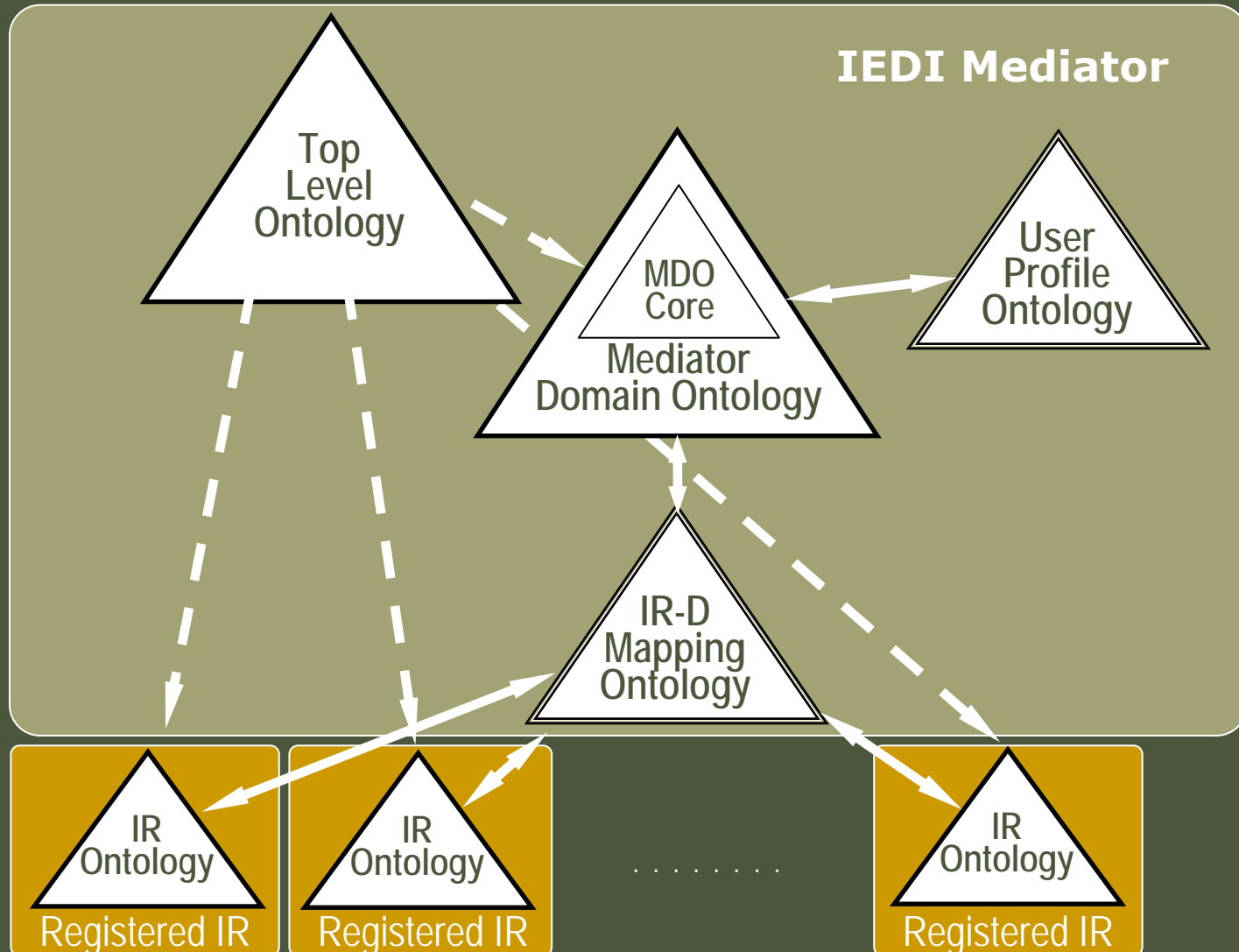


IEDI: Ontologies

- **Ontologies** are developed to provide a machine-processable semantics of IR-s that can be communicated between different software and humans
- **An ontology is a formal, explicit specification of a shared conceptualization***
 - **Conceptualization** - a simplified abstract model of some object or phenomenon in the world which identifies the relevant concepts of that object or phenomenon
 - **Formal** ...
 - **Explicit** ...
 - **Shared** ...

*Gruber, T. R.: A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition*, 5:199—220, 1993.

IEDI: Hierarchy of Ontologies



IEDI: Usage of Ontologies

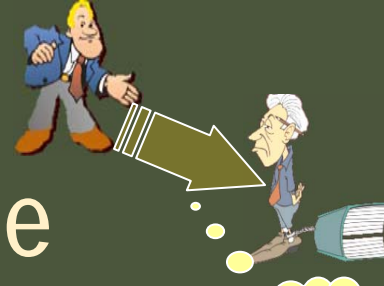
Mediator Knowledge Base (MKB)

