Demonstration and Report of training modules

Project Number: FP6-2005-IST-026996

Deliverable id: D 14.3

Deliverable name: Demonstration and Report of training modules

Date: 31 January, 2008

COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	ACGT
Project Full Name:	Advancing Clinico-Genomic Clinical Trials on Cancer: Open Grid Services for improving Medical Knowledge Discovery
Document id:	D 14.3
Document name:	Demonstration and Report of training modules
Document type (PU, INT, RE)	RE
Version:	2.0
Date:	31.01.2009
Editor: Organisation: Address:	Radu Gramatovici SIVECO 8-10 Bd. Maresal Averescu, Bucharest, Romania

Document type PU = public, INT = internal, RE = restricted

ABSTRACT:

The present deliverable is a report on the first demo of a training module integrated in the ACGT Portal. The training module answers the question **How to register a workflow?**

Part 1 of this document presents the principles of online training on which our modules are built. **Part 2** of this document presents recommendations for the development of online training materials. **Part 3** of this document presents an online training demo module, more exactly a training module answering the question **How to register a workflow?**. **Part 4** of this document presents the integration of the online training modules in the ACGT Portal.

KEYWORD LIST: Online Training, Web Portal

MODIFICATION CONTROL			
Version	Date	Status	Author
0.1	14.10.07	Draft	Radu Gramatovici
0.2	31.10.07	Draft	Radu Gramatovici
0.9	30.11.07	Draft	Radu Gramatovici
1.0	06.12.07	Release	Radu Gramatovici
1.3	25.11.08	Draft	Radu Gramatovici
1.4	30.11.08	Draft	Radu Gramatovici
2.0	31.01.09	Release	Radu Gramatovici

List of Contributors

- Sorin Portase, SIVECO Romania
- Otto Zelch, SIVECO Romania
- Fatima Schera, Fraunhofer IBMT
- Nicolas Spalinger, Healthgrid
- Andreas Persidis, Biovista
- Johan Karlsson, UMA
- Thierry Senstag, SIB

Contents

EXECUTIVE SUMMARY **9**

PART 1 ..... **10**

1 NOBODY READS **11**

1.1 NOBODY READS HELP DOCUMENTS 11
1.2 NOBODY READS LONG DOCUMENTS 12
1.3 NOBODY READS 13

2 ONLINE TRAINING DELIVERY **14**

2.1 WHEN? 14
2.2 WHERE? 14
2.3 WHAT? 14
2.4 HOW? 14

PART 2 ..... **15**

3 GENERAL SPECIFICATIONS **16**

3.1 TECHNICAL REQUIREMENTS 16
3.1.1 Technologies *16*
3.1.2 Functionalities *16*
3.2 INSTRUCTIONAL DESIGN REQUIREMENTS 16
3.2.1 Structure, organization *16*
3.2.2 Metadata *17*
3.2.3 Scenarios *17*
3.2.4 Learning items *17*
3.2.5 Content *18*
3.2.6 Language, terminology *18*
3.2.7 Pedagogical approach *18*
3.2.8 Interaction *19*
3.2.9 Aesthetical value. Presentation *19*
3.2.10 Ease of usage. Browsing *20*
3.2.11 Multimedia *20*
3.2.12 Other requirements *21*

4 DETAILED DESCRIPTION OF THE LEARNING ITEMS **22**

4.1 TEXT 22
4.2 ADDITIONAL SOURCES OF INFORMATION 24
4.3 IMAGES 27
4.4 THE MAP 29
4.5 DIAGRAM 31
4.6 AUDIO MATERIAL 32
4.7 AUDIOVISUAL MATERIAL 33
4.8 ANIMATION 34
4.9 SIMULATION 36
4.10 INTERACTIVE MATERIAL 38
4.11 SOLVING PROBLEMS 40
4.12 EDUCATIONAL GAME 42
4.13 TEST (EVALUATION) 44

5 STYLES **47**

6 GLOSSARY **48**

PART 3 ..... **52**

7 USE CASE **53**

7.1 DEFINITION 53
7.2 MODULE CODIFICATION 53
7.3 RESTRICTIONS 53
7.4 PREREQUISITES 53
7.5 ACTORS 53
7.6 FLOW 54

8 INTEGRATION IN THE ACGT PORTAL **62**

8.1 THE "HOW TO" SECTION 62
8.2 SECTION FILTERING 63
8.3 SERVICE/WORKFLOW LINKING 64

9 THE TRAINING MODULE DEMO **66**

PART 4 ..... **67**

10 ONTOLOGY VIEWER TUTORIAL **68**

10.1 TAKING ACTION WITH ONTOLOGY CONCEPTS 68

11 VICTORIA SEARCH ENGINE WITHIN MAGALLANES CLIENT **70**

11.1 COMPOSING COMPLEX WORKFLOWS WITH MAGALLANES 70

12 JORCA TUTORIALS **81**

13 DATA MINING TOOLS TUTORIAL **82**

13.1 USER ACCOUNT CREATION 82
13.2 USER ACCOUNT CREATION 85
13.3 THE GRIDR PORTLET 88
13.4 CREATION OF A GRIDR-BASED DATA MINING WORKFLOW 89
13.5 WORKFLOW EXECUTION 94

Table of Figures

Figure 1: Types of text: Work tasks....22

Figure 2: Types of text: Scientific text....23

Figure 3: Types of text: Scientific Text within supplementary sources of information....23

Figure 4: Additional sources of information, attached to an image....24

Figure 5: Additional source of information, attached to a hyperlink....25

Figure 6: Additional source of information, attached to an additional source of information....25

Figure 7: Additional source of information, attached to a diagram....26

Figure 8: Additional source of information, attached to an external link (webpage)....26

Figure 9: Image attached to a reminder moment....27

Figure 10: Image used for acquiring attention....28

Figure 11: Image resulted after an interactive experiment....28

Figure 12: Static map....29

Figure 13: Dynamic map, associated to a learning item of interaction/ simulation type....30

Figure 14: Diagram attached to a dynamic map....31

Figure 15: Audio material associated to an historical film....32

Figure 16: Audiovisual material exemplifying a virtual experiment....33

Figure 17: An animation which illustrates a physical phenomenon....34

Figure 18: An animation through which a war setting is recreated....35

Figure 19: Simulation by which various situations in solving a problem are explored....37

Figure 20: Simulation of a html editor....37

Figure 21: Simulation of modification of parameters of a spectrum....38

Figure 22: Simulation of the emission or absorption spectra of elements....39

Figure 23: Problem solving....40

Figure 24: Problem solving with choice of answer versions at every step....41

Figure 25: Educational game....42

Figure 26: Stages of the game by which reaching the operational objective is verified....43

Figure 27: Multiple choice test 44

Figure 28: Test with association of elements 45

Figure 29: Test associating images with text 45

Figure 30: Test with chronological ordering 46

Figure 31: Test with filling in of a text 46

Figure 32: ACGT Portal – Log in 54

Figure 33: ACGT Portal – Select the *Workflow* tab 55

Figure 34: ACGT Portal – Select the *Register* subtab 56

Figure 35: ACGT Portal – Fill-in the workflow registration form 57

Figure 36: ACGT Portal – Conclude the workflow registration form 59

Figure 37: ACGT Portal – The visualization of the workflow metadata after registration 60

Figure 38: ACGT Portal – The visualization of the workflow in the list after registration 61

Figure 39: ACGT Portal – The *How to* section 62

Figure 40: ACGT Portal – The *Tutorials* section 63

Figure 41: ACGT Portal – Publication of tutorials related to a specific section 64

Figure 42: ACGT Portal – Linking the tutorial related to a specific service 65

Figure 43: ACGT Portal – The acgt_guest homepage 66

Figure 44: Ontology viewer homepage 68

Figure 45: Search ontology 69

Figure 46: Take action menu 69

Figure 47: Complex workflow example 71

Figure 48: Search for object data types 73

Figure 49: Workflows tab in Magallanes 73

Figure 50: Search tab in Magallanes 74

Figure 51: Workflow generated in Magallanes 75

Figure 52: Pop-up service descriptions 75

Figure 53: Select service 76

Figure 54: Export the workflow to Taverna 77

Figure 55: Taverna view (1) 77

Figure 56: Taverna view (2) 78

Figure 57: Workflow completion 78

Figure 58: Taverna view (3) 79

Figure 59: Taverna view (4) 79

Figure 60: Taverna view (5) 80

Figure 61: User account creation in the ACGT Portal 84

Figure 62: ACGT MyProxy tool 84

Figure 63: The ACGT Portal before logging in 85

Figure 64: The ACGT DMS portlet 87

Figure 65: The ACGT GridR portlet 89

Figure 66: A GridR component 90

Figure 67: ACGT services tree 91

Figure 68: ACGT RScript definition 92

Figure 69: ACGT Workflow editor commands 93

Figure 70: GridR icon from the workflow editor toolbar 93

Figure 71: Service connection into a workflow 94

Figure 72: Service connection into a workflow 95

Executive Summary

ACGT is an Integrated Project (IP) funded in the 6th Framework Program of the European Commission under the Action Line "*Integrated biomedical information for better health*". The high level objective of the Action Line is the development of methods and systems for improved medical knowledge discovery and understanding through integration of biomedical information (e.g. using modelling, visualization, data mining and grid technologies). Biomedical data and information to be considered include not only clinical information relating to tissues, organs or personal health-related information but also information at the level of molecules and cells, such as that acquired from genomics and proteomics research.

The ultimate objective of the ACGT project is the development of European Knowledge Grid infrastructure offering high-level tools and techniques for the distributed mining and extraction of knowledge from data repositories available on the Grid, leveraging semantic descriptions of components and data and offering knowledge discovery services in the domain of Cancer research.

In the architecture of the ACGT Grid infrastructure, the portal is the gate to the services. Structured on several layers of access rights, the ACGT Portal is offering information to all the actors playing in the ACGT project and more general in the domain of Cancer research.

The present document reports on the is structured in four parts.

Part 1 of this document presents the principles of online training on which our modules are built. Part 1 is formed by a section of DON'Ts, where we followed by a section of DOs, where we present the principles that we use in the development of online training modules.

Part 2 of this document presents a recommended methodology for the development of online training modules.

Part 3 of this document presents an online training demo module, more exactly a training module answering the question **How to register a workflow?**.

Part 4 of this document presents the integration of the online training modules in the ACGT Portal.

PART 1

Principles of online training modules

1 Nobody reads

1.1 Nobody reads Help documents

When delivering the online help for using an IT application, users are not willing to frequently switch to a Help section or to frequently open an user manual. Instead, what they really need is to find fast and simple answers to the ordinary problems they encounter when performing specific actions.

Online help was originally thought of as an independent support system. It allows the user to access information easily and immediately, whenever it is needed. As such, it was thought to be the most often used support system on a user's computer.

However, from a number of sources, we have learned that online help is used less frequently than we originally thought. Usability tests, especially on web applications, often show that users do not use the online help; or if they do, they do not find the required information. Field engineers and customer support representatives often report that they need to explain information that is documented in the help because users do not use it. To provide more effective assistance, we must answer the question: "Why is online help not used?"

Is it because of poor user interface design? Poor usability? That is: Are users unaware that there is a help system? Is the term "Help" discouraging to use? Does it take too long to access?

Is it because the help is designed poorly in terms of content and structure? That is: Once users request help, are they getting the information they need, designed in a format they can use? Are users expecting one kind of information and getting something different?

Is it because users are smarter these days and need less online help?

When using traditional Help systems the following usability problems were found to be the most severe:

- Not knowing how to begin interacting with the application.
- Misunderstanding the meaning and implications of pushing certain buttons.
- Not knowing what else can be done with the system.
- Difficulties in finding information using the online help.

First two usability problems can be solved using embedded help that is integrated into the interface. The other two usability problems should be approached through a better organization of and a better access to the information.

Help systems usability include elements like:

- Easy to learn – can people use the help system the first time they open it?
- Easy to remember – can people use the help system more easily the next time?

- Effective – can people easily navigate through the help system, understand the content and put it to the use of solving their problem?
- Efficient – can people find in the help system what they need and accomplish their goals in a reasonable amount of time?
- Satisfying – do people have a good feeling about using the help system? Do they feel it was worth their time to use it? Will they use it again?

1.2 Nobody reads long documents

Many times training materials are long and comprehensive, like usage manuals. These are not well perceived by users that usually have no time to read all the manual, but are rather interested in a specific topics. Then, how to organize a help system in order to avoid the

Embedded Support

For the most critical usability problems, primarily those that are associated with conceptual issues, the help embedded into the user interface itself. Embedded support explains:

- **How to begin interacting:** On the first screen that is displayed when opening the system for the first time, the user is told what they can do and how to proceed.
- **What the buttons mean:** Instead of a one-word definition of each icon, we should use tool tips to provide meaningful explanations of what each button would do or mean if pressed.

Most thoughtful usability experts are of two minds when it comes to embedded help systems. On the one hand, it is good to add helpful information about the task at hand right on the page, sometimes in the form of a patch on the screen that reveals the helpful text upon mouse-over.

However, there is something in the soul of the usability professional that believes that if the interface were properly designed, no help would be necessary. But if you cannot build the perfect system, you should at least provide necessary help information without forcing the user to search for it.

Design principles

- Make the information **task-oriented** and highly structured, because the purpose of online help is to get the user back on task as fast as possible
- **Separate information** into distinct information types and include only one information type in each online help topic. Information types can be by content type (task, concept, reference) of user models (beginner, intermediate, advanced).
- Follow **minimalist principles** in designing online help topics. Many online help designers misconstrue minimalism to simply mean brevity; however, minimalist design hinges on being able to make good decisions on what to do, say, or show, and on what not to include.

- **Keep the length of topics short**, perhaps to no more than two screens long. However, there is no infailible length prescription: online help topics need to be "as long as necessary and as short as possible".

1.3 Nobody reads

Written training materials are usually hard to digest. Basically users are not able to relate the information they read with the actions they have to accomplish.

The help system of computer applications attempts to provide useful and usable information on their usage. While standardization and innovations in user interfaces have made computer application easier to use, it is clear that getting started with unfamiliar interfaces, learning advanced features and understanding application domain concepts remains a challenge for many users. Although the help system is typically used only when users cannot proceed with their task and most of the help information is never consulted at all, there is no doubt about the necessity of providing help. Different applications will benefit from different style of help: for example a complex CAD system to be used daily by an engineer will require elaborate tutorials and detailed online reference manuals. On the other hand a web application used once or twice by novice users may benefit from a video demonstration or a one-page set of “get started” directions. Such web applications are even likely to be abandoned by users who can’t accomplish their goal within a few seconds or minutes. Therefore, the challenge is often to choose the best method of providing help for a given application.

In the context of building an infrastructure for non computer specialists, we have to explore what help system might be best suited to improve the learnability of tools available to the general audience, which is formed by clinicians and medical researchers. General access information systems imply that many of the users will be first-time users of the interface, and that they will have limited time and interest in learning the system. Users will want an answer to their question, not necessarily learn all that a tool can do for them.

2 Online training delivery

Following the critiques presented in the previous section, we have established a series of principles to be followed in the ACGT online training modules.

2.1 When?

The online training materials should be available when the user needs them. This means that the most effective training is the one that happens when the user effectively attempts to use the service. Thus, the training should be available in the same time with the usage of the service.

2.2 Where?

The online training materials should be available where the user needs them. This means that the training materials should be available in the same page with the usage of the service.

2.3 What?

The online training materials should contain exactly the information that the user looks for. This means that a specific training material should describe the usage of specific service and nothing more.

2.4 How?

Graphical training materials, simulating the usage of services are more suitable for this online approach than text materials or other related forms. In other words, the training for using a service should be delivered as the service itself is delivered.

PART 2

Development methodology

3 General specifications

3.1 Technical requirements

3.1.1 Technologies

- *Adobe Flash™ (Actionscript 2.0 +),*
- *Adobe Shockwave™,*
- *JavaScript, HTML.*

3.1.2 Functionalities

- The Content has to function under normal parameters on the browsers IE 6.0 + / Mozilla Firefox 1.5 +.
- The content has to check whether the user has the necessary Flash Player version installed on computer and to provide the possibility to download it, otherwise.
- The whole text that appears in lessons must be loaded from an external file, in order to facilitate a subsequent localization. As a support, we provide producers the HTML Area component.
- All content-type-objects must have a preloader.

3.2 Instructional design requirements

3.2.1 Structure, organization

Each interactive multimedia lesson is presented under the form of lesson moments called *Reusable Learning Objects (RLO)*. In their turn, these moments are divided into submoments that are built by concatenating different learning items, by which the educational objectives are associated to each lesson, according to the corresponding curricula, to be approached and achieved.

Every lesson contains autonomous and reusable lesson moments. This organization and structure must provide the possibility to:

- Localize ("search" function) all lessons for every package.
- Localize a certain lesson, depending on the title for a certain subject and level.
- Localize a lesson moment depending on the title, type, level and subject, as well as any combination between them.

3.2.2 Metadata

- Every submoment is accompanied by the following metadata:
- Title (depending on the content)
- Subject
- Level of expertise
- Types of learning items existent in the submoment
- Basic Key-words (depending on the content)
- Size in bits

The labeling is made according to LOM standard (Learning Object Metadata), an initiative of IEEE Learning Technology Standard.

3.2.3 Scenarios

The development of lessons is made by following scenario.

Types of scenarios:

- The scenario for a usual lesson. Includes lesson moments smaller than 20 minutes. Thus, a one-hour lesson includes 3-4 lesson moments.
- The scenario for a synthesis or revision lesson. Includes lesson moments that may exceed 20 minutes, under condition that submoments should not exceed 20 minutes
- The scenario for a lesson of carrying out a didactical project. Includes a single moment of 50 minutes and submoments which are not longer than 20 minutes that guides the student towards the project's realization.

3.2.4 Learning items

- text
- additional information sources
- images
- diagrams
- maps
- audio materials
- audiovisual material
- interactive material

- animation
- simulation
- solving problems
- test
- educational game

3.2.5 Content

The content of every lesson must:

- be related to the objectives of the corresponding curricula for every subject
- facilitate the understanding of notions
- address the known learning difficulties that appear when the teaching of subject is made in a traditional manner
- be compatible with the knowledge level of students
- include complementary sources and references
- be in conformity with medical specific cultural nature
- to be correct from the scientific point of view

3.2.6 Language, terminology

Language has to be compatible with the students' background and level of knowledge.

- The text must be easy to read and written in a simple and comprehensible language.
- The vocabulary must be rich and coherent.
- The text must be characterized through syntactical and grammatical coherence, and long phrases should be avoided.
- The terminology must be in compliance with the international one (wherever this applies, for example in mathematics, physics, biology etc.), but also with the terminology used by the corresponding curricula.

3.2.7 Pedagogical approach

The proposed lesson must:

- raise and maintain the student's level of interest
- involve the student by means of creative activities (ex.: experiments,

- research)
- stimulate the cooperation between students
- encourage the critical thinking
- use a variety of means to present the information
- offer the possibility of multiple representations of the same notion or of the same phenomenon
- contain activities related to real situations of day-to-day life
- contain activities with level of difficulty that increase gradually
- include a variety of learning items
- offer the possibility to evaluate the student (in time)
- offer the possibility of final evaluation

3.2.8 Interaction

The lesson must offer the possibility of interaction in the following modes:

- feedback from the user in different ways (ex.: text, graphics and sound)
- limitating the attempts to find out the right answer
- treating the mistake with gradual feedback, meaning providing gradual multiple explanations of the mistake depending on the stage
- possibility to modify the values of variables in order to obtain a different result
- possibility of choice for solving the problem
- possibility to determine procedures during the learning game

3.2.9 Aesthetical value. Presentation

The graphics must be adequate from the aesthetical point of view and attractive for the students' background, but must not distract their attention from activities.

- The multimedia resources used (sound, image, animation) must have a good quality, high resolution (images), good quality (sounds).
- The graphical structure of the screens must be functional and followed in a coherent way, with the possibility to go back to the previous screen
- The graphics must have a functional role, not only ornamental.
- Must contain adequated visual and/or sound messages and feedback

- Generally, the entire multimedia presentation will support the objectives of learning

3.2.10 Ease of usage. Browsing

The necessary skills for using the lesson must be in conformity with the level of student's skills

- The lesson must ensure an easy browsing (content, accessible navigation tools, access to options menus)
- Browsing must be without excessive waiting times
- There must exist help available in all lesson's screens
- The instructions must be simple, clear and comprehensible
- The students must have the possibility to chose, use and abandon an activity anytime they like
- The student must have the possibility to chose activities
- The instructor must have the possibility to set the sequence of activities

3.2.11 Multimedia

The audiovisual and sounds presentations can offer possibilities of seting the representations of images, subtitles or transcripts for certain texts, as well as sound adjustment.

Lessons use accepted multimedia formats such as:

- Sound: mp3, wav, midi, wma;
- Mode: mono or stereo;
- Bits rate: 16Kbps – 48Kbps;
- Image: jpeg, gif, png;
- Video: minimum resolution 320 X 240 pixels;
- Formats: Windows Media Video (.wmv), Flash, Video (.flv – through flash files).
- Other types which are used for lessons development: Video for Windows (.avi), Real Media (.rm), QuickTime (.mov), MPEG-1 (.mpeg, .mpg), MPEG-2 (.mpg, .m2v);
- Simulations and animations: Macromedia Flash (.SWF), Sun Java, Adobe Shockwave, Quick Time Video and Quick Time VR.