WKB

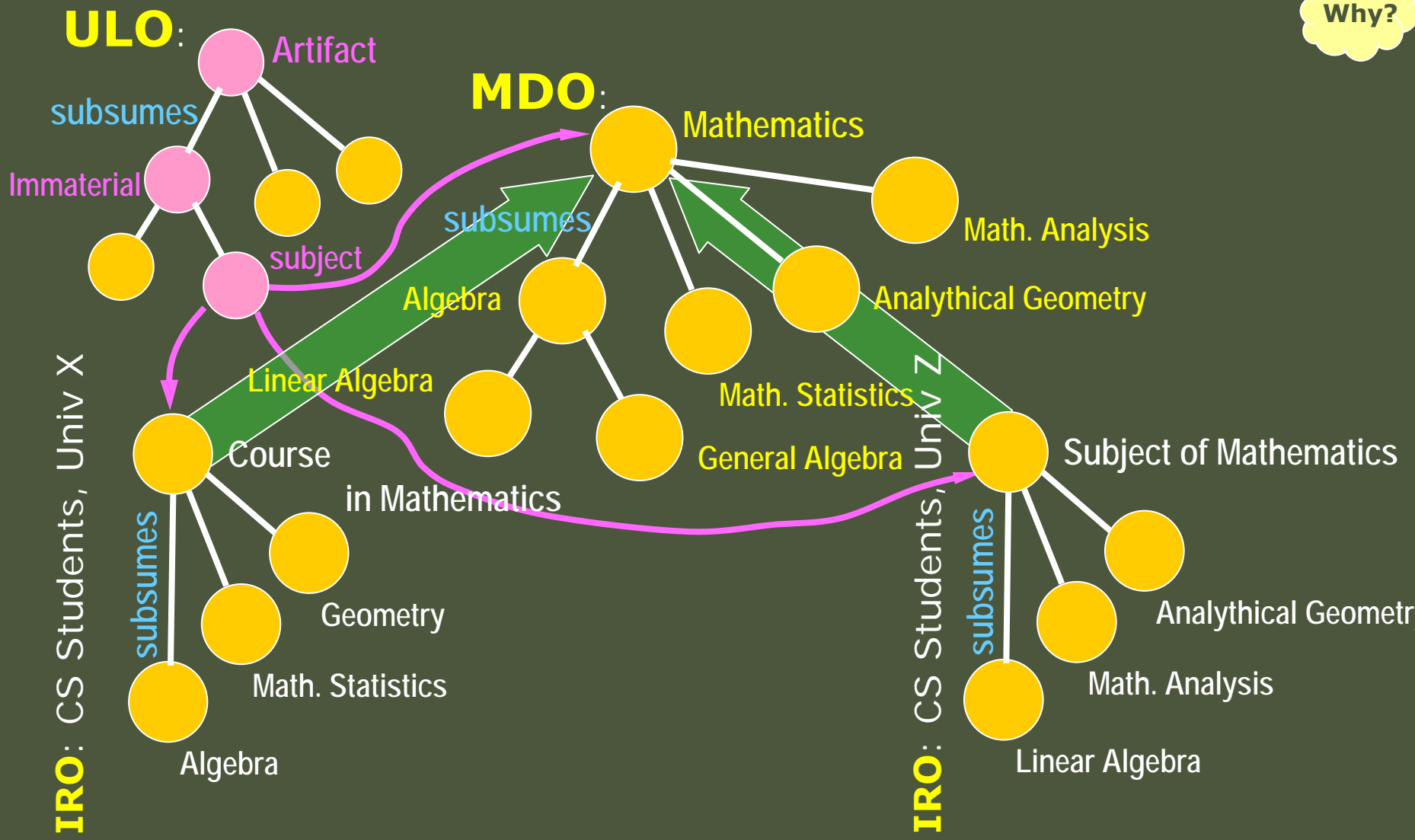
Processes \ Ontologies	Mediator Knowledge Base (MKB)					WKB
	ULO	MDO Core	MDO	IRDMO	UPRO	IRO
Query distributed autonomous semantically heterogeneous information resources	--	R	R	R	R/U	R
Register new information resource	R	R	R/U	R/U	--	R
Maintain coherent semantic descriptions	R	R/U	R/U	R/U	R/U	R/U

- **R** – usage for reference purposes only
- **R/U** – used as a reference and is updated
- **--** – not used

Ontologies: Walkthrough Example

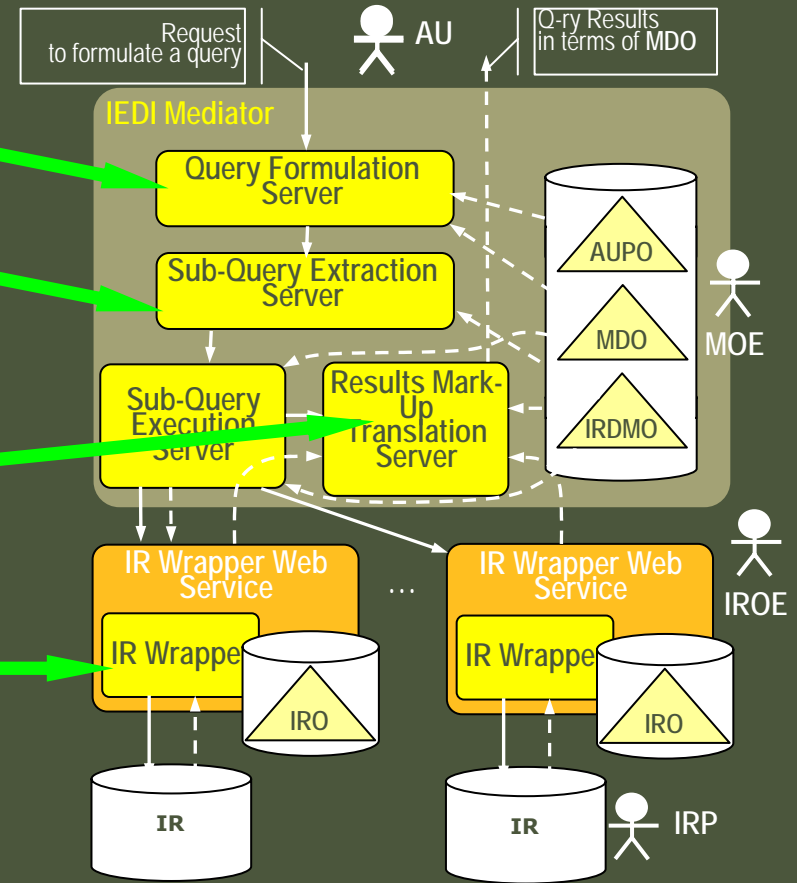


Why?



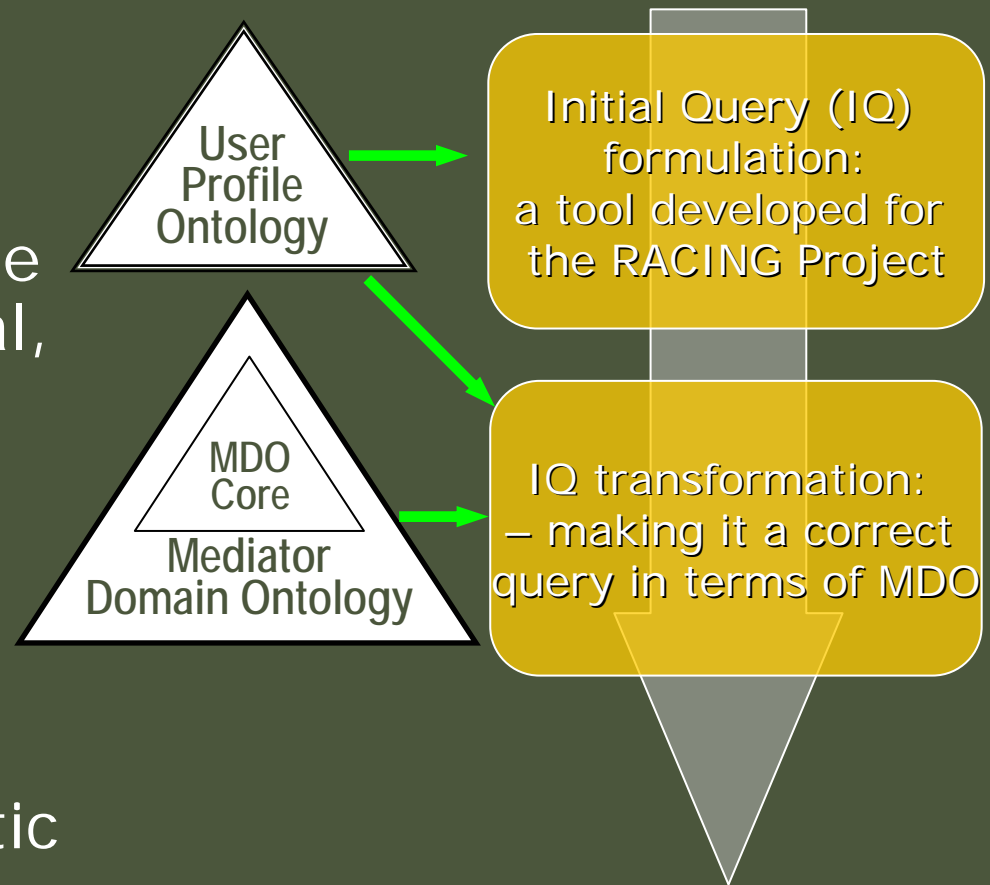
Semantically Reinforced Components

- ❑ Query Formulation Server
- ❑ Sub-query Extraction Server
- ❑ Results Mark-up Translation Server
- ❑ IR Wrapper



IEDI Mediator: Query Formulation*

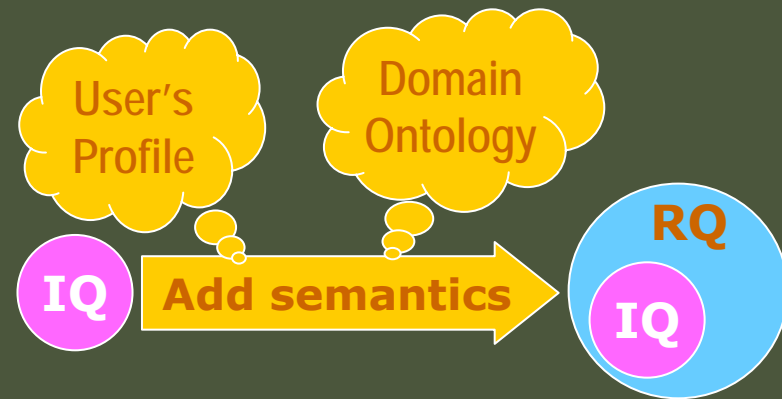
- **Query formulation**
(in the terms which are convenient and understandable for the specific AU) – manual, by the Tool
- **Query Transformation:**
reformulating the query in the terms of MDO (preserving the recall) – automatic



*Ermolayev, V. et al.: Capturing Semantics from Search Phrases: Incremental User Personification and Ontology-Driven Query Transformation. In: Proc. of the 2-nd Int. Conf. on Information Systems Technology and its Applications (ISTA'2003), Kharkiv, Ukraine, June 19-21, 2003, pp. 9-20,