The size of the file for a lesson moment does not exceed 1 MB, for an easy download, in a time interval that should not create waiting moments for the user. Exceptions from this rule are the audio-video files that can exceed this size if the instructional needs will require it (for films or film sequences).

3.2.12 Other requirements

Lessons should be perfectly compatible with the Internet standards, while lessons usage has to take place using a web browser.

- The design and the unrolling of every lesson must be based on adequated, modern technologies; compatible with the international standards for Internet applications.
- Concerning the content (of all forms) from the lessons, the lesson should not violate the rights (intellectual property, restricted commercial rights, etc.) of any third party.
- Auxiliary lessons and materials must be used, both within a compatible eLearning platform, and installed directly on a computer.

4 Detailed description of the learning items

The lesson scenarios will identify and describe in detail the content of every lesson moment and will underline the adequated learning items, which implement the operational objectives of the lesson.

4.1 Text

The text is an item present in all moments and submoments in different ways and having different roles.

Apart from the scientific text on which the education approach is based, there are texts providing navigation help or contextual help (the one which help in complete fulfillment of work tasks).

Displaying the text on the screen and the percentage of optimal text on a screen are established depending on the international standards with the goal of reaching maximal results in the field of memorizing or information assimilation. The properly text should occupy between 25 ~50% of the total page space, the rest of text form information being distributed in windows that the user opens according in proportion as he goes through the course.

The hypertext provides the opening of screens where supplementary information is offered or images are presented.

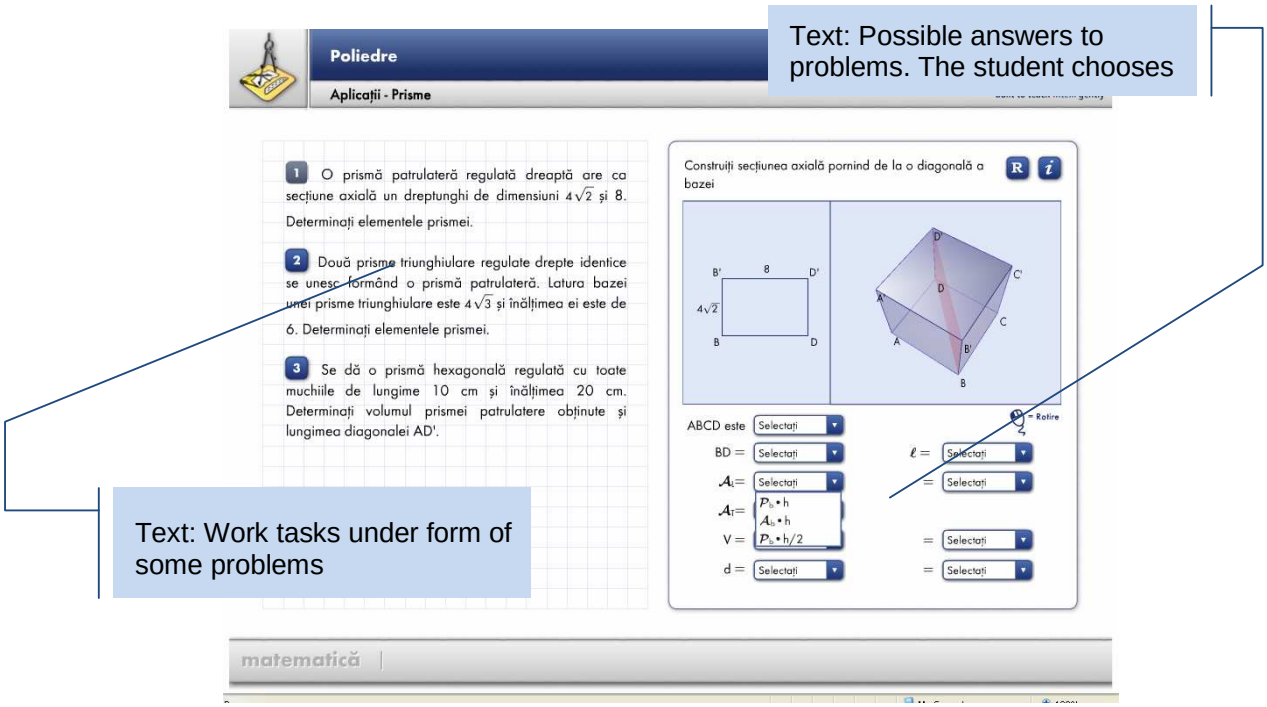


Figure 1: Types of text: Work tasks

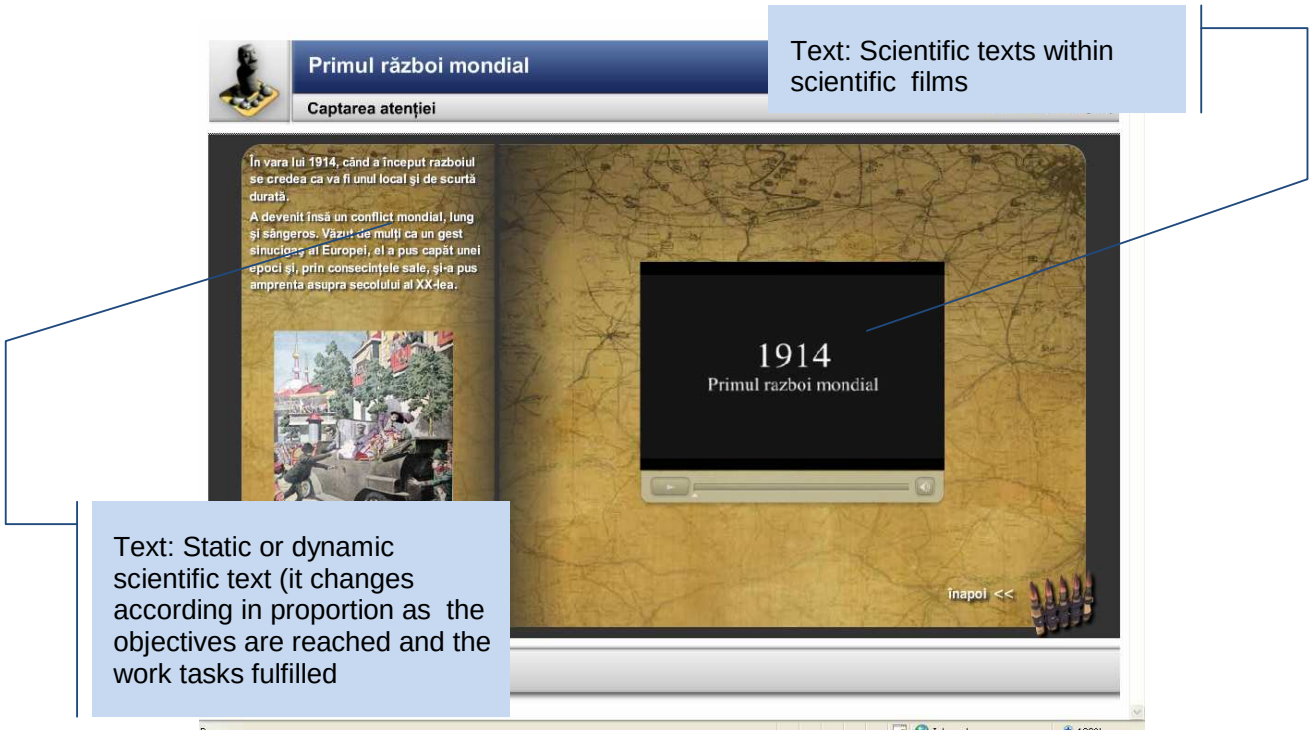


Figure 2: Types of text: Scientific text

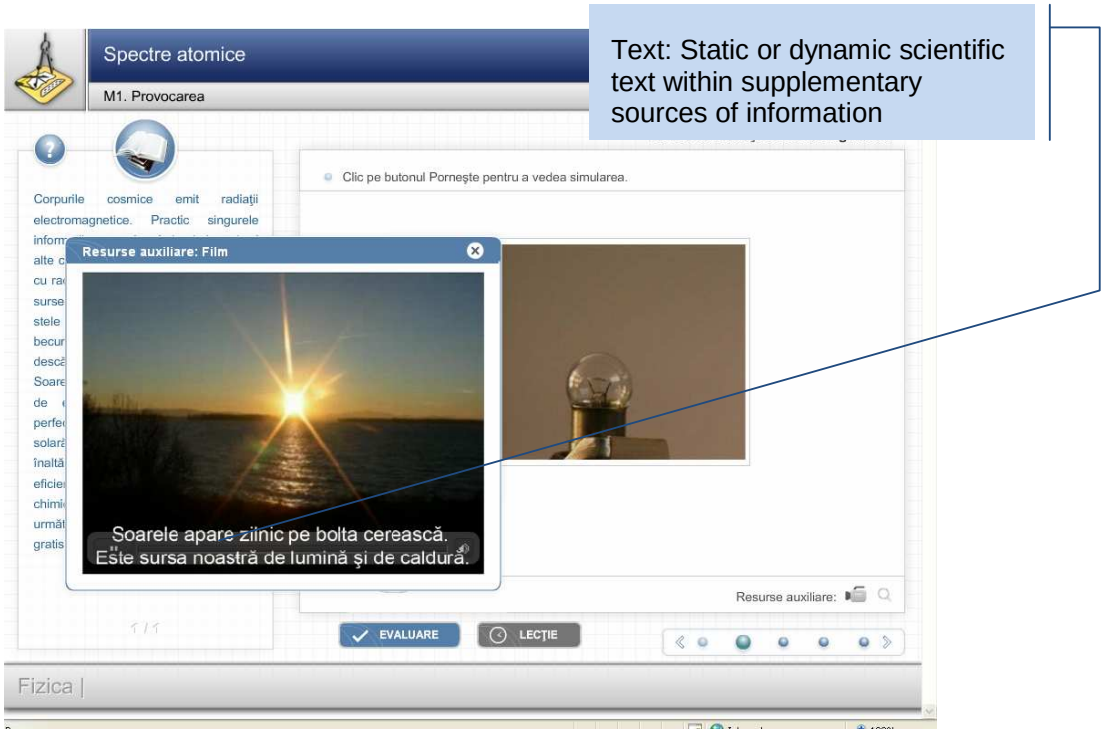


Figure 3: Types of text: Scientific Text within supplementary sources of information

4.2 Additional sources of information

The additional sources can be web addresses to which students have access with the instructor's consent, or bibliography elements or dictionary search. This information sources provide accurate information that harmonizes with the lesson strategy. Usage of these supplementary sources is up to the instructor.

Figure 4: Additional sources of information, attached to an image

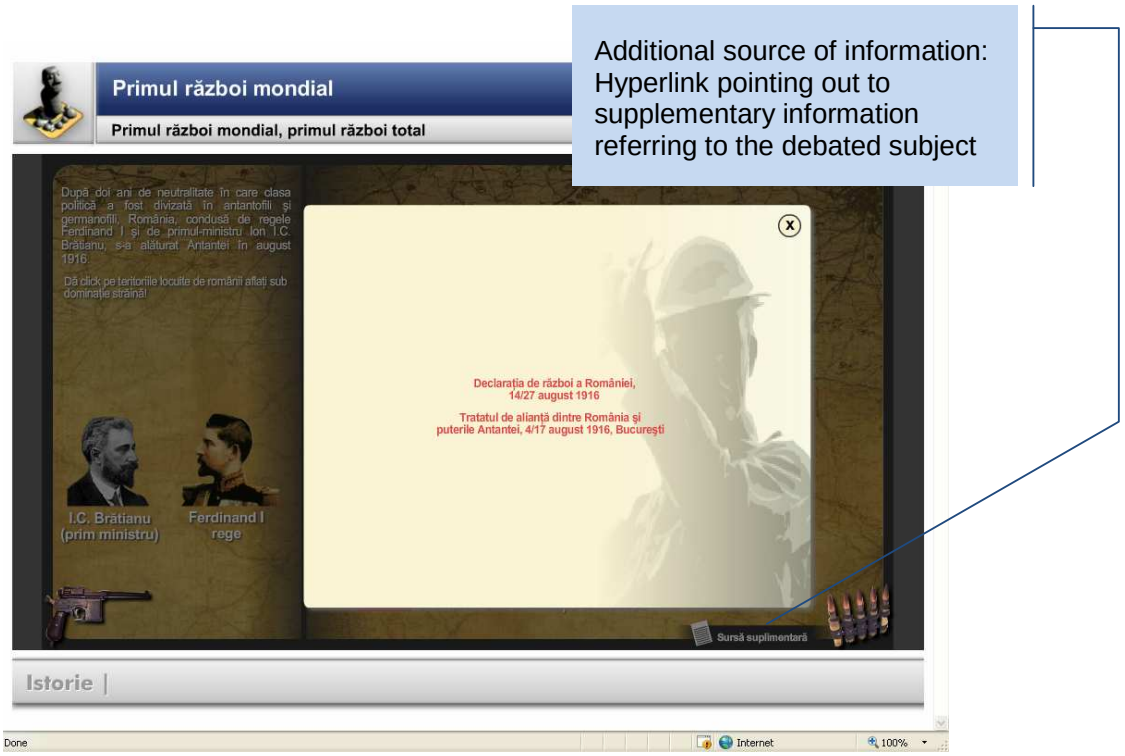


Figure 5: Additional source of information, attached to a hyperlink

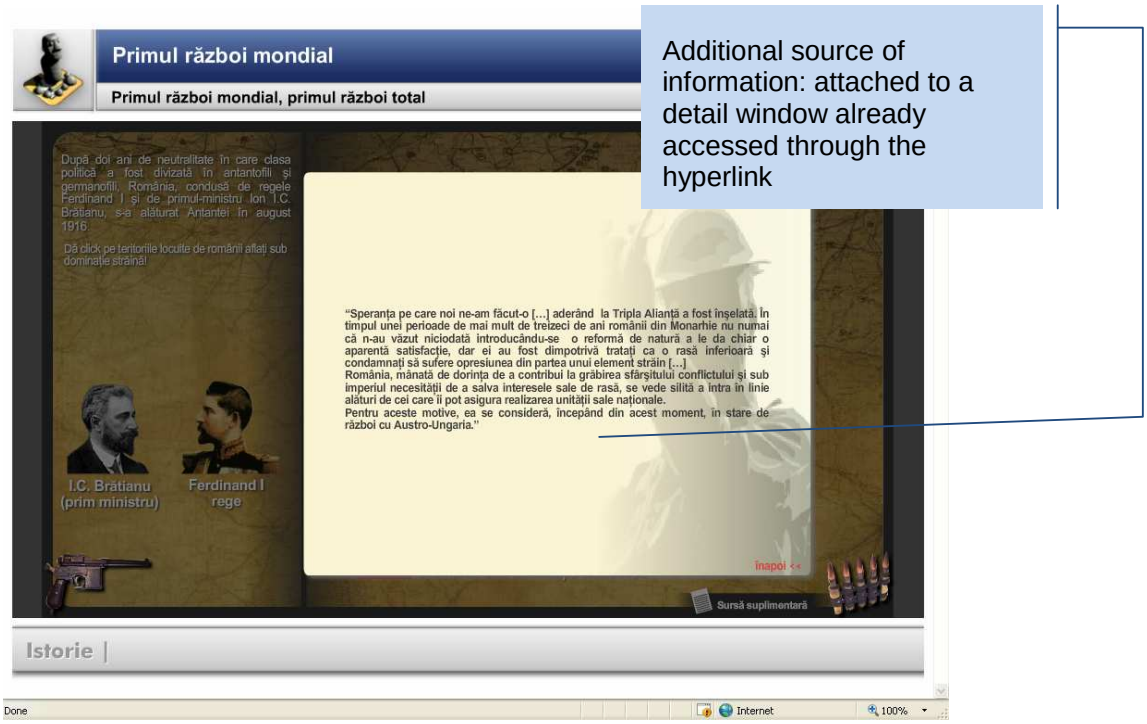


Figure 6: Additional source of information, attached to an additional source of information

Figure 7: Additional source of information, attached to a diagram

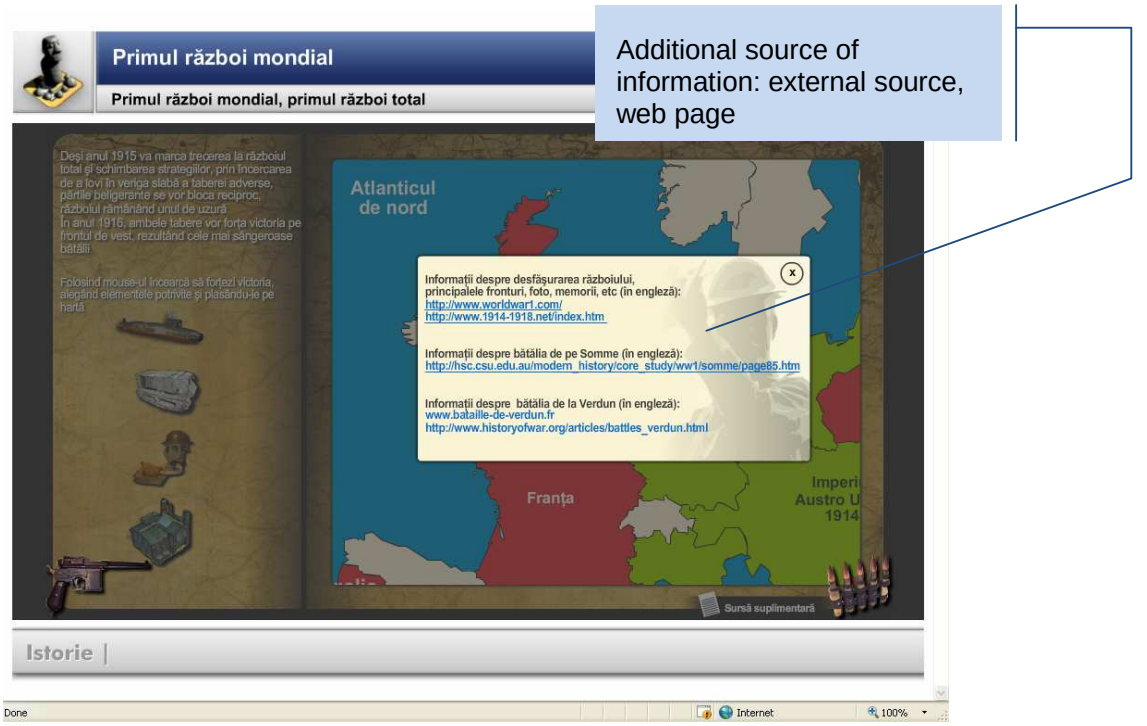


Figure 8: Additional source of information, attached to an external link (webpage)

4.3 Images

Images are another element present in the most moments and submoments, forming, for most cases, along with the text, a first acquaintance for the student with the subject. The hyperlinks (hypertexts) lead in general to items of text or image type providing adjacent information in additional screens.

The image contributes to memorizing, clarification or classification of notions.

The use of images contributes to representation of reality.

The content and character of images are carefully verified, as well as their good presentation both in the page and from the resolution's point of view.