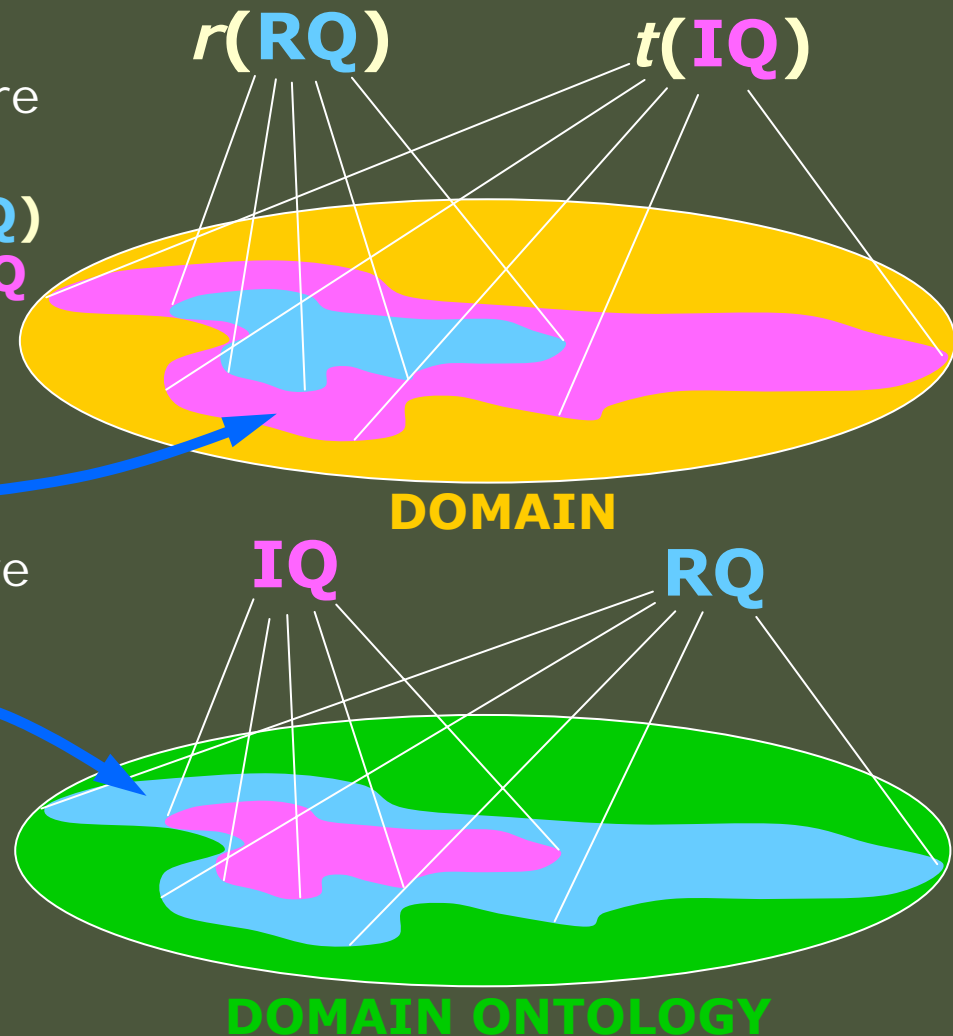
IQ Transformation

- The **GOAL**: transform the **Initial Query (IQ)** to the **Resulting Query (RQ)** in the terms of the **MEDIATOR DOMAIN ONTOLOGY**
- The **procedure**:
 - Form the **Query Plan (QP)** by parsing the **IQ**
 - Use the **User's Profile** to map the **key words** of **QP** to the **concepts** of the **DOMAIN ONTOLOGY**
 - Use **semantic relationships** between the **concepts** of the **DOMAIN ONTOLOGY** to add more semantics to **RQ**
- The **Basic Principle**: – **IQ** preservation



IQ Preservation

- **IQ preservation principle:**
 - **Strategic goal** is to gain more **Recall** and more **Precision**
 - I.e., relevant **RQ** results $r(\mathbf{RQ})$ should be the sub-set of all **IQ** results $t(\mathbf{IQ})$ and, **ideally**, the difference $t(\mathbf{IQ}) \setminus r(\mathbf{RQ})$ should contain only **irrelevant results**
 - Consequently, **RQ** should have **the same or the broader meaning** than **IQ**
- Transformation mappings are produced in the way providing that the **recall** of the **RQ** is **at least the same** than the **recall** of the **IQ**



IEDI: Implemented Components

- **User Profile Ontology Editor**
- **Tool for IQ Plan editing and approval**
 - Contribution of the RACING project
<http://racing.zsu.zp.ua>

	Mero	Relev	Remove
ATION	is-a	0.7	<input type="checkbox"/>
			<input type="checkbox"/>

- **Parts of IEDI Mediator Query Formulation Tool**

#	I agree	Source	Keyword	Mero	Concept	Relev...	Edit
1	<input checked="" type="checkbox"/>	Profile	Database Management System	is-a	SOFTWARE	0,7	EDIT
2	<input checked="" type="checkbox"/>	Profile	Database Management System	component-of	DATABASE MANAGEMENT	0,9	EDIT
3	<input checked="" type="checkbox"/>	Profile	Indexing	is-a	DATA STORAGE REPRESENTATION	0,7	EDIT

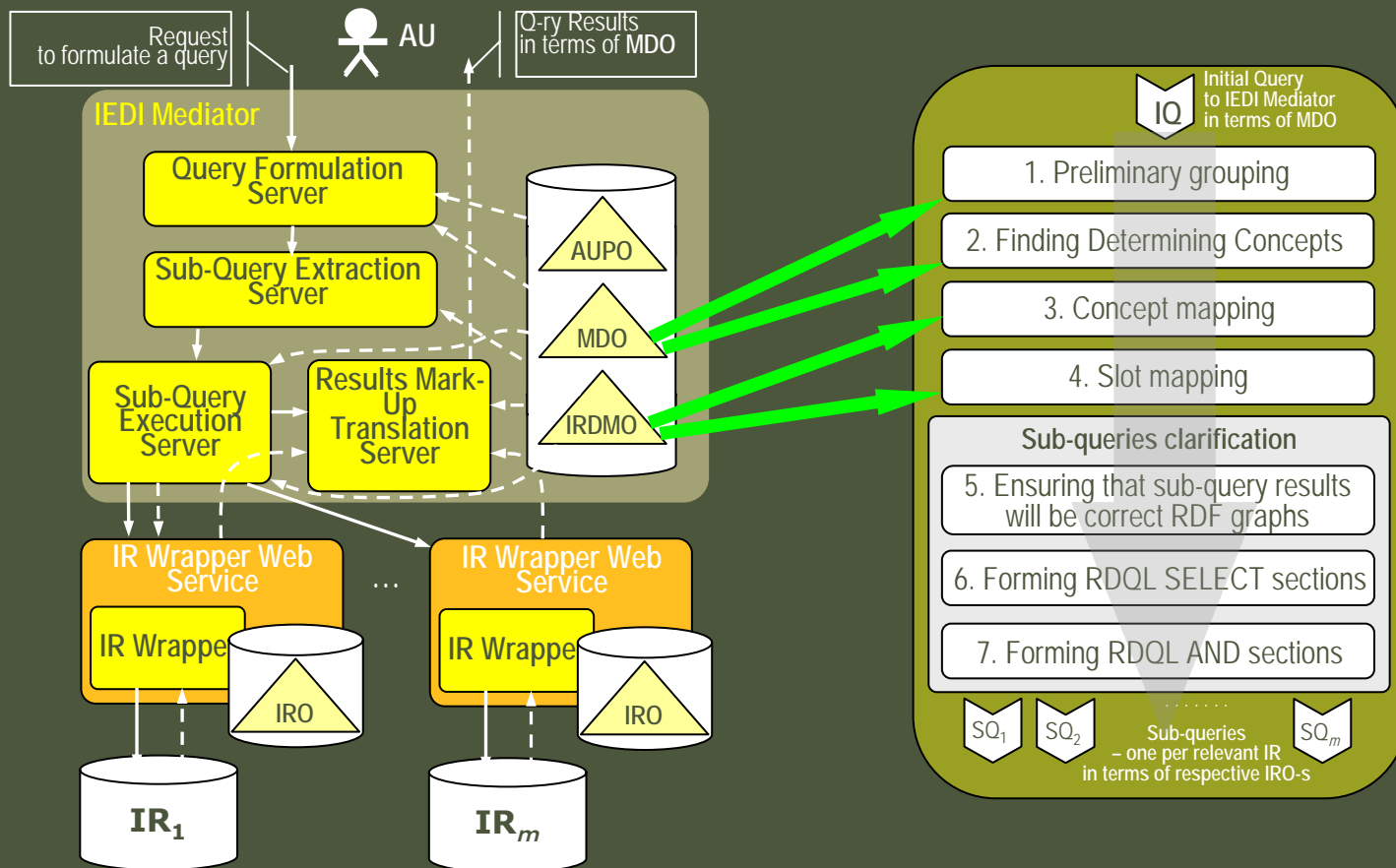
Concept detail:

Concept:
Data \ Data Storage Representations
URI: http://racing/xml_classification.dami#ACMTopic__Data__Data_Storage_Representations

Parent:
Data

Children:
Data \ Data Storage Representations \ Hash Table Representations
Data \ Data Storage Representations \ Linked Representations
Data \ Data Storage Representations \ Object Representation

IEDI Mediator: Sub-Query Extraction*



*Ermolayev. V. et al.: Ontology-Driven Sub-Query Extraction for Distributed Autonomous Information Resources in UniT-Net IEDI. Proc. 3-d Intl. Conference on Information Systems Technology and its Applications (ISTA'2004), Salt Lake City, Utah, USA, July 14-16, 2004.

Walkthrough Example: IQ (RDQL)

Retrieve the list of the 1-st year students who have received maximum grade (5) in Mathematics at the University entrance examinations and have failed to pass the 1-st semester examination in any basic course in Mathematics (got unsatisfactory grade - 2).

```
SELECT ?firstName, ?secondName, ?lastName, ?specialityName, ?sessionExTitle
WHERE
  (?x, stud:first_name, ?firstName), (?x, stud:second_name, ?secondName),
  (?x, stud:last_name, ?lastName), (?x, stud:exams_passes, ?y),
  (?x, stud:exams_passes, ?z), (?x, stud:on_spec, ?a),
  (?y, stud:exam_title, ?entrantExTitle), (?y, stud:exam_type, ?examType1),
  (?y, stud:entrant_grade, ?entrantGrade), (?y, stud:examOnDiscipline, ?r1),
  (?z, stud:exam_title, ?sessionExTitle), (?z, stud:exam_type, ?examType2),
  (?z, stud:session_grade, ?sessionGrade),
  (?z, stud:semesterNum, ?semesterNum),
  (?z, stud:examOnDiscipline, ?r2),
  (?a, stud:specialityName, ?specialityName)
  (?r1, stud:disciplineName, ?entrDiscName), (?r1, stud:includes, ?i1),
  (?r2, stud:disciplineName, ?sessionDiscName), (?r2, stud:includes, ?i2),
  (?i1, stud:disciplineName, ?discName1),
  (?i2, stud:disciplineName, ?discName2)
AND (?examType1 eq "Exam"), (?examType2 eq "Exam")
AND (?entrDiscName eq "Mathematics"), (?sessionDiscName eq "Mathematics")
AND ((?entrantExTitle eq ?discName1) || (?sessionExTitle eq ?discName2))
AND ((?sessionExTitle eq "Linear Algebra") ||
      (?sessionExTitle eq "Mathematical Analysis"))
AND (?entrantGrade eq "5")
AND (?sessionGrade eq "2")
AND (?semesterNum eq "1")
USING stud FOR <MDO-URL#>
```