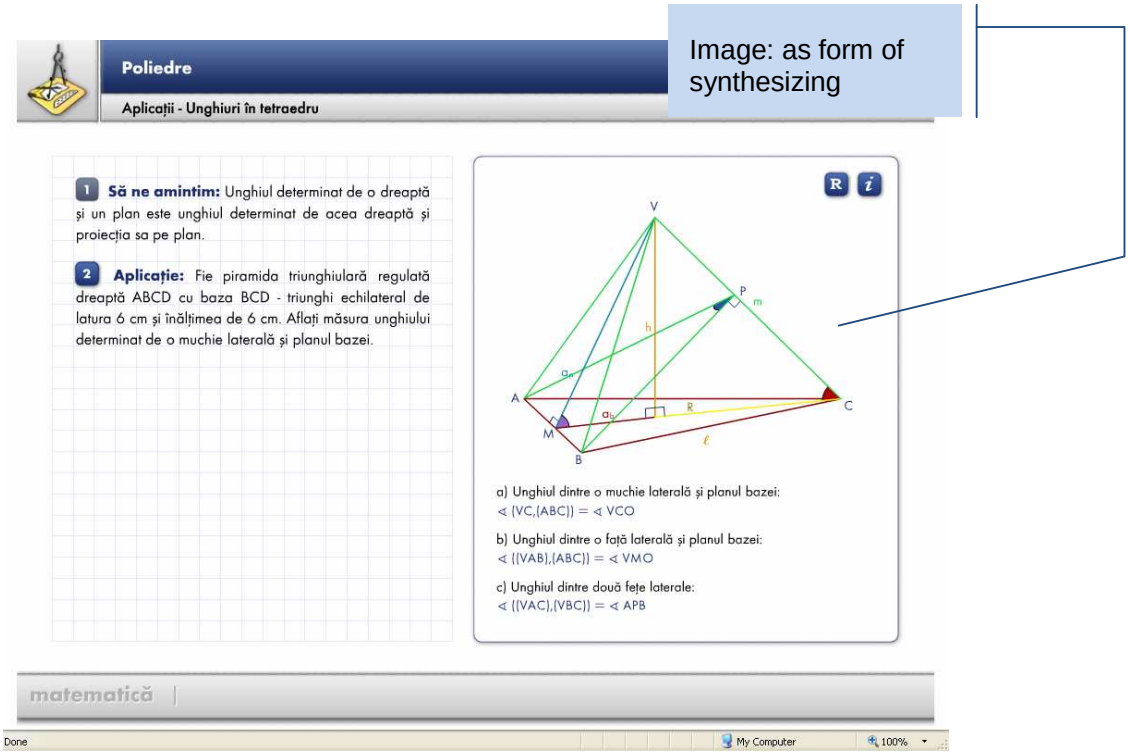


Figure 9: Image attached to a reminder moment

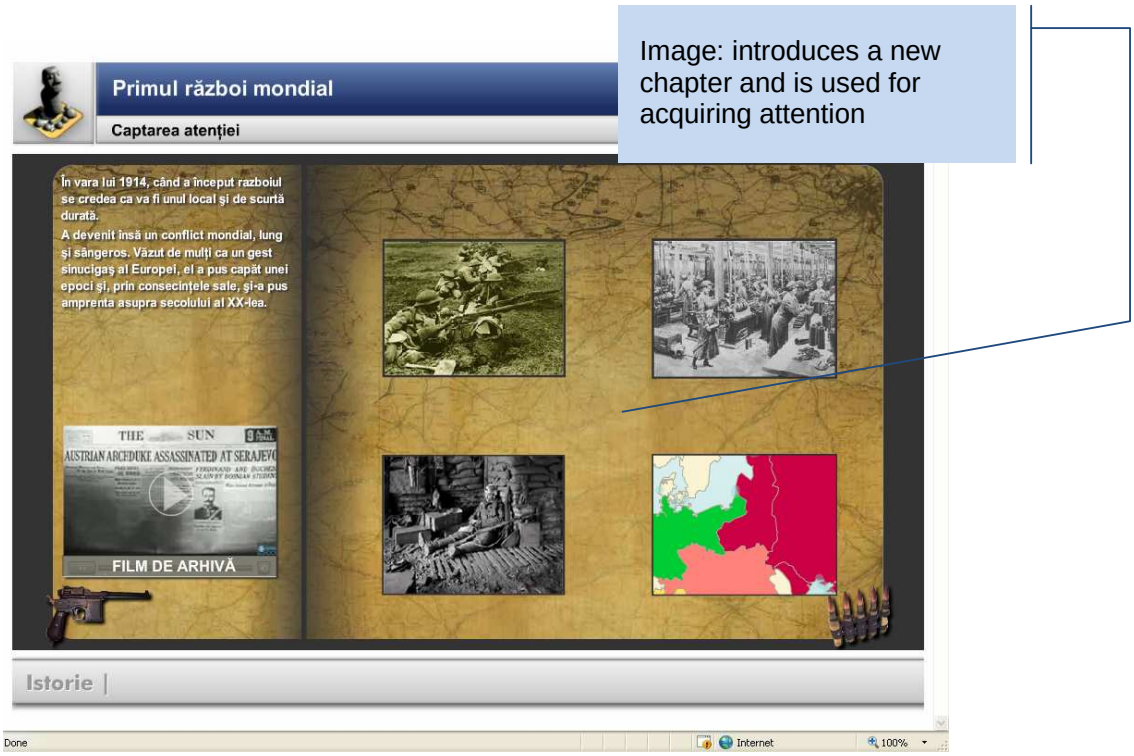


Figure 10: Image used for acquiring attention

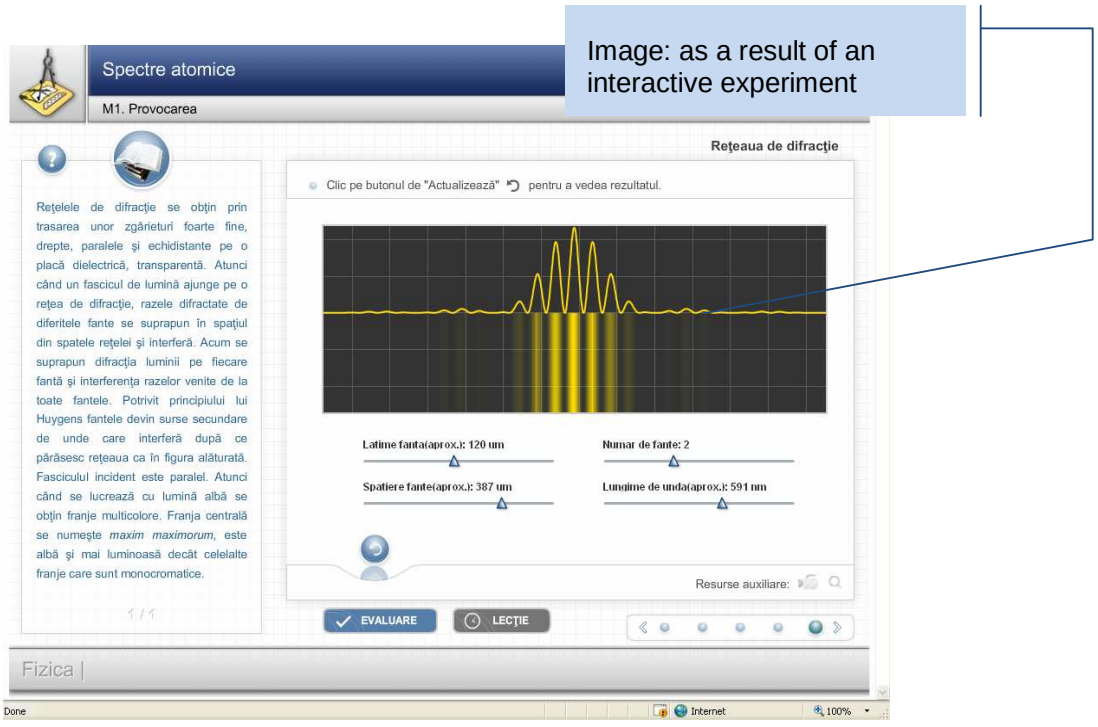


Figure 11: Image resulted after an interactive experiment

4.4 The map

The map is a first active and/or interactive item. The map offers different degrees of interactivity being a basic instrument for creating the orientation spirit and in making decisions. The map will allow the student to simply and quickly locate different objectives, to discover, to explore and to participate. The map solves the problems related to the difficulty of representing data of large dimensions.

Figure 12: Static map

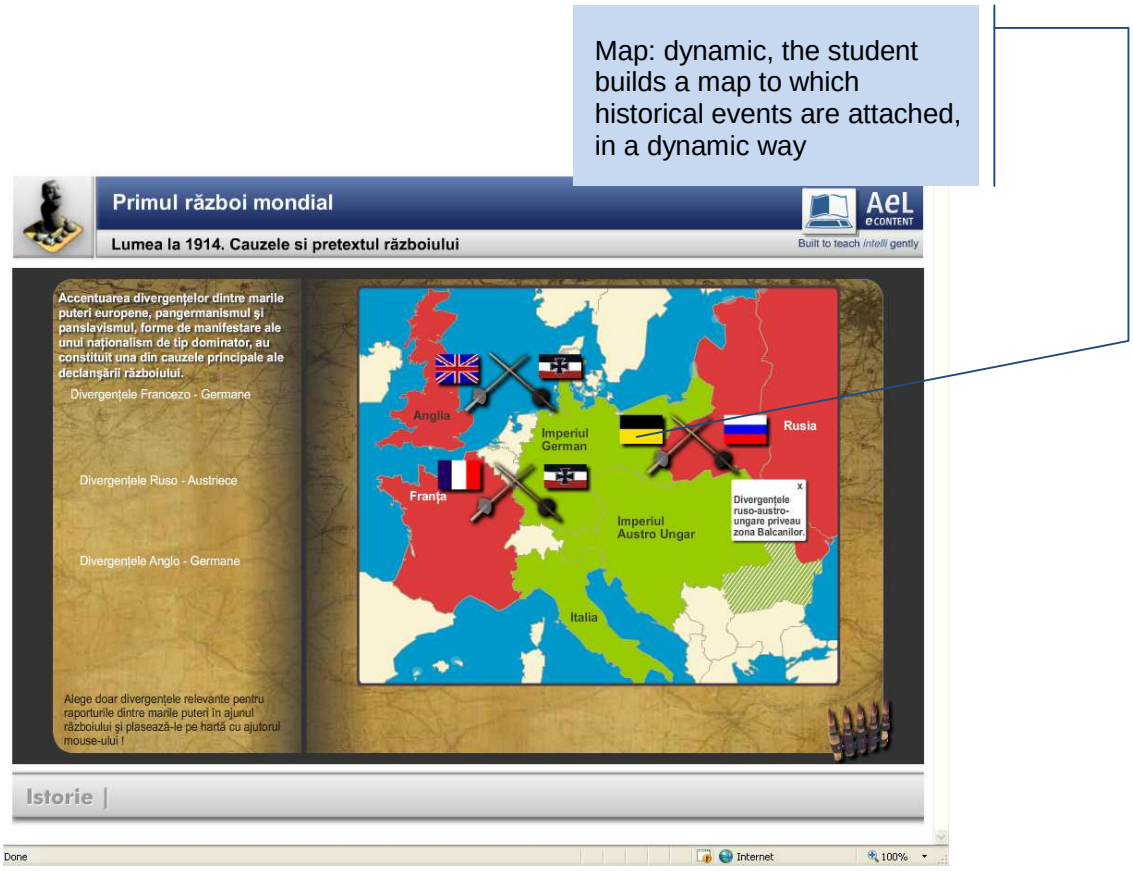


Figure 13: Dynamic map, associated to a learning item of interaction/ simulation type

4.5 Diagram

The conventional map is developed on a single dimension and does not use the entire power of analysis, imagination and creativity of the student.

The interactive diagram uses information in a format much closer to the mental representation of the user and allows the operation of objects in a manner close to the real one, it is a high level description of actions, eliminates the difficulties imposed by the real representation and allows a higher level of abstracting. One can view information and it is much more richer than the textual display. The graphical specifications describe in a simpler but intuitive manner, complex actions, such as processes that cannot be viewed or real-time systems.

The degree of interactivity is decided by the nature of the described processes and by the abstracting capacity of the student.

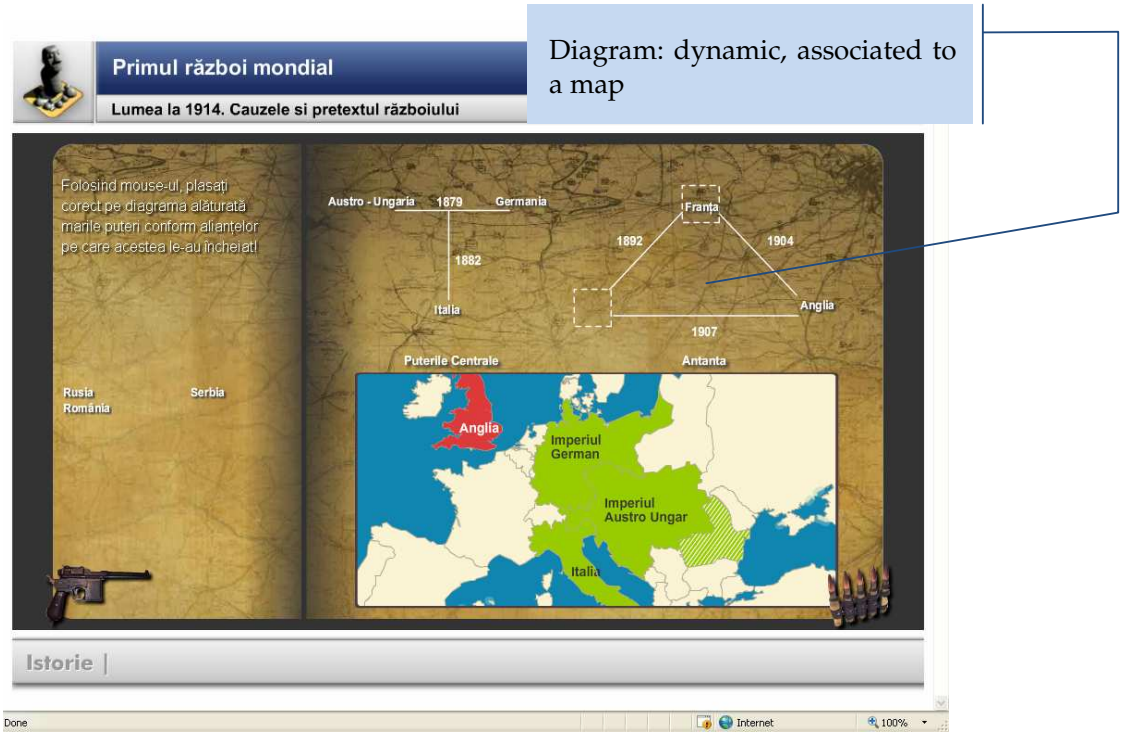


Figure 14: Diagram attached to a dynamic map

4.6 Audio material

The audio material is used for amplifying the educational message in combination with other learning items. The signal level is not strictly considered as an indicator of the audio content quality. Neverheless, the result of the process of recording –> transport –> play –> listening depends very much on the correct choice of the signal levels all along the informational track.

In order to obtain an audio material which satisfies the highest exigencies, the studio equipments (starting from the microphone and ending with a CD writer) are operated at parameters delivering the maximum performance out of them. An audio material is considered good, if both from the content and the technical point of view, there is nothing to claim against it.

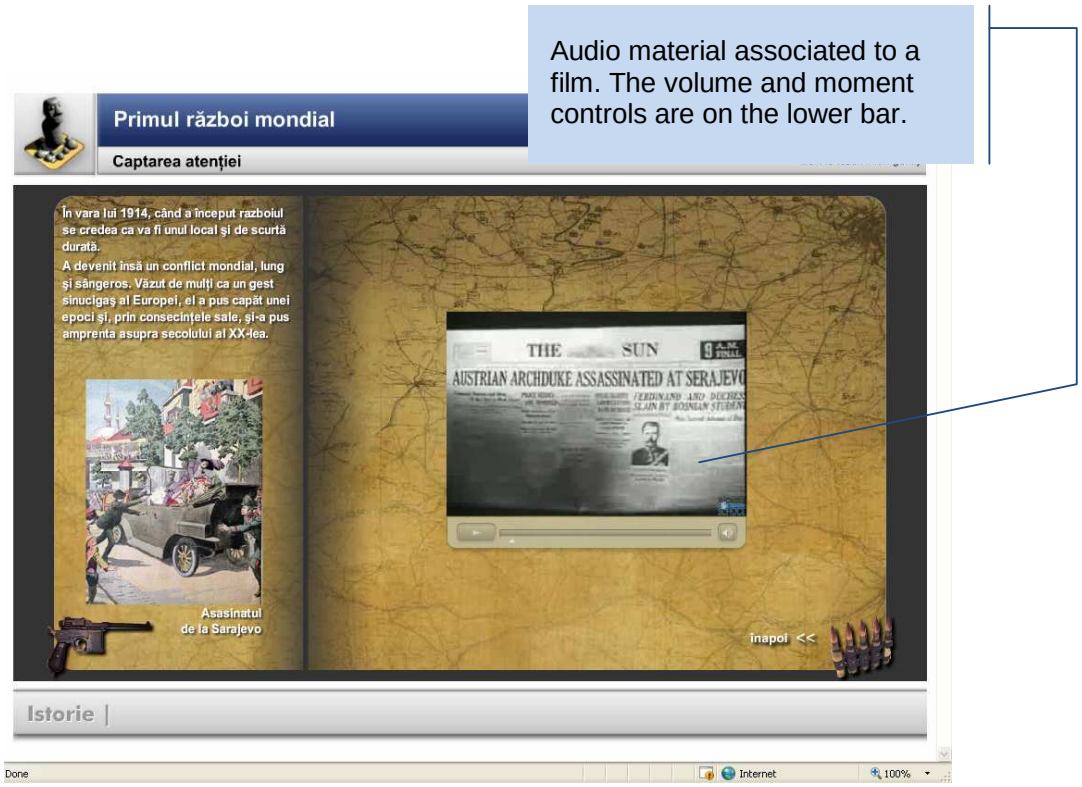


Figure 15: Audio material associated to an historical film

4.7 Audiovisual material

The audiovisual material is, generally, transposed in larger files that makes the transmission to student stations more difficult and increases the waiting time, therefore one uses only materials reduced as duration, but with high impact in representing information (film sequences).

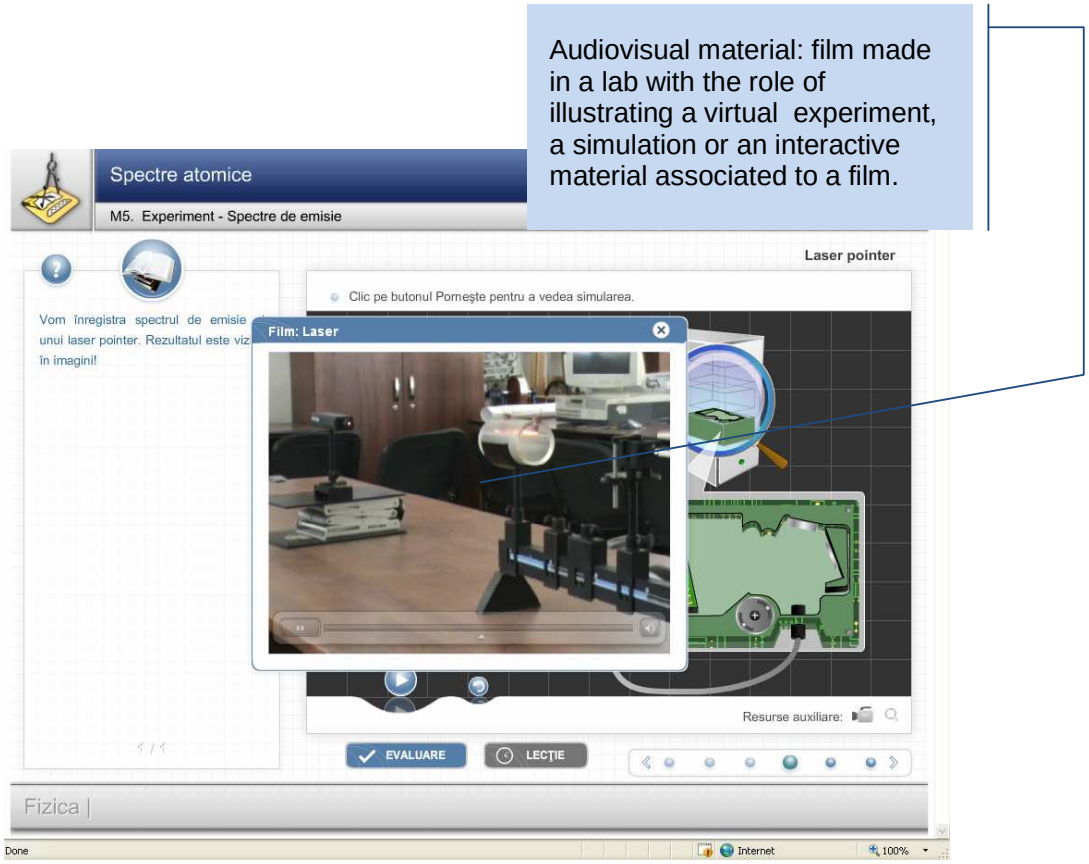


Figure 16: Audiovisual material exemplifying a virtual experiment

4.8 Animation

Animations are used in order to offer a more dynamic aspect, more attractive, for students, through the visual impact they generate to the receiver in the performance of the communication. The student can view the order, coherence, duration, the bulk of the data received through the transmission channel – the computer screen – a process that will then influence his perception and view on the debated subject. The animations are present in all lessons developed, being one of the most important learning items. Animations significantly contribute to creating the sensation of real space and time of the multimedia lessons, being considered a building element of the virtual educational environment.