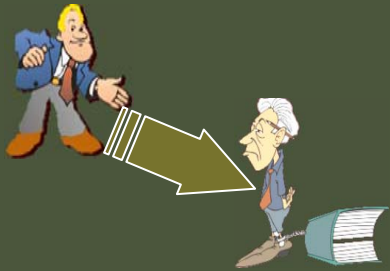
Walkthrough Example: Results (RDQL)

```
SELECT ?firstName, ?secondName, ?lastName, ?specialityName, ?sessionExTitle
WHERE
  (?x, stud:first_name, ?firstName), (?x, stud:second_name, ?secondName),
  (?x, stud:last_name, ?lastName), (?x, stud:exams_passes, ?y),
  (?x, stud:exams_passes, ?z), (?x, stud:on_spec, ?a),
  (?y, stud:exam_title, ?entrantExTitle), (?y, stud:exam_type, ?examType1),
  (?y, stud:entrant_grade, ?entrantGrade), (?y, stud:examOnDiscipline, ?r1),
  (?z, stud:exam_title, ?sessionExTitle), (?z, stud:exam_type, ?examType2),
  (?z, stud:session_grade, ?sessionGrade), (?z, stud:exam_type, ?examType2),
  (?z, stud:semesterNum, ?semesterNum),
  (?z, stud:examOnDiscipline, ?r2),
  (?a, stud:specialityName, ?specialityName)
  (?r1, stud:disciplineName, ?entrDiscName), (?r1, stud:includes, ?i1),
  (?r2, stud:disciplineName, ?sessionDiscName), (?r2, stud:includes, ?i2),
  (?i1, stud:disciplineName, ?discName1),
  (?i2, stud:disciplineName, ?discName2),
  AND (?examType1 eq "Exam"), (?examType2 eq "Exam")
  AND (?entrDiscName eq "Mathematics"), (?sessionDiscName eq "Mathematics")
  AND ((?entrantExTitle eq ?discName1) || (?sessionExTitle eq ?discName2))
  AND ((?sessionExTitle eq "Linear Algebra") ||
        (?sessionExTitle eq "Mathematical Analysis"))
  AND (?entrantGrade eq "5")
  AND (?sessionGrade eq "2")
  AND (?semesterNum eq "1")
USING stud FOR <MDO-URL#>
```

```
SELECT ?firstName, ?secondName, ?lastName,
       ?specialityName
WHERE
  (?x, abo:aboName, ?firstName),
  (?x, abo:secondName, ?secondName),
  (?x, abo:surname, ?lastName),
  (?x, abo:passes, ?y),
  (?x, abo:AboSpec, ?a),
  (?y, abo:EntrantExamName, ?entrantExTitle),
  (?y, abo:examType, ?examType1),
  (?y, abo:grade, ?entrantGrade),
  (?y, abo:examOnDiscipline, ?r1),
  (?a, abo:specialityName, ?specialityName)
  (?r1, abo:disciplineName, ?entrDiscName),
  (?r1, abo:includes, ?i1),
  (?i1, abo:disciplineName, ?discName1),
  AND (?examType1 eq "Exam")
  AND (?entrDiscName eq "Mathematics")
  AND ((?entrantExTitle eq ?discName1)
  AND (?entrantGrade eq "5")
USING abo FOR <IRO Entrant-URL#>
```

```
SELECT ?firstName, ?secondName, ?lastName,
       ?specialityName, ?sessionExTitle
WHERE
  (?x, stud:name, ?firstName),
  (?x, stud:secondName, ?secondName),
  (?x, stud:surName, ?lastName),
  (?x, stud:examPasses, ?z),
  (?x, stud:onSpec, ?a),
  (?z, stud:examName, ?sessionExTitle),
  (?z, stud:examType, ?examType2),
  (?z, stud:grade, ?sessionGrade),
  (?z, stud:semesterNum, ?semesterNum),
  (?a, stud:specialityName, ?specialityName)
  AND (?examType2 eq "Exam")
  AND ((?sessionExTitle eq "Linear Algebra") ||
        (?sessionExTitle eq "Mathematical Analysis"))
  AND (?sessionGrade eq "2")
  AND (?semesterNum eq "1")
USING stud FOR <IRO-Student URL#>
```

Walkthrough Example: Results (English)



Retrieve the list of the 1-st year students who:

- have received maximum grade (5) in Mathematics at the University entrance examinations
- and have failed to pass the 1-st semester examination in any basic course in Mathematics (got unsatisfactory grade - 2).



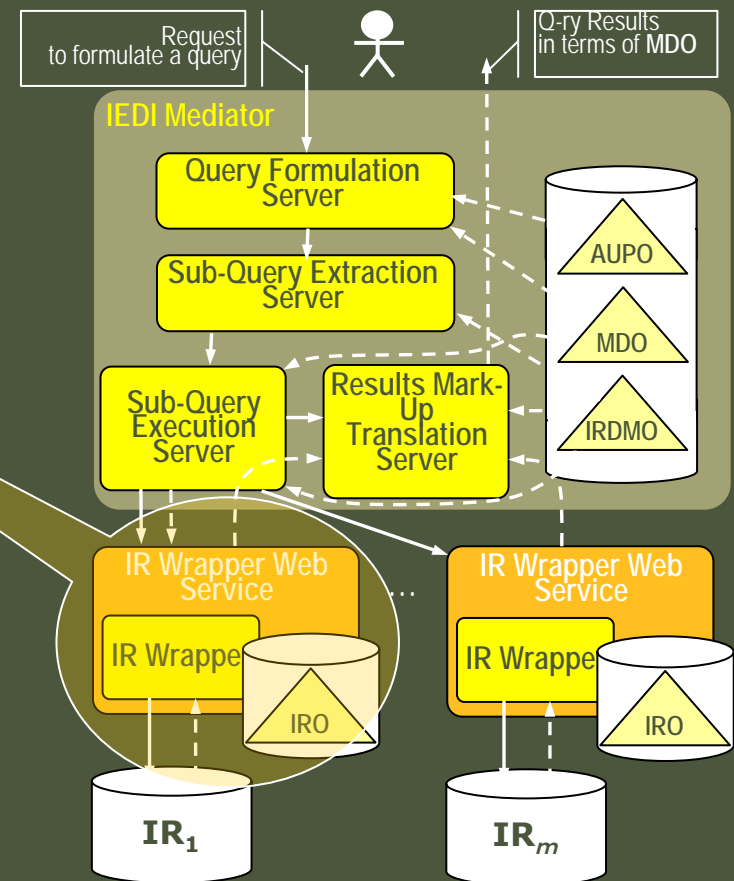
Retrieve the list of the 1-st year students who have received maximum grade (5) in Mathematics at the University entrance examinations



Retrieve the list of the 1-st year students who have failed to pass the 1-st semester examination in any basic course in Mathematics (got unsatisfactory grade - 2).

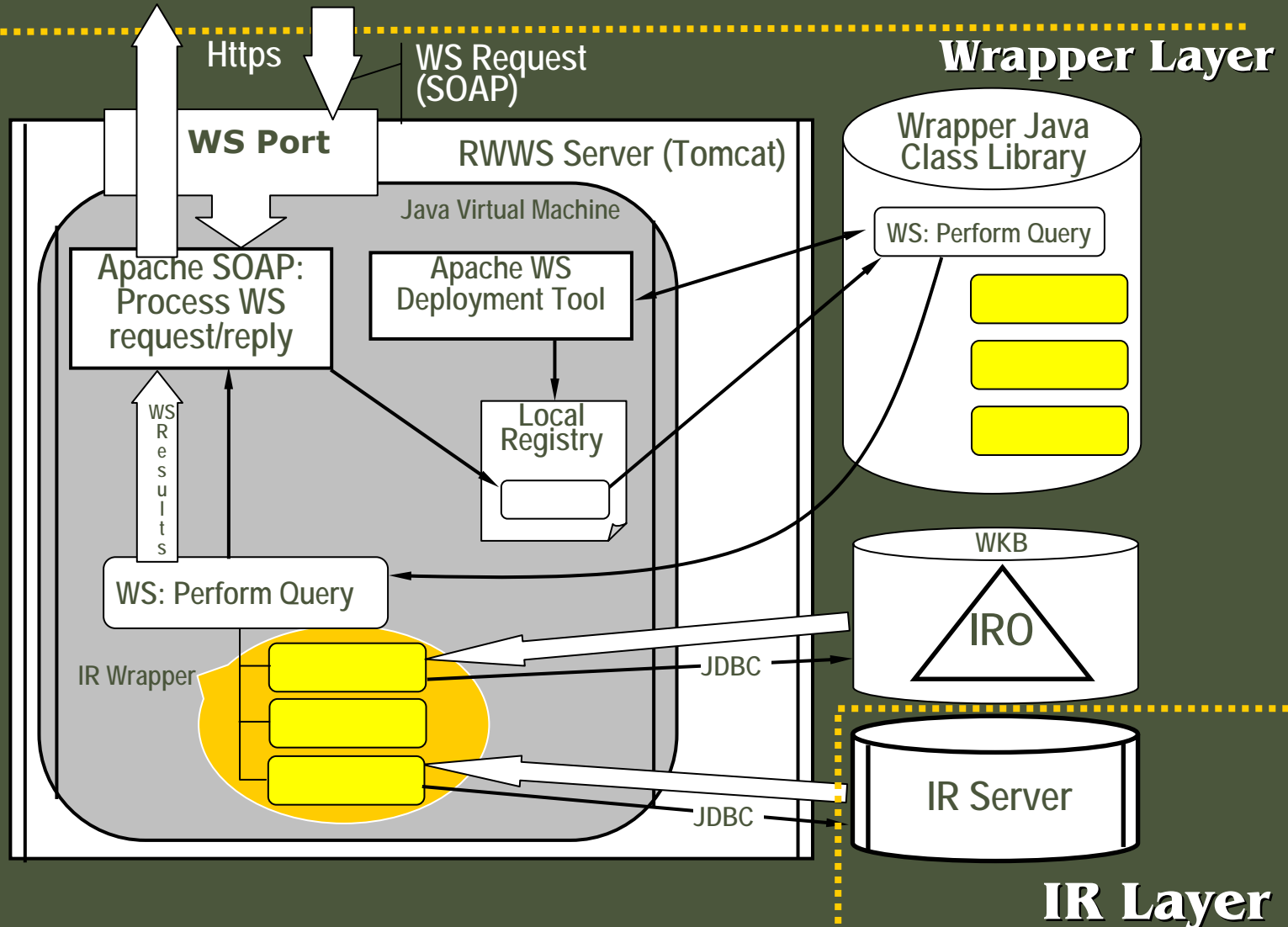
IEDI: IR Wrapping*

- **IR Wrapper** design is based on **Web Service Technology**
- IR Wrapper Web Service is **Semantically Reinforced** by:
 - **Generic IR Wrapper**
 - Specific **IR Wrapper binding**
- which use **IRO** for their operations