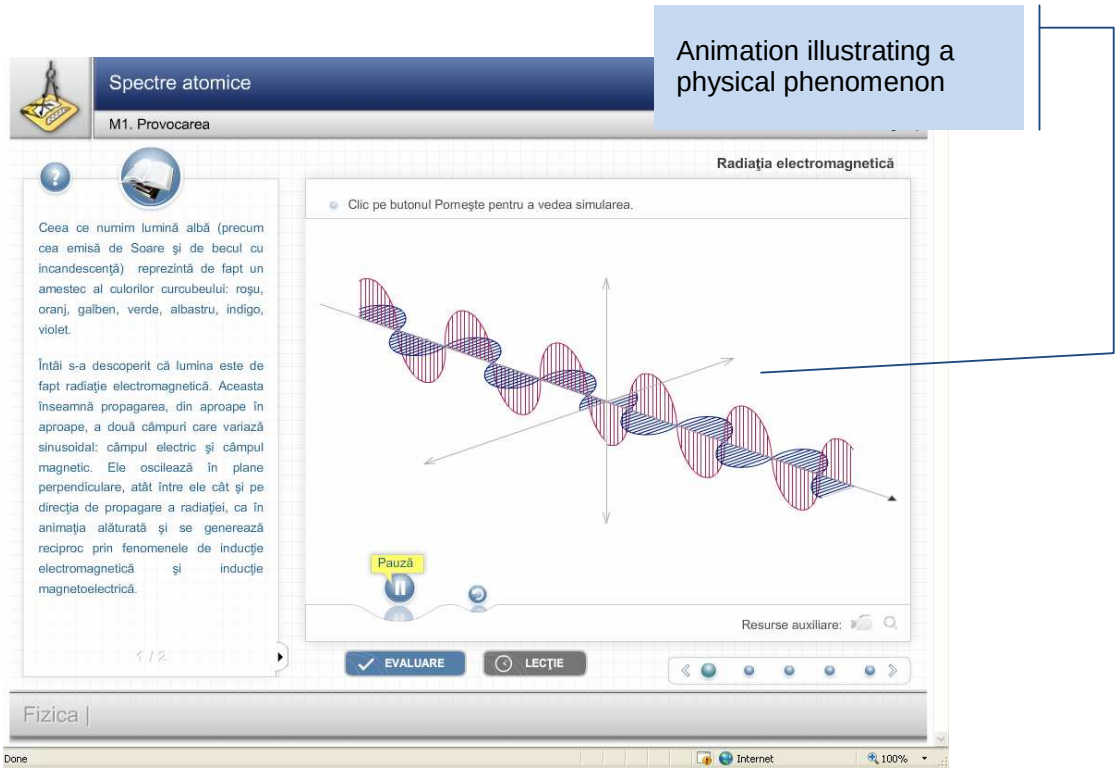


Figure 17: An animation which illustrates a physical phenomenon

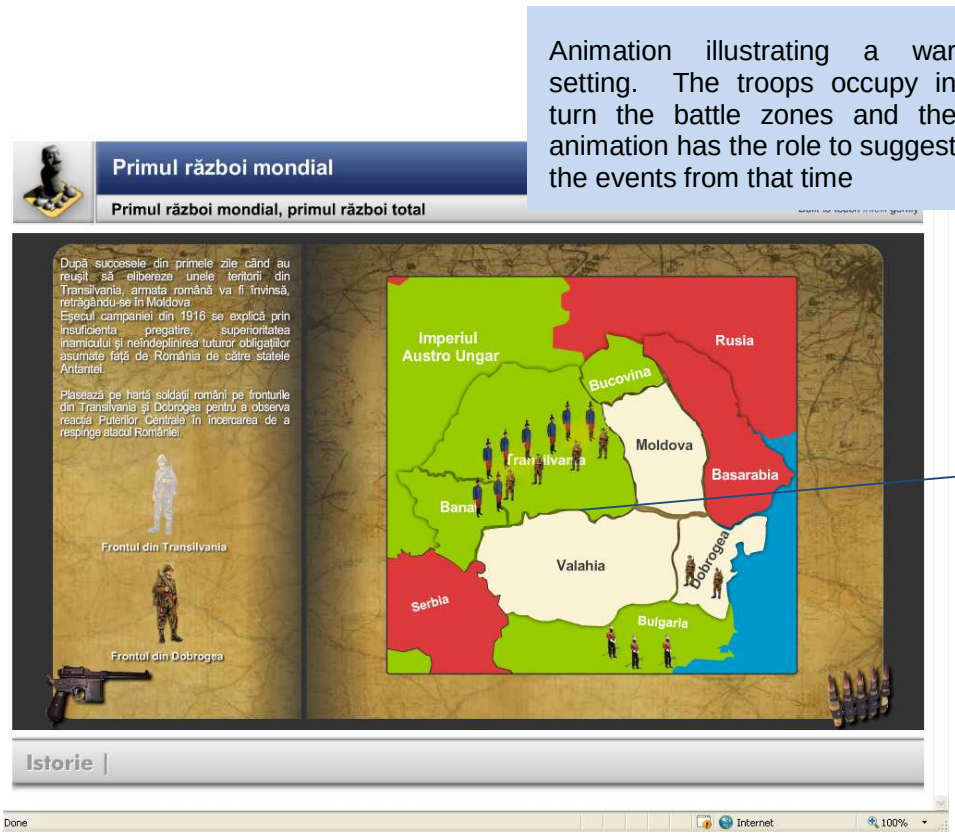


Figure 18: An animation through which a war setting is recreated

4.9 Simulation

Simulation is interactivity, modeling, visualization in the scientific field, by which one obtains the image and the study of different models, processes or phenomena unavailable for direct observation (flows of information, atomic structures, weather systems, cosmic system etc.).

Simulations can be:

- Experiments and analysis for different study subjects, for learning different procedures without endangering the users;
- Systems of simulation for obtaining different competences;

The perceptive- visual learning is realized, mainly through simulation.

Simulation can be realized through a *direct manipulation of real objects* (but which are presented on the screen) – for example the mounting/dismounting of an equipment. Another type of simulation is the one where *the development of a process is accelerated/ slowed down* – trajectory of a missile etc.

The procedural simulations are the ones used especially for forming of abilities, for instance for driving a car or an airplane, or for appropriation of some algorithms, for example for setting of trajectories.

A last type of simulations is the one of problem-situations by which the user is placed in a context: by analyzing it, he must take a series of decisions, for example to lead a research project, to realize a track under the most suitable conditions etc.

Using simulations allows obtaining some training elements that the traditional lesson cannot ensure:

- The intuitive representation on the screen of processes that took place during extremely long or very short periods of time;
- Individual involvement of the student with his own responsibility for the resulted product;
- Avoiding hazardous situations (in case of explosives);
- Observing some phenomena impossible to be seen “live”;
- Repetition /replay of sequence;
- Control on the model with the possibility to modify some parameters.

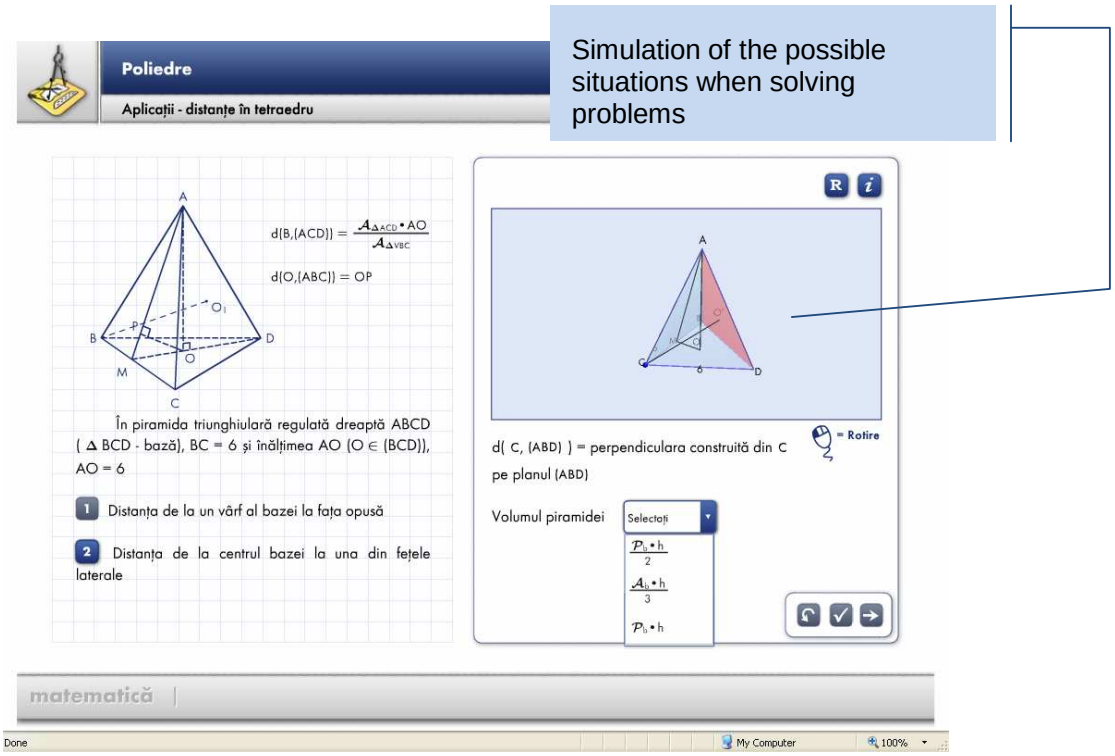


Figure 19: Simulation by which various situations in solving a problem are explored

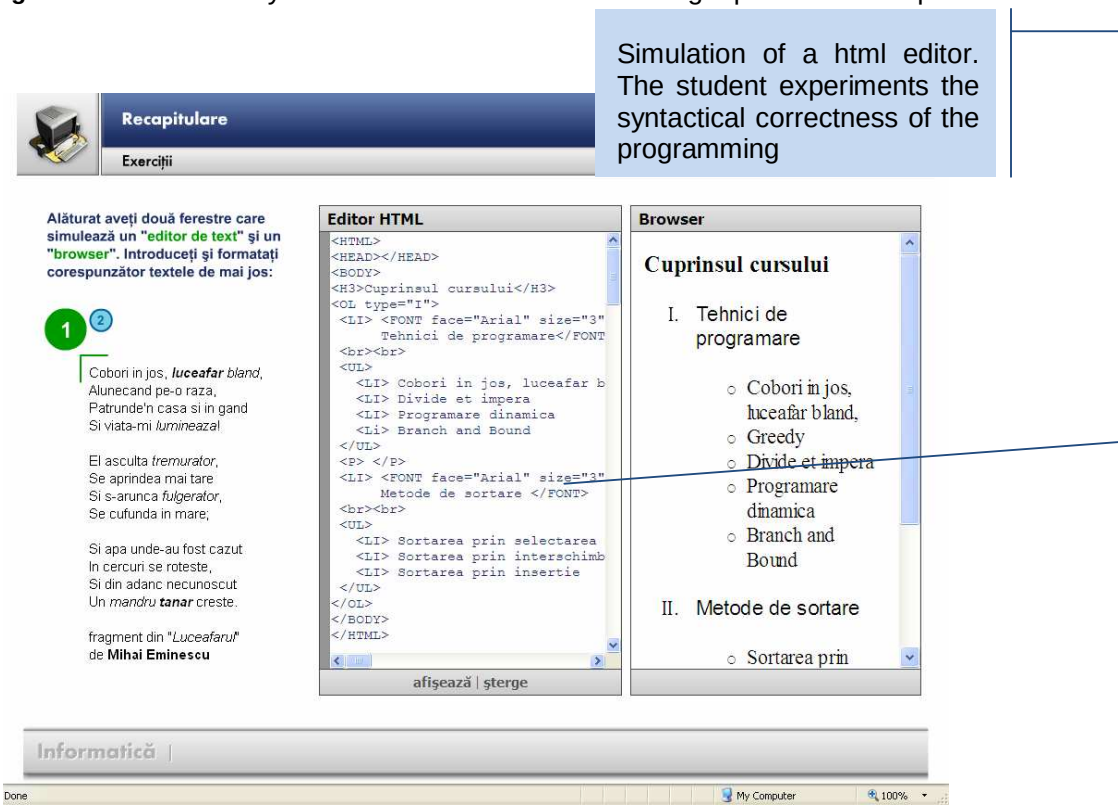


Figure 20: Simulation of a html editor

4.10 Interactive material

Different interactive materials are conceived to serve the learning process. The simple interaction techniques are used for specifying the value of a single input variable. The complex interaction techniques allow the introduction of more comprehensive information, with focus on a field. An interaction technique includes the *student's entrance and exit* from the programme. User's entry is executed through actions at the entrance/ input devices of text or graphical form. The actions performed at the graphical input devices, such as the mouse or the keyboard, are called events or input events. The input events are, for example, pressing a mouse button, moving the mouse coursor, releasing the button, pressing a key etc. The interaction modality specifies the link between the input events and the communication concepts. The input events, generally, are not interpreted isolated, but in sequences called gestures. The most used gestures from the graphical user interfaces are: action (click), press (press-down), release, time (press-timer), (range) and the dragging gesture (drag). Simple interaction - simple tools, by which the input value of a single variable is given. The elementary interaction techniques are presented in the majority of graphical user interfaces: option buttons (radio buttons), option boxes (check boxes), (command button), press button (push button) and the scroll bars or slider. Complex interaction – tools allowing the definition and operation of complex information, by combining several elementary interaction techniques (dialogue boxes, menus – under all existing forms: bar menu, pull-down or pop-up menu, fixed of floating, text or graphical, working areas move by dragging).

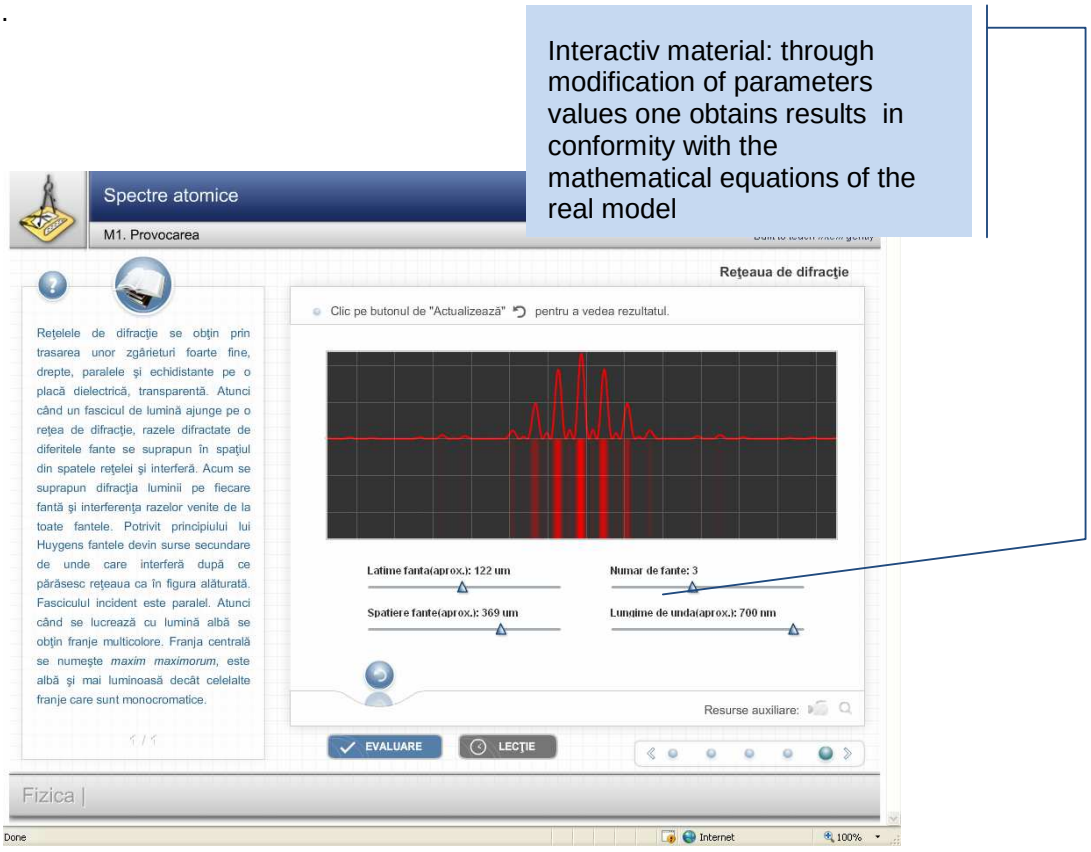


Figure 21: Simulation of modification of parameters of a spectrum

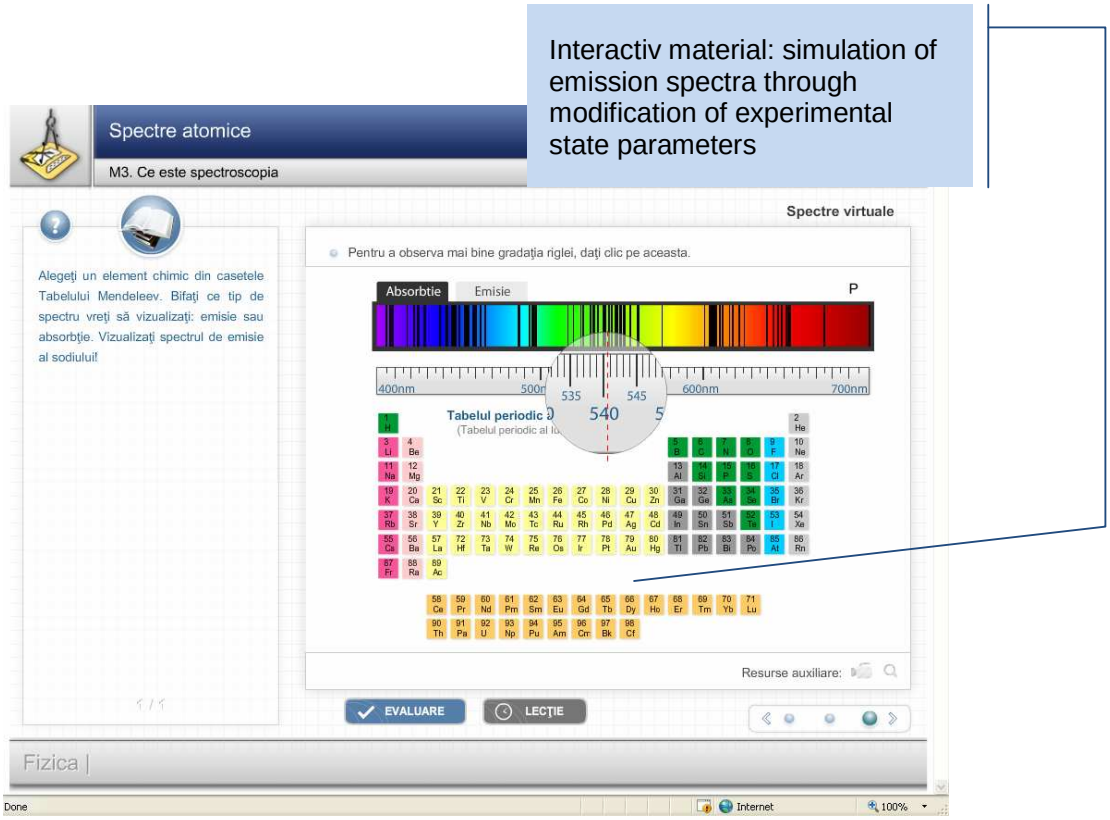


Figure 22: Simulation of the emission or absorption spectra of elements

4.11 Solving problems

These learning items have, through the psycho-pedagogical design, a strategy that aims - through an adaptive interaction - to ensure that the user reaches the objectives for which they have been designed. Some integrate simulations of objects, processes, procedures. The feedback and continuous control determines an individuality of the course, corresponding to the training level of the subject.

Problems solving is a complex process that brings together in a single pedagogical vision the degree of appropriateness and inventivity of the learning scenario, the existence of different possible solutions and the way in which it manages the mistakes and provides feedback.

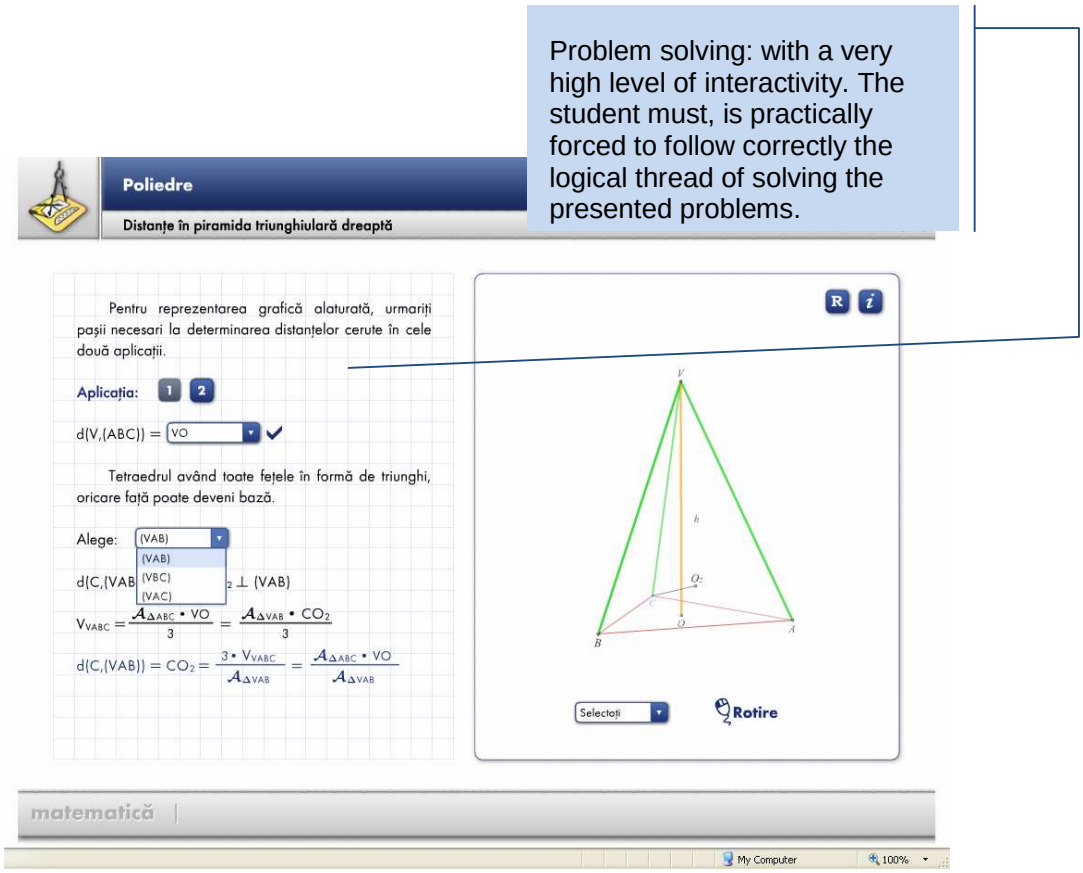


Figure 23: Problem solving

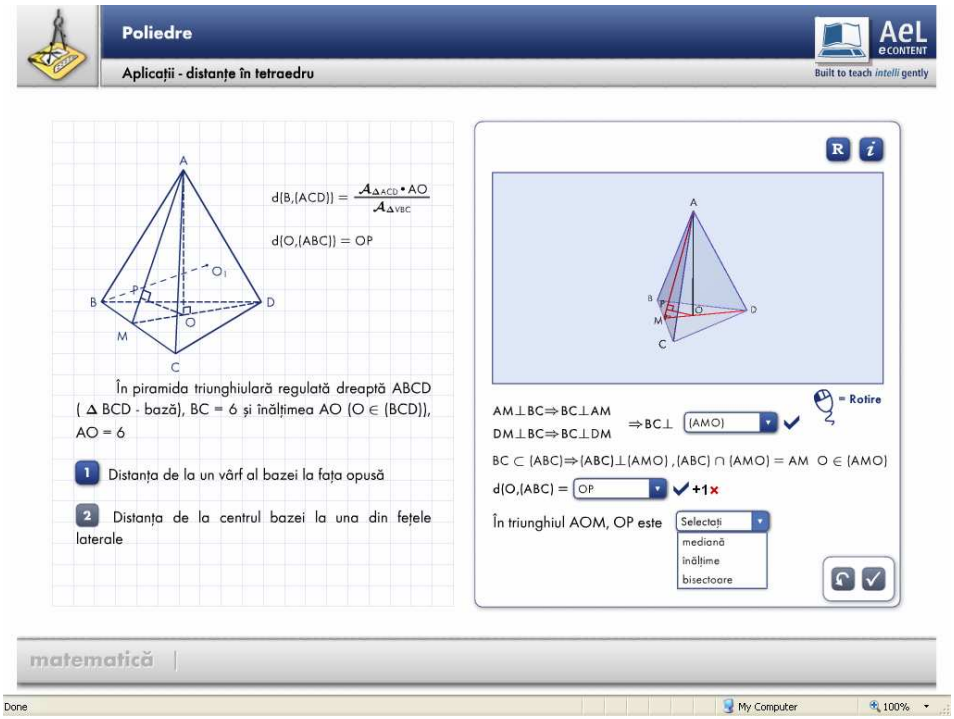


Figure 24: Problem solving with choice of answer versions at every step

4.12 Educational game

The educational item is under form of a game that proposes achieving a goal, by intelligent application of a set of rules – This kind of activity involves the student in a process of problems solving. Usually, a simulation of a real phenomenon is realized, offering the student various modalities to influence reaching the set goal.

The pedagogical structure uses the inventivity of the game combined with the management of mistakes in order to reach some didactical objectives.

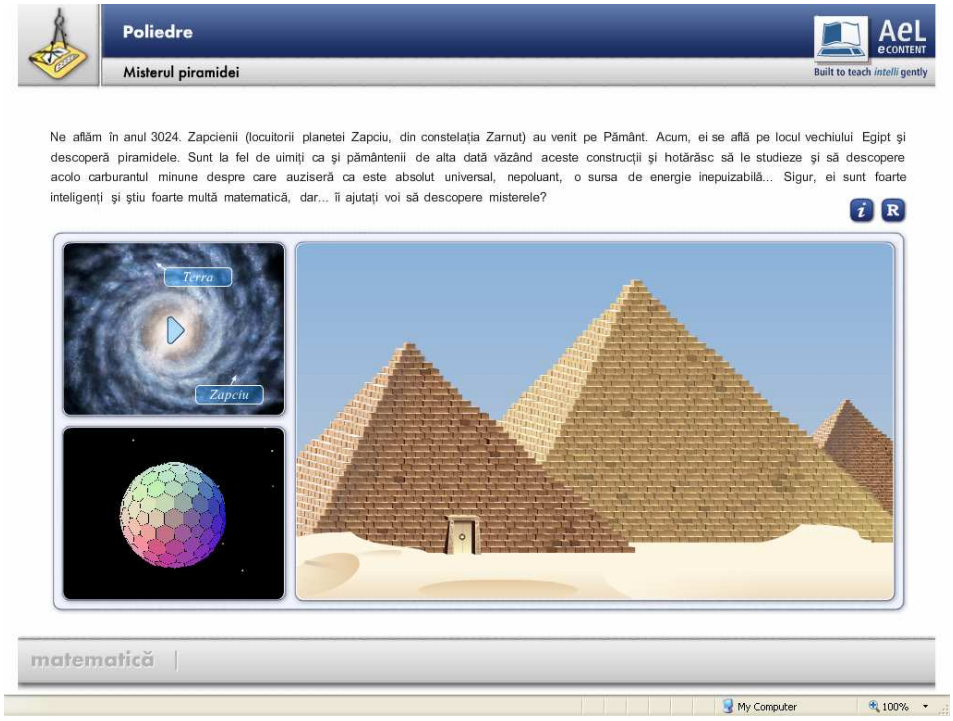


Figure 25: Educational game

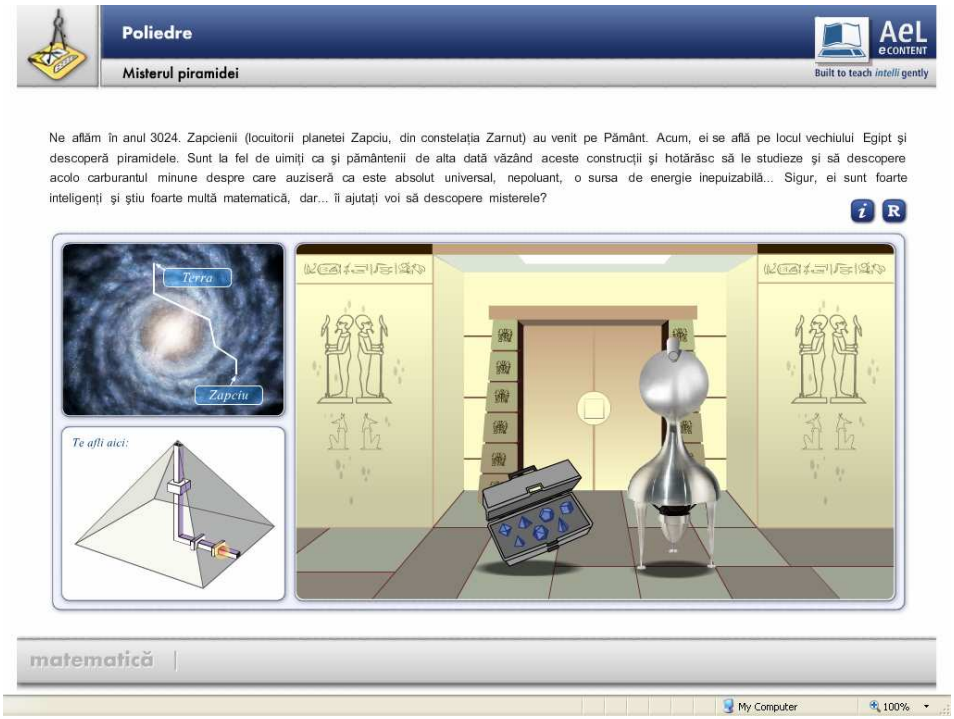


Figure 26: Stages of the game by which reaching the operational objective is verified

4.13 Test (evaluation)

Tests offer a new vision on the evaluation of students' progress through the existence in the lesons of examples of evaluation items, ellaborated in regard to the behavior objectives. These facilitate the instructor's access to a more relevant image on students' progress; thereby there is also the possibility to regulate the teaching-learning process, as well as the training differentiation. Tests also offer a manner of errors management and feedback with important educational valence.

The moments for testing knowledge can be found in lessons possibly in the most various range, as their specificity depends on several factors – the moment of testing, purpose of testing, interaction typology (immediate feedback or not) – these softwares sometimes appear independently, other times are included as integral part from a complex training moment.

Ellaborated for evaluating the training level of the examined person, according to certain standards, criteria or performances, the tests have in our vision also an educational goal, apart from the evaluating one.

Figure 27: Multiple choice test

Figure 28: Test with association of elements

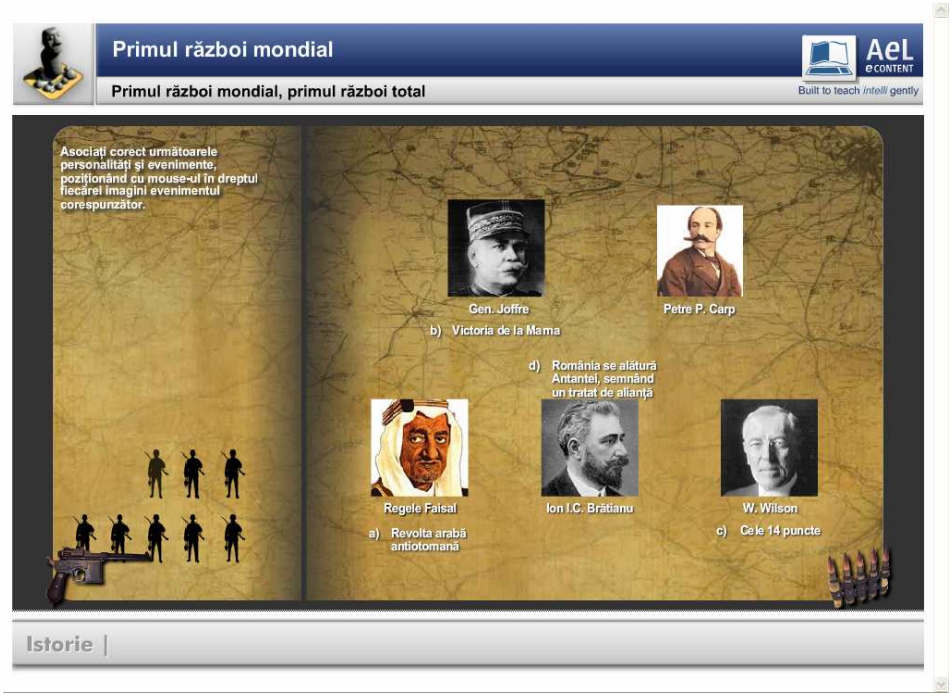


Figure 29: Test associating images with text

Figure 30: Test with chronological ordering

Figure 31: Test with filling in of a text

5 Styles

Things that should be avoided:

- words written with “all-caps”
- colors making the reading more difficult (e.g. red, violet, pale colors)

We recommend the formatting of fonts used according to the table below:

Type of text	**Style**
Link	Underline regular
Quotation	Italic
Scientific text	Regular
Indication of usability	Regular
Title	Bold
Subtitle	Bold
Help text	Regular
Key word	Bold
Hypertext keyword (popup)	Highlighted
Text on button	Bold
Paragraph description (Note:/Use/Form)	Italic
Title for text block (indications, usability)	Bold
Tooltip text	Regular
Educational activity text	Regular
Feedback for activity	Bold
Educational activity reference from Context Text	Regular
Proof of statement	Regular
Proof attached to negative feedback	Regular
Test scoring and conclusions	Bold
Combo box text	Bold

6 Glossary

assessment
Evaluating a learner's actual skill or knowledge level based on the expected skill or knowledge for a person in the same job, position, or assignment.

assessment item
A questionnaire or measurable activity used to determine if the learner has mastered a learning objective.

asynchronous learning
A self-paced learning event. Learners are online at different times and cannot communicate without time delay. Examples: courses taken via Internet, CD-ROM, Web presentation, or videotaped classes.

authoring tool
Software application used to produce interactive learning materials that bring together all components of a course, such as text presentation, graphics, tracking, and links.

broadcast
Transferring learning content to many learners simultaneously, as in a satellite broadcast or an IP multicast. In an IP multicast, numerous learners can participate in a learning event that is broadcast over the network using the Internet Protocol from a single source.

certification
Recognition and acceptance as a qualified professional in a specific area. Certification comes after the learner has successfully completed training and has passed an assessment with a minimum acceptable score. To increase validity and assure authentication, the certification process should be overseen by an independent agent.

class
Scheduled learning event that can take place at a centralized location or in a virtual environment.

cognitive level
A designation that identifies the knowledge and skills (mental or physical) a learner must display to prove mastery of a given reusable information object (RIO).

computer-based training (CBT)
Any instructional event that can be accessed via a standalone computer.

content item
Information stored in a database and used to communicate skills or knowledge. It can be in any media format, including text, graphics, animation, video, audio, and HTML plug-in. It is combined with practice items and assessment items to create a RIO.

content on demand (CoD)
Immediate availability via the network of an offering packaged in a media format such as audio on demand (AOD) and video on demand (VOD).

course
Activities and information, combined in electronic or print format, designed to help learners develop proficiency in a skill or gain knowledge about a given topic.

curriculum
A predefined series of learning events designed to meet a specific goal, such as certifying in a particular area or achieving required job skills and knowledge.

delivery
Any method of providing education. Methods include instructor-led training, Web-based distance learning, online laboratories, CD-ROMs, interactive TV, videos, and books.

distance learning
Situation in which the instructor and students are separated by time, location, or both. Courses are delivered to remote locations via synchronous or asynchronous means.

e-book
Information and graphics that have been organized in electronic or computerized lessons or chapters and made available via computer.

e-learning
Education via the Internet, network, or standalone computer. Network-enabled transfer of skills and knowledge. e-learning refers to using electronic applications and processes to learn. e-learning applications and processes include Web-based learning, computer-based learning, virtual classrooms, and digital collaboration. Content is delivered via the Internet, intranet/extranet, audio or video tape, satellite TV, and CD-ROM.

electronic classroom
A classroom equipped with multimedia devices that are used to enhance learning.

evaluation
Any method used to gather information about the impact or effectiveness of a learning event. Measurements might be used to improve the offering, determine if the learning objectives were achieved, or determine if the offering has been of value to the organization.

feedback
Interaction between the learner and the instructor or system. Feedback increases the quality of the learning experience.

instructional designer (ID)
An individual who, using systematic methodology and instructional theory, creates content for learning events.

Instructional management system (IMS)
IMS is a set of technical specifications developed by IBM defining how learning materials will be exchanged over the Internet and how organizations and individual learners will use these materials.

instructor-led training (ILT)
A scheduled event conducted by an instructor, either in a classroom or on the network. Sometimes called leader-led training (LLT) or lecture/lab training (if the course includes hands-on lab exercises).

lab
A physical or virtual hands-on activity designed to help learners meet skills-based learning objectives.

learner
Anyone who accesses information to increase his or her skills and knowledge.

learning event
An activity in which a learner increases his or her skills and knowledge.

Learning Object Metadata (LOM)
Data model, usually encoded in XML, used to describe a learning object and similar digital resources used to support learning. The purpose of learning object metadata is to support the reusability of learning objects, to aid discoverability, and to facilitate their interoperability

learning objective
A statement that establishes a measurable outcome. An objective is used as an advanced organizer to indicate how the learner's acquisition of skills and knowledge will be measured.

localization
Modifying an offering to meet the specific needs of a geographic area, product, or target audience. For example, America uses avoirdupois weight units while Europe uses metric units.

metadata
Information about content that allows it to be stored in and retrieved from a database.

practice item
A question or learning activity that gives learners immediate feedback, allowing them to determine whether they can apply the skills and knowledge just learned in a RIO.

practices
Activities that reinforce the material being learned and give the learner an opportunity to apply skills and knowledge. Some practices provide feedback and mentoring. The learner's performance may or may not be affected by the results of the practice activities. Practice activities include case studies, learning activities, practice quizzes, practice tests, testing quizzes, and practice labs.

prescriptive learning
Matching a learner with offerings designed to fill gaps in the learner's knowledge and skills.

reusable information object (RIO)
A collection of content, practice, and assessment items based on a learning objective.

reusable learning object (RLO)
Information, based on RIOs, overviews, summaries, and assessments, that supports a specific learning objective.

Sharable Content Object Reference Model (SCORM)
Collection of specifications that enable interoperability, accessibility, and reusability of e-learning content.

self assessment
Process in which the learner determines his or her level of knowledge and skills.

self-paced learning
Learning in which the pace and timing of content delivery are determined by the learner.

subject-matter expert (SME)
An individual who has substantial knowledge about and skills in a specific subject.

synchronous learning
Real-time learning situation that can include immediate, two-way communication between participants.

virtual classroom (VC)
A scheduled offering that is available at multiple locations (either desktop or classroom) via a network

Web-based training (WBT)
Any instructional event that can be accessed via the Internet or the Web.

PART 3

Online training module demo

7 Use Case

7.1 Definition

This training module answer the question **"How to register a workflow?"** This training module does not refer to the Workflow editor. It is supposed that the workflow that is registered was already edited using the ACGT Workflow Editor or other similar tool.

7.2 Module codification

ACGT-WRT-SIV-v1[1]

7.3 Restrictions

This training module is restricted to those users (groups) that are entitled to perform the specific action.

7.4 Prerequisites

Users performing this action may want to learn about other actions preceeding this one:

- How to register a functional category?
- How to register a service?
- How to edit an workflow?

7.5 Actors

The registration of a new workflow is an operation restricted to the users affiliated to the group RESEARCHER.

[1] ACGT-[MODULE NAME]-[INSTITUTION SHORT NAME]-[VERSION NUMBER]

7.6 Flow

The flow of this action starts in the frontpage of the ACGT Portal that can be accessed from any browser at the address http://rd.siveco.ro/acgt/.

1. Log into the portal

Indicate your ACGT *User Name* and the current ACGT portal *Password*.

Figure 32: ACGT Portal – Log in

2. Select the **Workflows** tab

Click on the *Workflows* tab, which appears on the horizontal bar, under the header of the page.

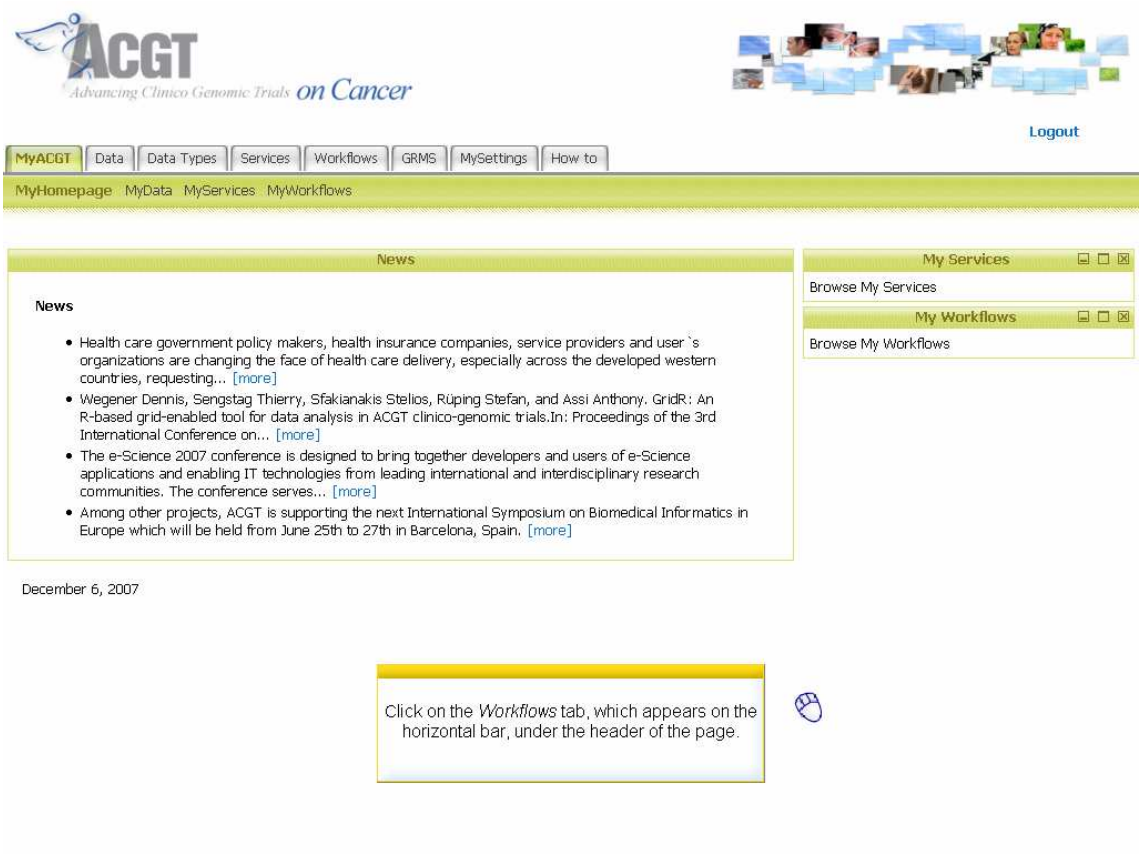


Figure 33: ACGT Portal – Select the *Workflow* tab

3. Select the **Register** subtab

Click on the *Register* subtab, which appears in the horizontal bar, under the *Workflow* tab.

Figure 34: ACGT Portal – Select the *Register* subtab

4. Fill-in the **Workflow Registration** form

Each field of the form should be completed with the corresponding information.