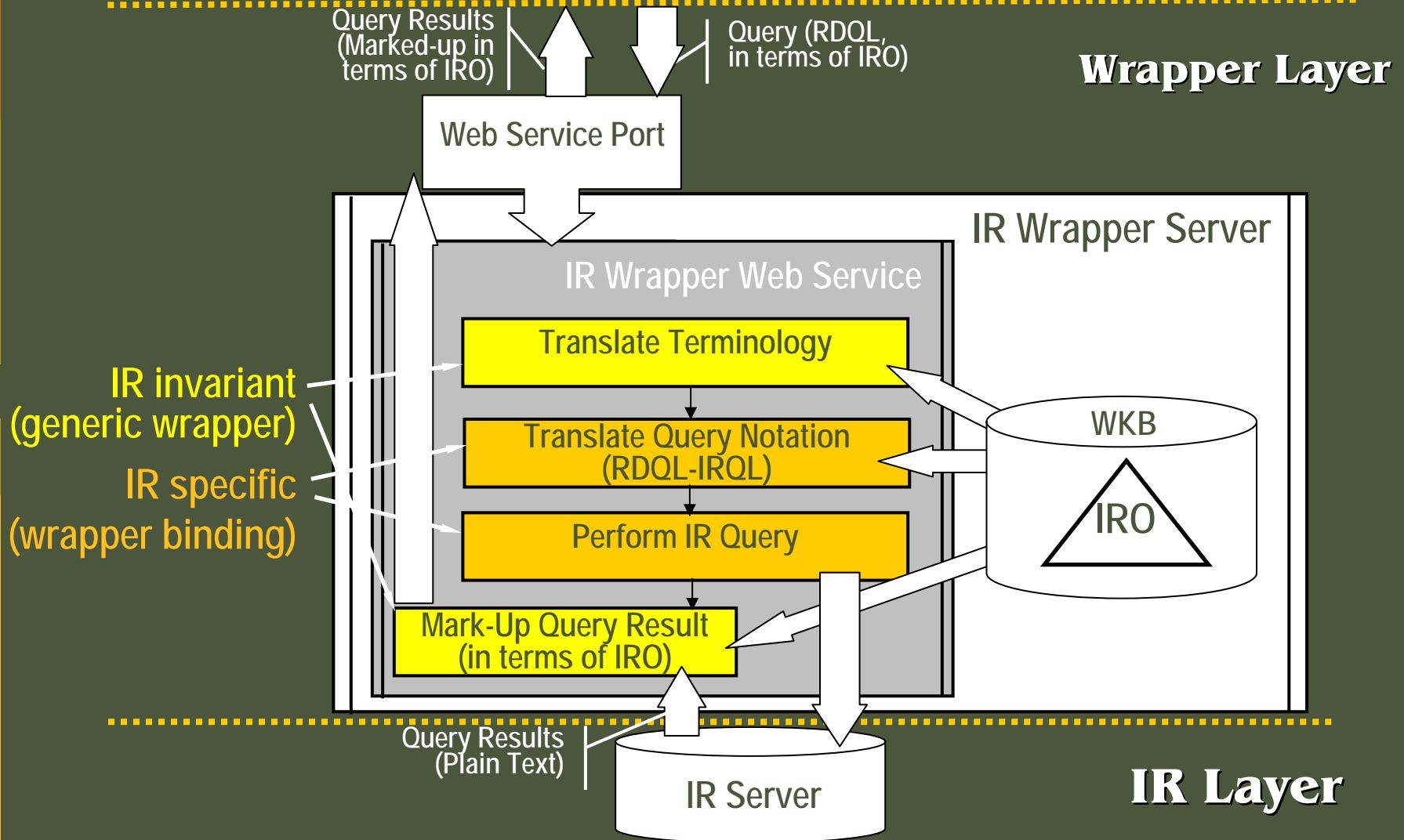


*Ermolayev, V. et al.: Semantically Reinforced Web Services for Wrapping Autonomous Information Resources. Submitted to: 2-nd European Conference on Web Services (ECOWS'04), Erfurt, Germany, September 27-30, 2004.

IEDI: IR Wrapper Web Service



IEDI Generic Wrapper and Wrapper Bindings



IEDI: Implemented Components

- ❑ Generic IR Wrapping Web Service
- ❑ Wrapper Testing Suite →
- ❑ IR Wrapper for ZSU University Entrant IR

The screenshot shows a Microsoft Internet Explorer window titled "Wrapper Web Service test Client - Microsoft Internet Explorer". The address bar shows the URL "http://www.zsu.zp.ua:88/cgi-bin/wrapperTest.htm". The main content area is divided into several sections:

- RQDL:** A text area containing an RDQL query:

```
#SQLONLY
SELECT ?aboName, ?secondName, ?surname, ?een, ?e1, ?sen, ?e2
WHERE (?x, stud:aboName, ?aboName), (?x, stud:secondName, ?secondName), (?x,
stud:surname, ?surname),
(?x, stud:passes, ?q), (?q, stud:EntrantExamName, ?een), (?q, stud:grade, ?e1),
(?x, stud:hasMarks, ?w), (?w, stud:SertExamName, ?sen), (?w, stud:grade, ?e2)
AND (?e1 = 9.0)
AND (?e2 = 9.0)
USING stud FOR http://owl.protege.stanford.edu#
```
- Perform...:** A button to execute the query.
- Wrapper:** A section header.
- SOAP Request:** A text area showing the XML request:

```
<?xml version="1.0"?><SOAP-ENV:Envelope xmlns:SOAP-
ENV="http://schemas.xmlsoap.org/soap/envelope/"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:SOAP-
ENC="http://schemas.xmlsoap.org/soap/encoding/"><SOAP-ENV:Body><NS1:RDQLToSQL
xmlns:NS1="urn:_wrapperIntf-Iwrapper" SOAP-
ENV:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"><sSource>#SQLONLY
SELECT ?aboName, ?secondName, ?surname, ?een, ?e1, ?sen, ?e2
WHERE (?x, stud:aboName, ?aboName), (?x, stud:secondName, ?secondName), (?x,
stud:surname, ?surname),
(?x, stud:passes, ?q), (?q, stud:EntrantExamName, ?een), (?q, stud:grade, ?e1),
(?x, stud:hasMarks, ?w), (?w, stud:SertExamName, ?sen), (?w, stud:grade, ?e2)
AND (?e1 = 9.0)
AND (?e2 = 9.0)
USING stud FOR http://owl.protege.stanford.edu#
</sSource></NS1:RDQLToSQL></SOAP-ENV:Body></SOAP-ENV:Envelope>
```
- SOAP Response:** A text area showing the XML response:

```
<?xml version="1.0"?>
<SOAP-ENV:Envelope xmlns:SOAP-ENV="http://schemas.xmlsoap.org/soap/envelope/"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:SOAP-
ENC="http://schemas.xmlsoap.org/soap/encoding/"><SOAP-ENV:Body SOAP-
ENC:encodingStyle="http://schemas.xmlsoap.org/soap/envelope/"><NS1:RDQLToSQLResponse
xmlns:NS1="urn:_wrapperIntf-Iwrapper"><return xsi:type="xsd:string">#SQLONLY
SELECT Profiles.name, Profiles.second_name, Profiles.surname,
List_OExam_dic.list_oexam, EntrantExam.grade, List_Osert_dic.list_osert,
EntrantExam.grade
FROM Profiles, List_OExam_dic, EntrantExam, List_Osert_dic
WHERE (EntrantExam.grade = 9.0)
AND (EntrantExam.grade = 9.0)
AND EntrantExam.oexam=List_Oexam_dic.code</return></NS1:RDQLToSQLResponse></SOAP-
ENV:Body></SOAP-ENV:Envelope>
```

Unit-net IEDI: to Round up ...

- That is what we have done in the project ... so far
- Semantic web technologies are used (and developed) for:
 - representing different aspects of knowledge
 - domain, resource, user profile, mapping, high-level
 - formulating, transforming, splitting down the queries to sub-queries
 - IR wrapping
 - Query results mark-up

That's it ...

**Shall be happy
to know the answers**

Mentioned papers and these slides are available from:

http://eva.zsu.zp.ua/eva_personal/evapubs.htm