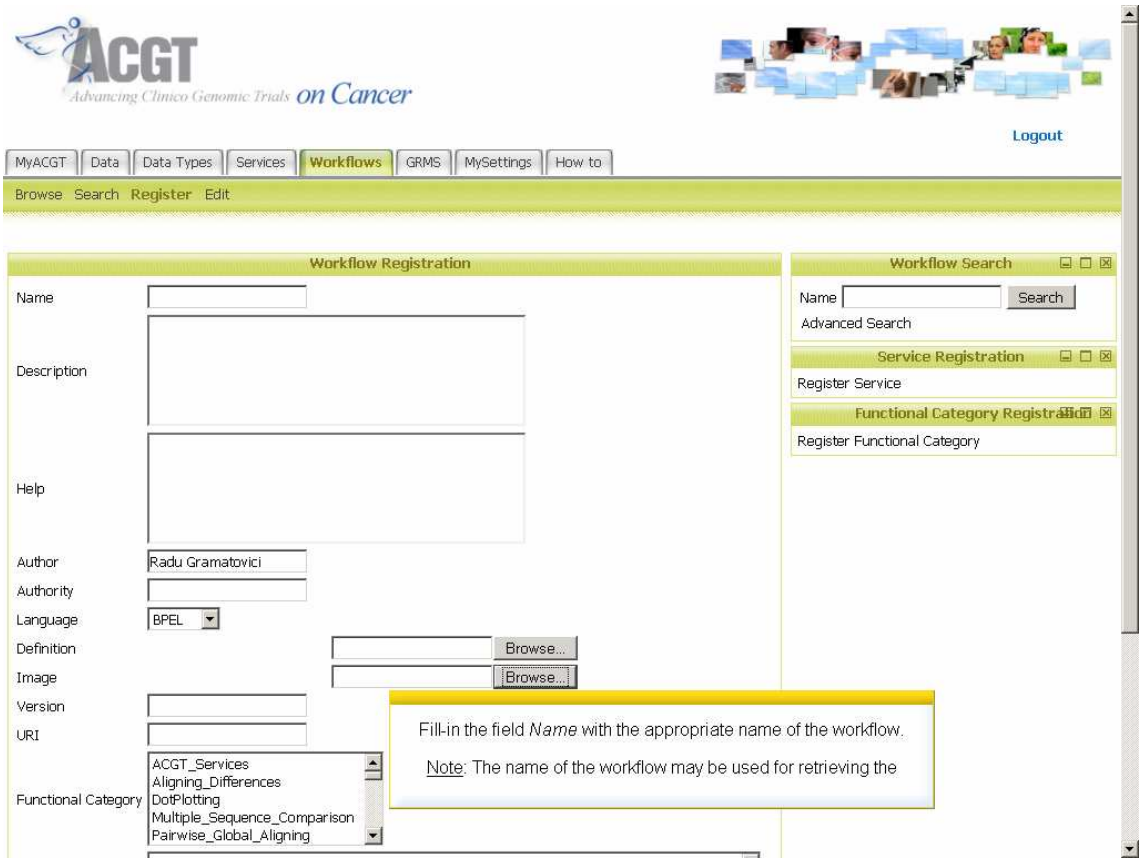


Figure 35: ACGT Portal – Fill-in the workflow registration form

4.1. Fill-in the *Name* of the workflow

Fill-in the field *Name* with the appropriate name of the workflow.

Note: The name of the workflow may be used for retrieving the workflow.

4.2. Fill-in the *Description* of the workflow

Fill-in the field *Description* of the workflow as an appropriate descriptive concise text.

Note: The description of the workflow may be used for retrieving the workflow.

4.3. Fill-in the *Help* reference of the workflow

Fill-in the *Help* reference with a link to the *How to use the* <*workflow*> tutorial in the *How to* section.

4.4. Fill-in the *Author* of the workflow

Fill-in the *Author* field with the name of the author of the workflow.

Note: This field has an implicit value equal to the name of the user that performs the operations. The implicit value can be overwritten. For correct filtering purposes, it is recommended to keep the implicit value.

4.5. Fill-in the *Authority* of the workflow

Fill-in the *Authority* field with the name of the organization that has generated the workflow.

4.6. Select the *Language* of the workflow

Select the value of the *Language* field from the list. The value indicates the language in which the formal description of the workflow is written. The actual values can be BPEL, SCUFL or Other.

4.7. Upload the *Definition* of the workflow

Upload the definition of the workflow in the *Definition* field. To accomplish this task, use the button *Browse* displayed on the *Definition* row.

Note: The file to be attached to this field should contain the definition of the workflow in the language indicated in the previous field.

4.8. Attach the *Image* of the workflow

Upload the diagram representing the workflow as an image in the *Image* field. To accomplish this task, use the button *Browse* displayed on the *Image* row.

4.9. Fill-in the *Version* of the workflow

Fill-in the *Version* field with a corresponding (possibly dotted) number.

Note: The version field should distinguish different versions of the same workflow.

4.10. Fill-in the *URI* of the workflow

Fill-in the *URI* field with a corresponding URI, where the operation execute of the workflow is located.

4.11. Select the *Functional Category* of the workflow

Select the value of the *Functional Category* field from the list. The selected functional category should be the one in which the new workflow fits.

Note: If the workflow doesn't fit in any of the registered functional categories, than a new functional category can be registered using the portlet *Functional Category Registration*, which is present on the page, in the right column. This operation is restricted to selected users.

4.12. Select the component *Service Operations* of the workflow

Select the values of the *Service Operations* field from the list. The selected service operations should be the ones from which the workflow is made.

Note: If some service or some service operation does not exist in the list of registered service operations, than a new service can be registered or a new operation can be added to a registered service using the portlet *Service Registration*, which is present on the page, in the right column.

4.13. Select the *Nested Workflows* of the workflow

Select the values of the *Nested Workflows* field from the list. The selected nested workflows should be the ones from which the workflow is made.

Note: All the workflows nested in the new workflow should be registered prior to the new one, following the same procedure. It is possible for a workflow to have no nested workflows.

5. Conclude the **Workflow Registration** form

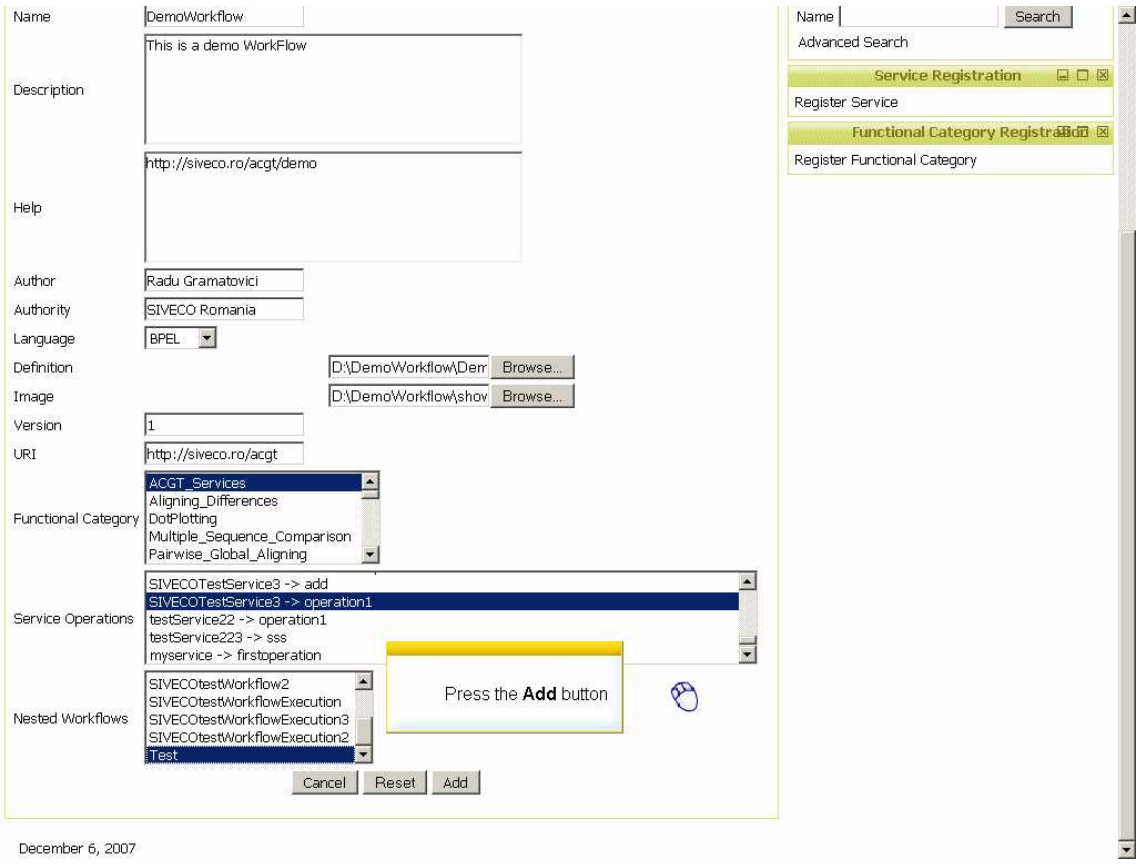


Figure 36: ACGT Portal – Conclude the workflow registration form

The *Workflow Registration* form can be concluded in three ways, by using the corresponding buttons displayed in the bottom of the form:

- Cancel – aborts the current registration and returns to an empty *Workflow Registration* form
- Reset – cleans all the information filled in the fields of the current form

- Add – registers a new workflow with the specific description in the Workflows table of the Metadata Repository database. After this operation the portal switches to the Workflow Browse portlet displaying the new workflow as it was registered in the database.

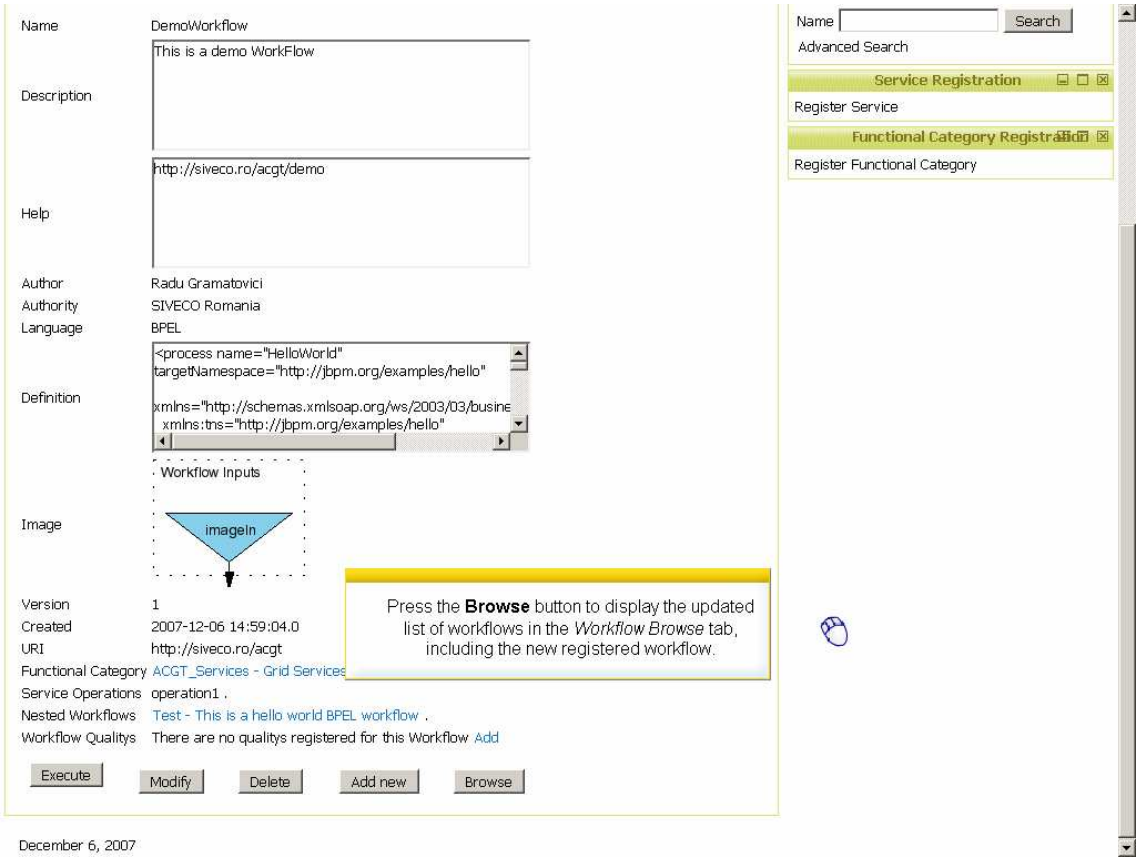


Figure 37: ACGT Portal – The visualization of the workflow metadata after registration

6. Conclude the action in the **Workflow Browse** view

After registering the new workflow, the action can be continued in five ways, by using the corresponding buttons displayed in the bottom of the view:

- Execute – executes the workflow (this action is not part of the Workflow Registration form)

- Modify – reenters the *Workflow Registration* form in order to modify the current values of the workflows

- Delete – deletes the current wokflow from the Workflows table, returning the updated list of workflows in the *Workflow Browse* tab

- Add new – opens a new empty *Workflow Registration* form in order to register a new workflow

- Browse – displays the updated list of workflows in the *Workflow Browse* tab, including the new registered workflow.

Figure 38: ACGT Portal – The visualization of the workflow in the list after registration

7. Log out from the portal

   Click the button Logout in the right upper corner of the header in order to log out from the portal.

8 Integration in the ACGT Portal

8.1 The "How to" section

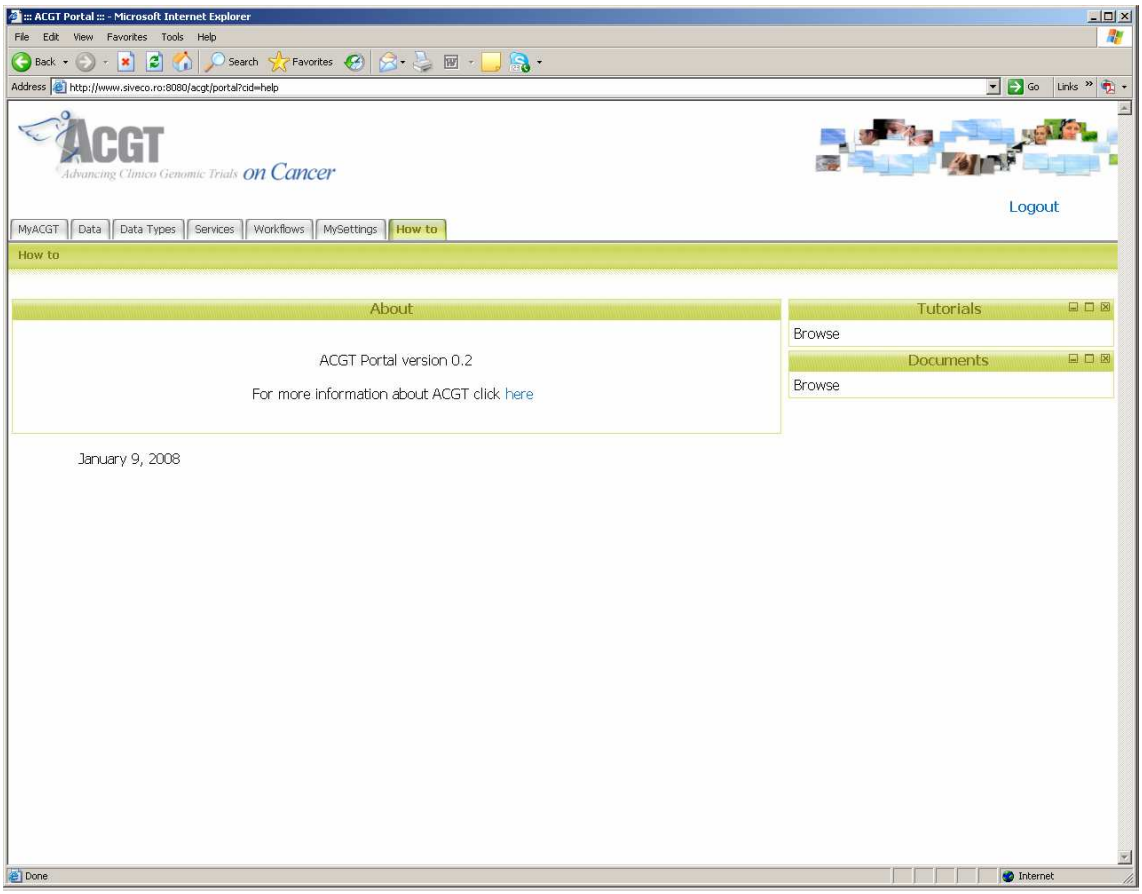


Figure 39: ACGT Portal – The *How to* section

All the electronic training materials are stored in the *How to* section. The training materials are split in two subsections:

- *Documents*

  This section contains the documents that are associated to the services exposed by the ACGT platform. Documents can be manuals, regulations, scientific papers, etc.

- *Tutorials*

  This section contains the tutorials that are associated to the services exposed by the ACGT platform. The online training demo module *How to register a workflow?* is listed in the *Tutorials* subsection.

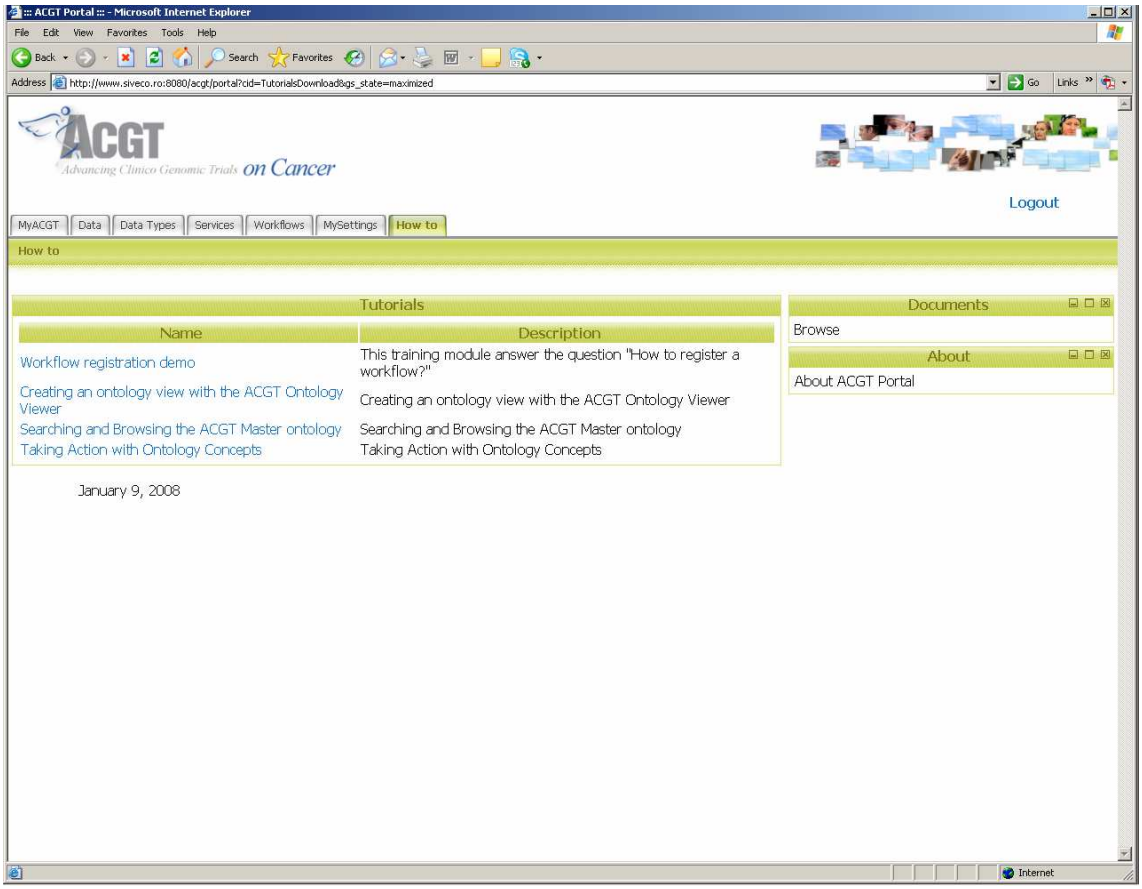


Figure 40: ACGT Portal – The *Tutorials* section

8.2 Section filtering

All the tutorials that corresponds to a specific section of the portal are filtered by the section name and can be directly accessed from the section, without being necessary to switch to the *How to* section. The list of tutorials referring a specific section of the portal can be accessed through a minimized portlet, which is displayed in every page of the section on the right column.

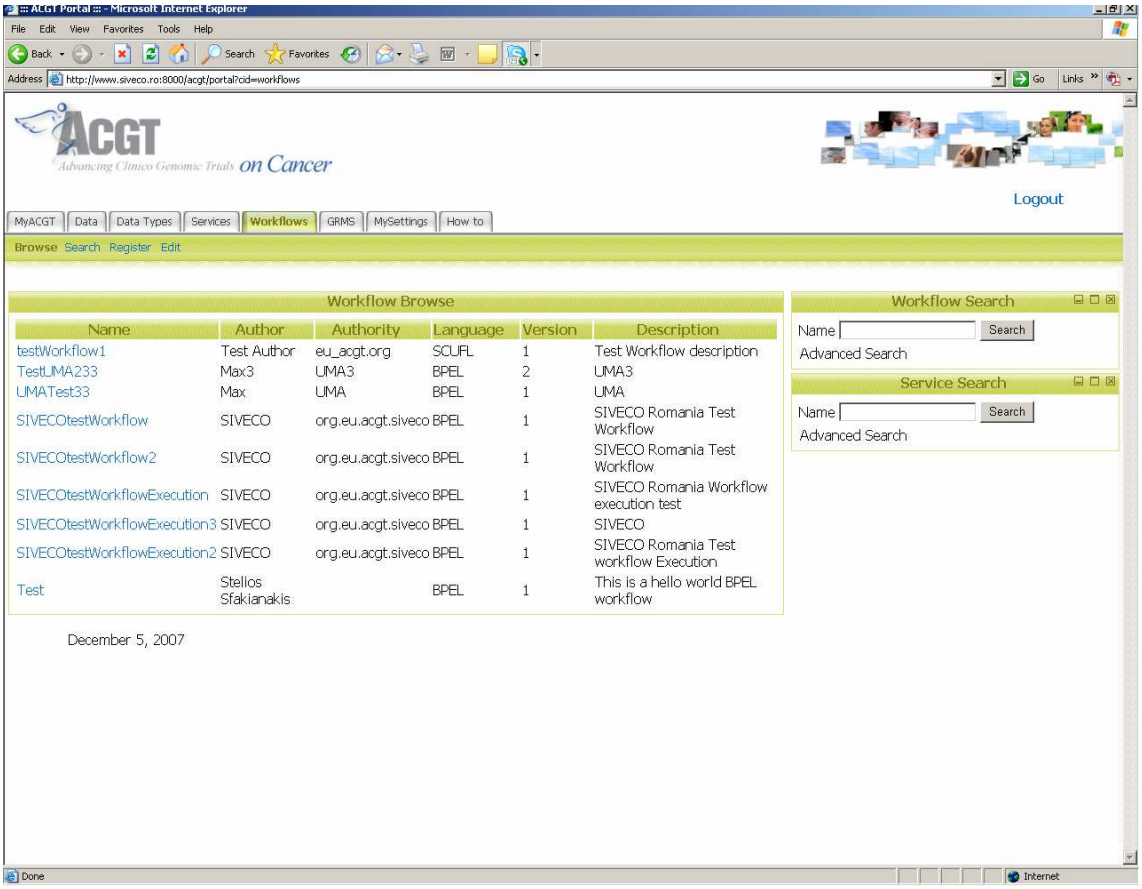


Figure 41: ACGT Portal – Publication of tutorials related to a specific section

8.3 Service/Workflow linking

The tutorial associated to specific service or workflow is also displayed as a link in the metadescription of the service, respectively workflow. When visualizing the corresponding item from the Metadata Repository

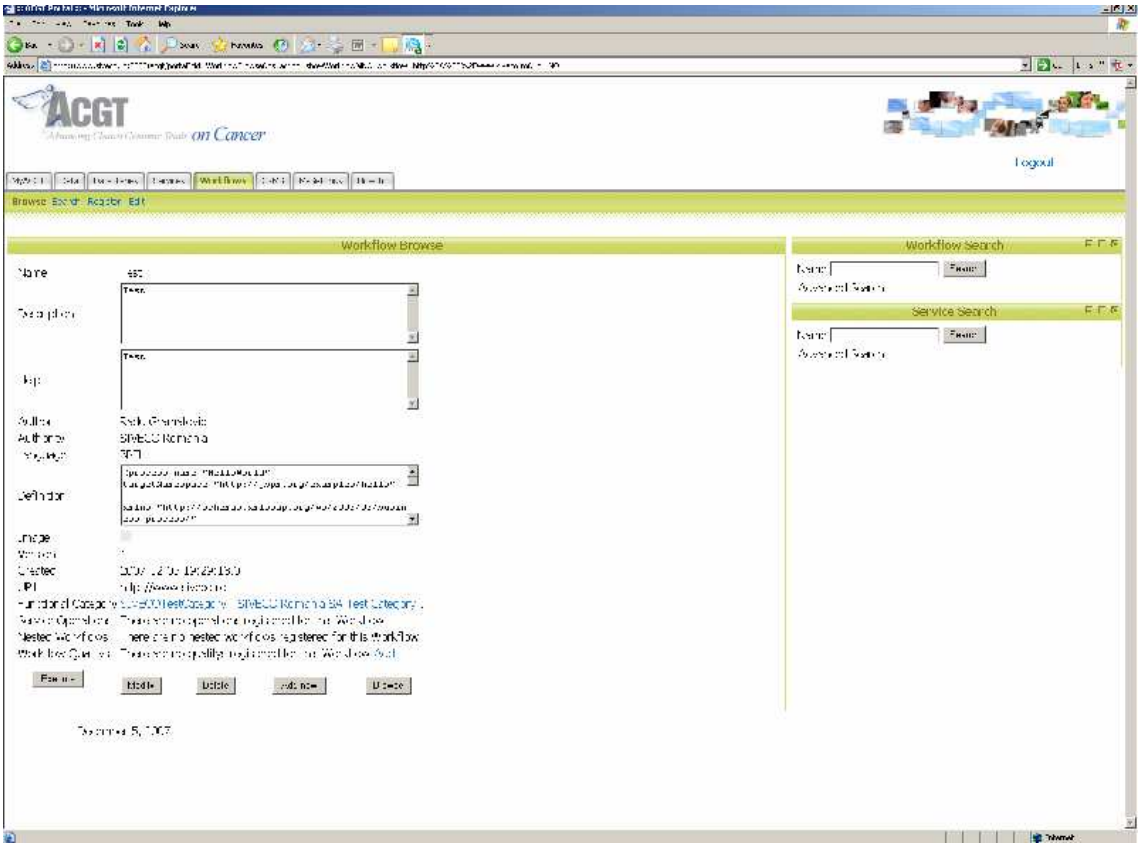


Figure 42: ACGT Portal – Linking the tutorial related to a specific service

9 The training module demo

The demo realised for the Workflow registration training module is publicly available through the ACGT Portal.

In order to access the demo, one has to connect to the portal following the link:

http://rd.siveco.ro/acgt/

Once in the welcome page of the portal a guest user is necessary to log in. Use:

User: *acgt_guest*
Password: *guest*

From the guest homepage the tab *HowTo* has to be selected.

From the *HowTo* portlet, the option *Browse* from the Tutorials box has to be selected.

From the list of tutorials, the *Workflow registration demo* has to be selected.

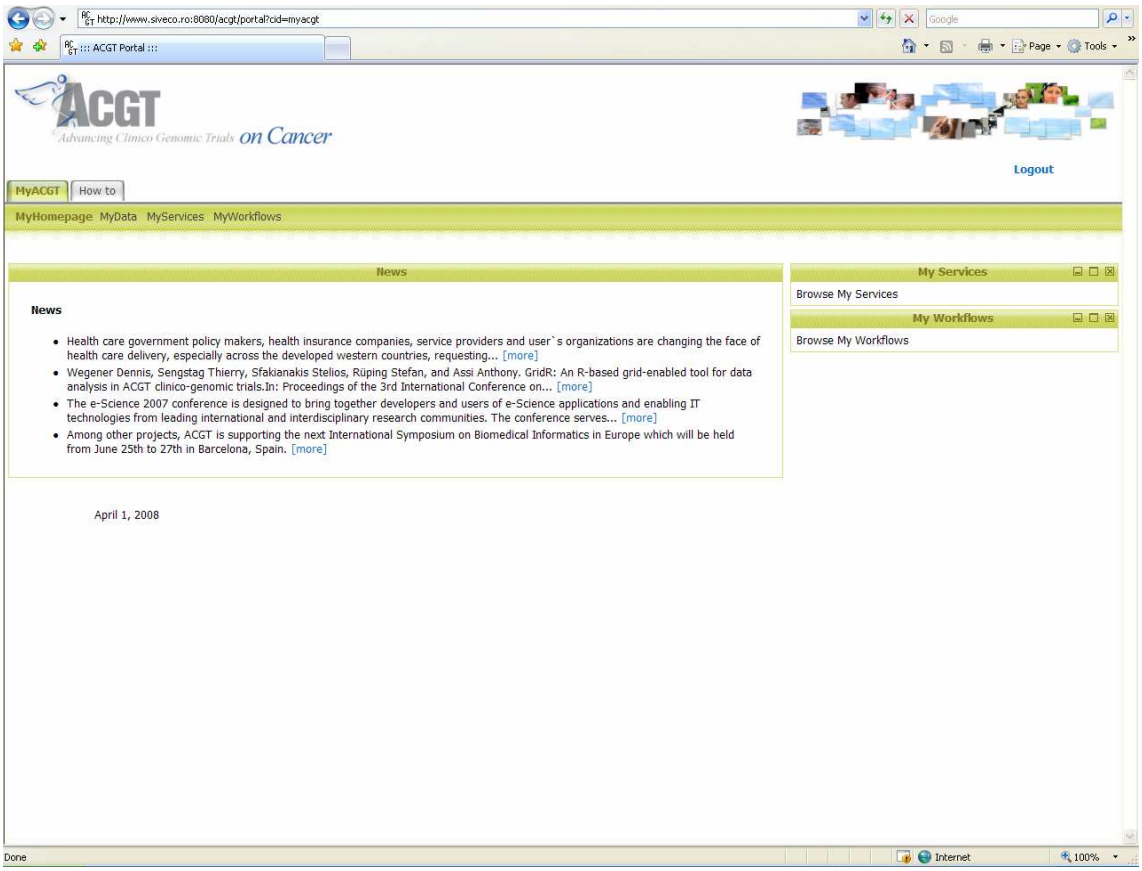


Figure 43: ACGT Portal – The acgt_guest homepage

PART 4

Other training modules

10 Ontology viewer tutorial

The Ontology viewer tool was developed by Biovista. The tool comes with a series of tutorials, some of them text-based, but also a video, which can be played from the ACGT Portal.

In the following, an extract from the Ontology viewer tutorial is presented.

10.1 Taking Action with Ontology Concepts

Follow these steps to complete this tutorial:

Step 1: Go to the Ontology viewer web page

Open a browser and type in the following URL: http://62.103.163.162:8080/OntologyViewer/

Your screen should look something like this:

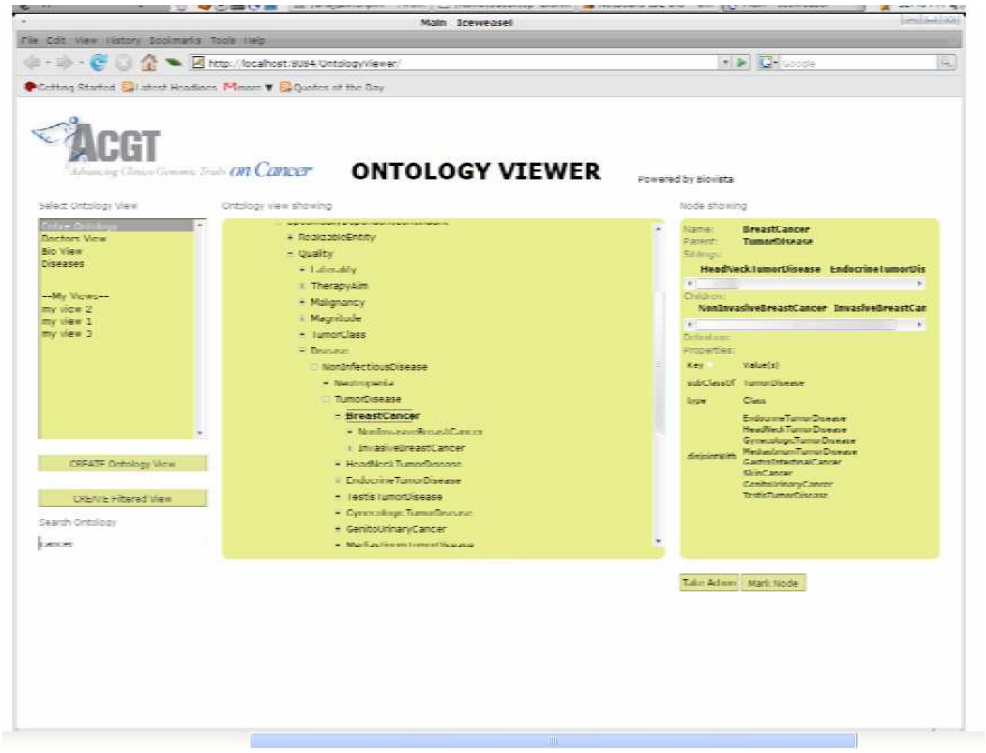


Figure 44: Ontology viewer homepage

Step 2: Search for your concept of interest

In the Search Ontology text box at the lower left hand side of the interface, type in 'BreastCancer' and press Enter.

Figure 45: Search ontology

The OV will now display the Entire Ontology tree zoomed on the concept BreastCancer (highlighted in the middle pane) and will also display all the concept data details in the right hand side of the interface.

Step 3: Take Action

Now that the OV displays your concept of interest you can access a number of concept-specific actions.

Click on the button Take Action on the lower right hand side of the interface.

Figure 46: Take action menu

From the drop-down menu select the Search Literature item.

The OV will now search the MEDLINE database and return publications on Breast Cancer and display the results in a pop-up window.

11 Victoria search engine within Magallanes client

Web-services are increasingly used for bioinformatics analysis. However, as the number of services grows, finding suitable services to process a given data becomes complex. Typically, service interfaces are annotated with metadata and shared in public repositories. Such annotations can vary in complexity; from simple text-descriptions to annotating with concepts from a formal ontology. However, in practice, service annotations are often limited to text-descriptions and some annotation of inputs and outputs (BioMOBY etc).

Victoria is a search engine library developed by University of Malage, able to use metadata from various service metadata repositories. Service discovery is supported by using user-specified keywords which are matched against service descriptions (matches can be exact or based on text-distance using Levenshtein's distance). Additionally, Victoria is able to automatically connect services into workflows based on user-specified input and output data type annotations which can solve the request. Workflows can be exported to the Taverna workflow system.

Victoria functionality is implemented as a proof-of-concept in the graphical client Magallanes. The library can easily be used in other applications.

In the following we present a step-by-step example of how to compose complex workflowa with Magallanes.

11.1 Composing Complex Workflows with Magallanes

The Workflow

For this example, we have selected the next workflow from IWWEM [http://ubio.bioinfo.cnio.es/biotools/IWWEM/workflowmanager.html]:

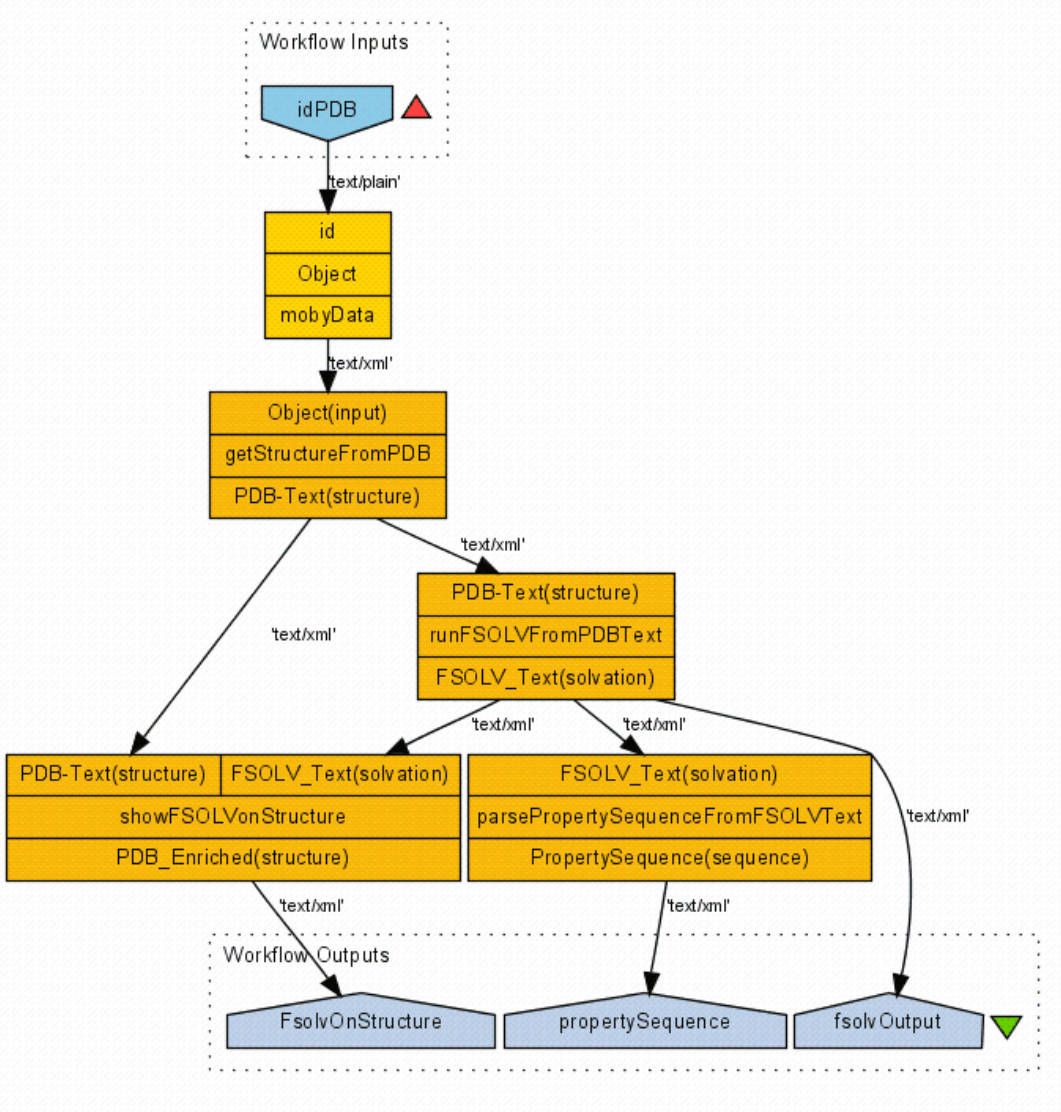


Figure 47: Complex workflow example

Author: Adam Hospital (adam@mmb.pcb.ub.es)

Uploader: Adam Hospital

License: CC Attribution-ShareAlike 3.0

Inputs:

- idPDB *(text/plain)*

Outputs:

- FsolvOnStructure *(text/plain)*
- propertySequence *(text/plain)*

- fsolvOutput *(text/plain)*

Description:

From a structure (PDB code), runs a fractional solvation analysis (FSOLV). It returns three different outputs:

- A FSOLV report with information about fractional solvation for each proteína residue.
- A Jmol 3D representation of this FSOLV report.
- A parsed Sequence with FSOLV report information, where N means NORMAL and O means OUTSIDE.

Steps (Web Services Used):

1. getStructurePDB Service

Getting the structure from the Protein Data Bank.

2. runFSOLVFromPDBText

Evaluates fractional solvation based on Linear Response Theory (LRT) method.

3. parsePropertySequenceFromFSOLV

Parses a FSOLVText report into a PropertySequence (where N means NORMAL and O means OUTSIDE).

4. showFSOLVonStructure

FSolv report 3D Jmol representation.

The Inputs/Outputs

We want to analyze the solvation of a protein structure, we only have the PDB's id and we want to get a Jmol 3D representation of that analysis.

We know that the ids are represented with the Object data type. Now we must search for the Object data type on Magallanes in order to add it as Workflow's source.

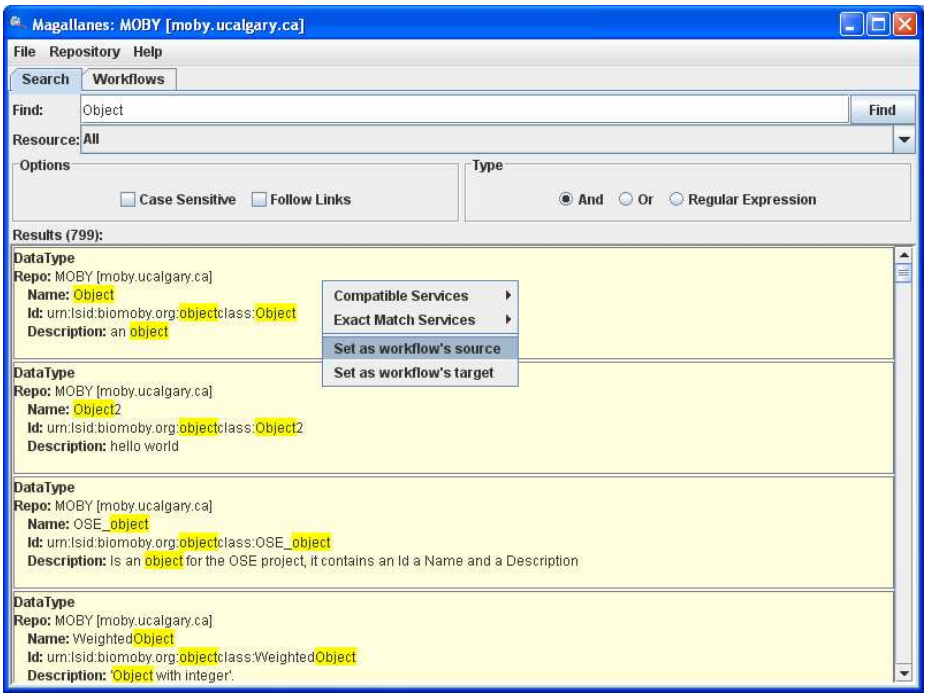


Figure 48: Search for object data types

If you try to search “Object” you will find the data type named Object, you must set it as workflow’s source by right clicking over it.

Magallanes will show you the workflow’s tab with the selected source.

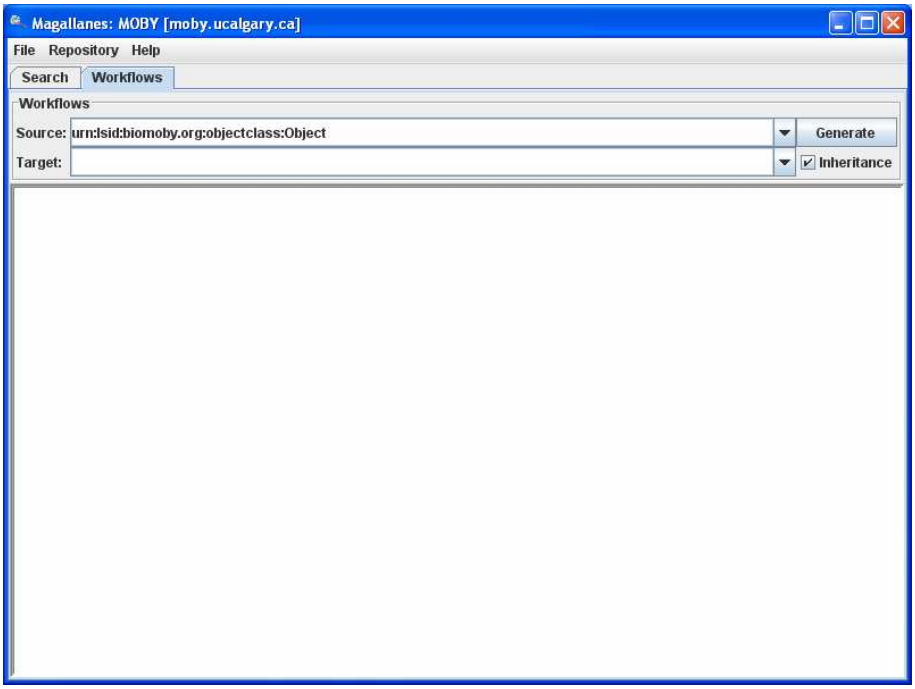


Figure 49: Workflows tab in Magallanes

Now you can go back to the Search tab in order to find the target data type. We want a Jmol representation then we will search jmol.

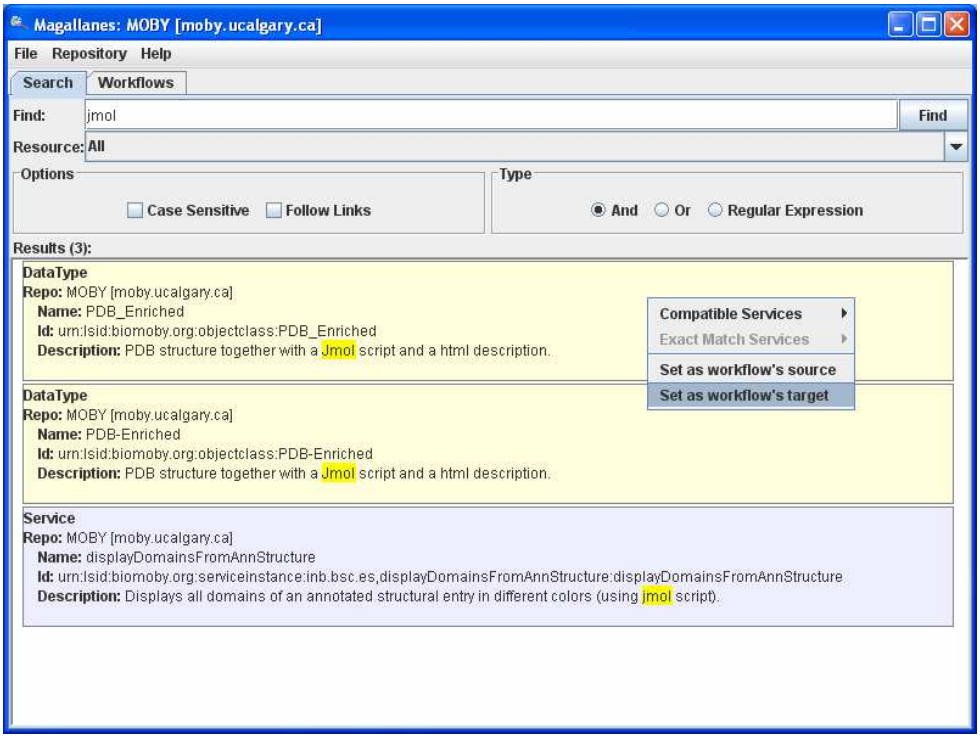


Figure 50: Search tab in Magallanes

We select PDB_Enriched as workflow's target.

Workflow generation

We can generate the workflow by clicking over the 'Generate' button.

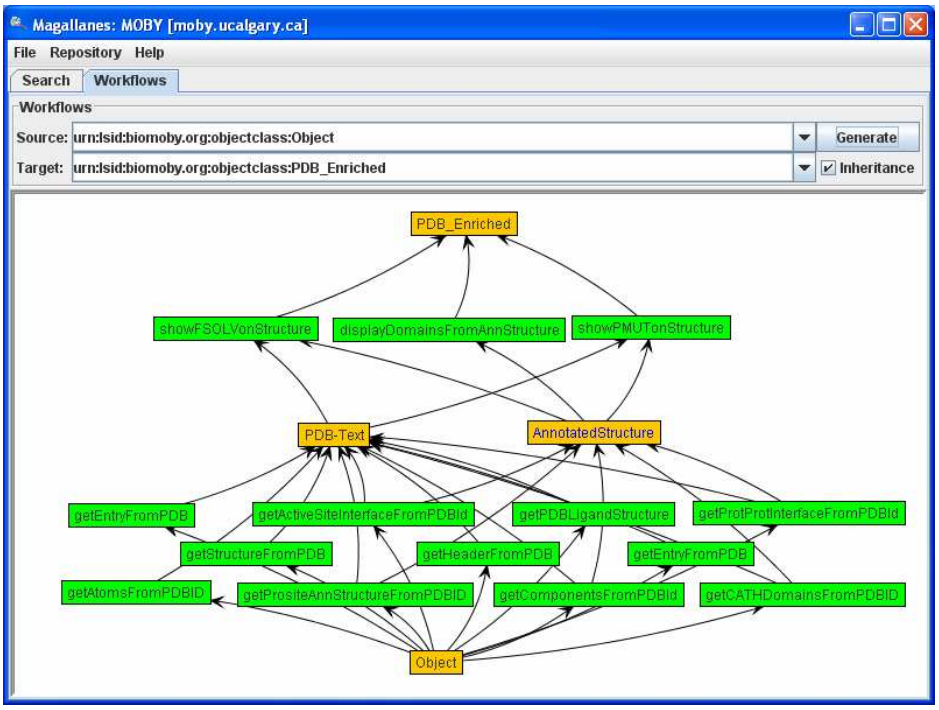


Figure 51: Workflow generated in Magallanes

As you can see we have a lot of services that receive Object. We must select a service that retrieves the PDB structure of the Object's id. If you keep the mouse over a service (green box) a popup message will be show with the service's description.

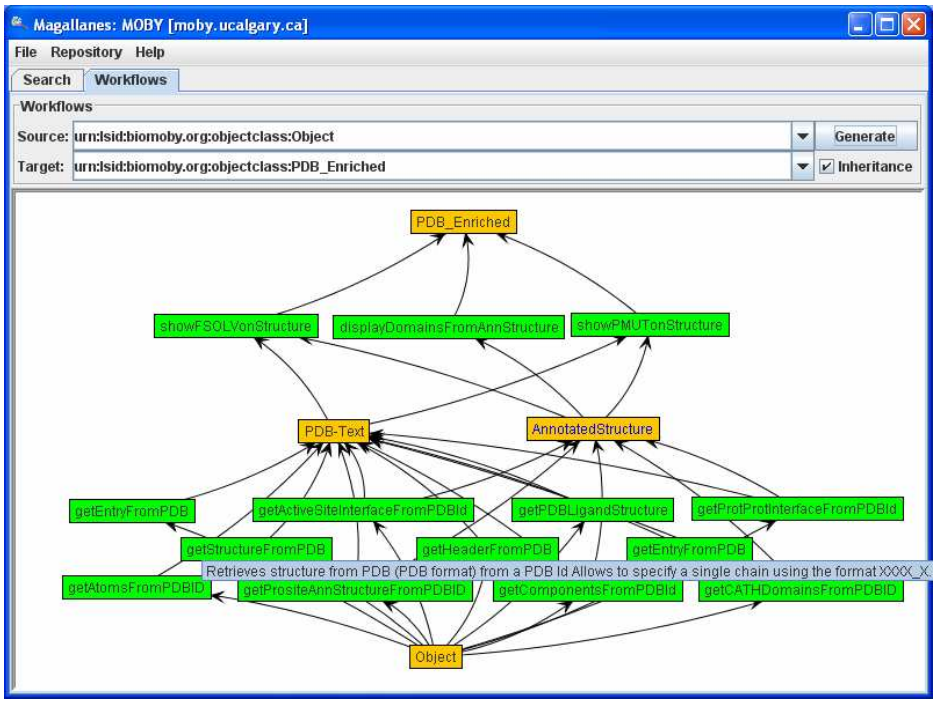


Figure 52: Pop-up service descriptions

To select the service you must double-click over it. We must select “getStructureFrom PDB”.

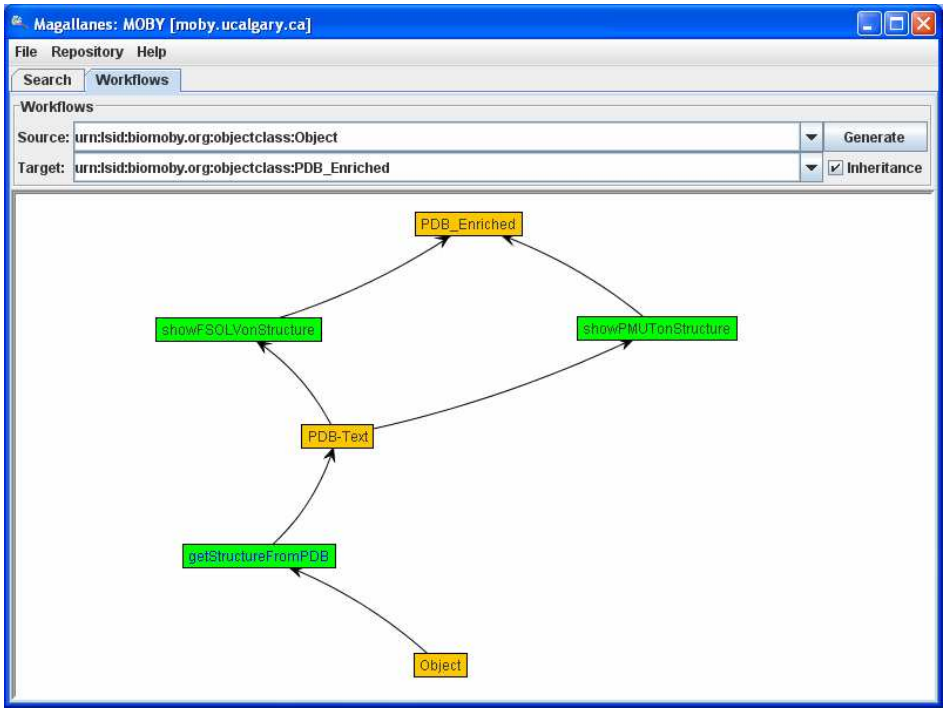


Figure 53: Select service

Do the same with the services that returns PDB_Enriched. As we want to analyze the fractional salvation, we will select showFSOLVonStructure.

Export to Taverna

We already have the pipeline without branches then we can export it to Taverna.

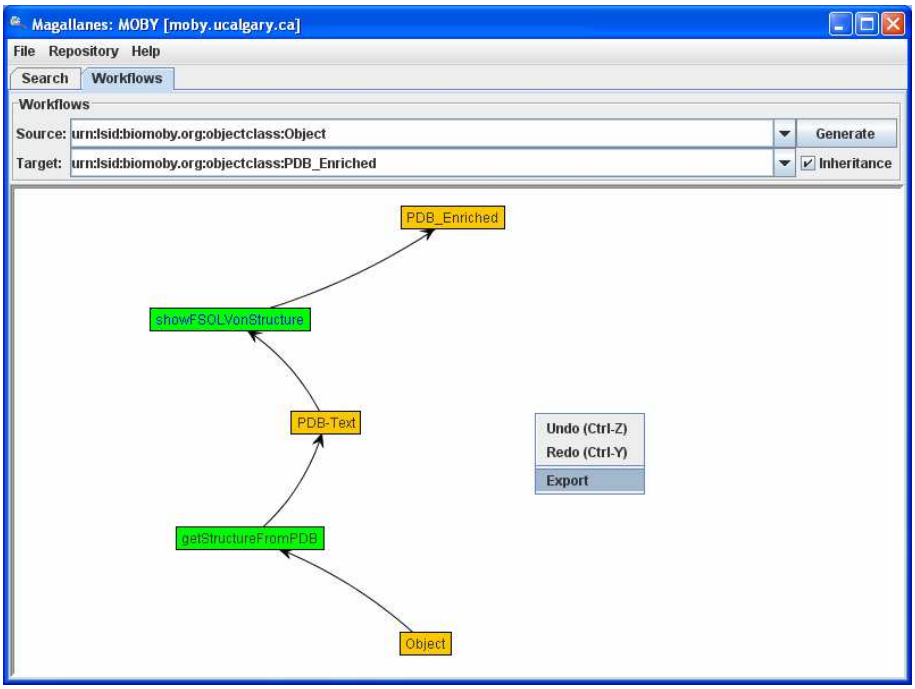


Figure 54: Export the workflow to Taverna

Now you can import the workflow on Taverna.

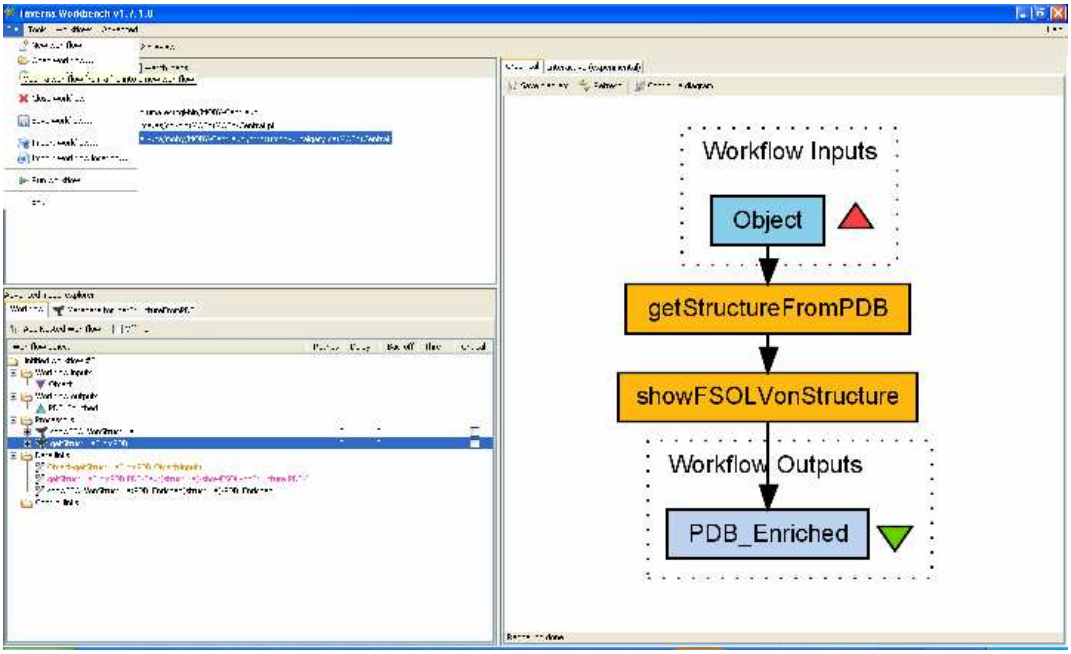


Figure 55: Taverna view (1)

By default Taverna doesn't show all the service's inputs. To enable it select 'All ports' and 'Top to Bottom' on the 'Configure diagram' menu.

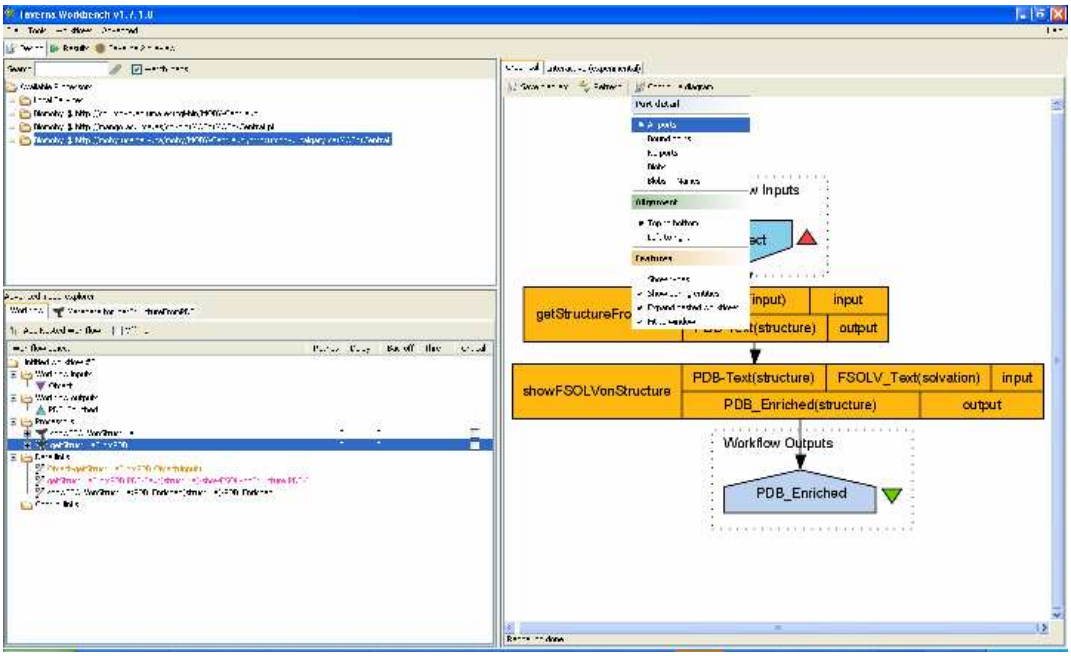


Figure 56: Taverna view (2)

Complete the Workflow

You can see that 'showFSOLVonStructure' must receive also a FSOLV_Text. We will try to obtain a little pipeline to go from PDB-Text (getStructureFromPDB's output) to the required FSOLV_Text.

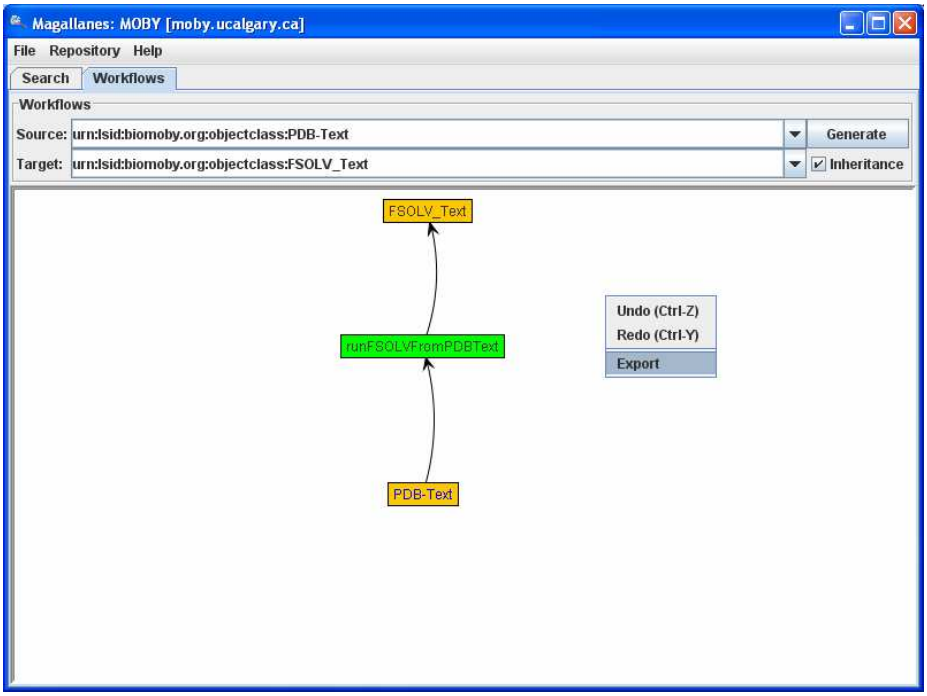


Figure 57: Workflow completion

Now export the workflow like the previous step and import it into Taverna as nested workflow.

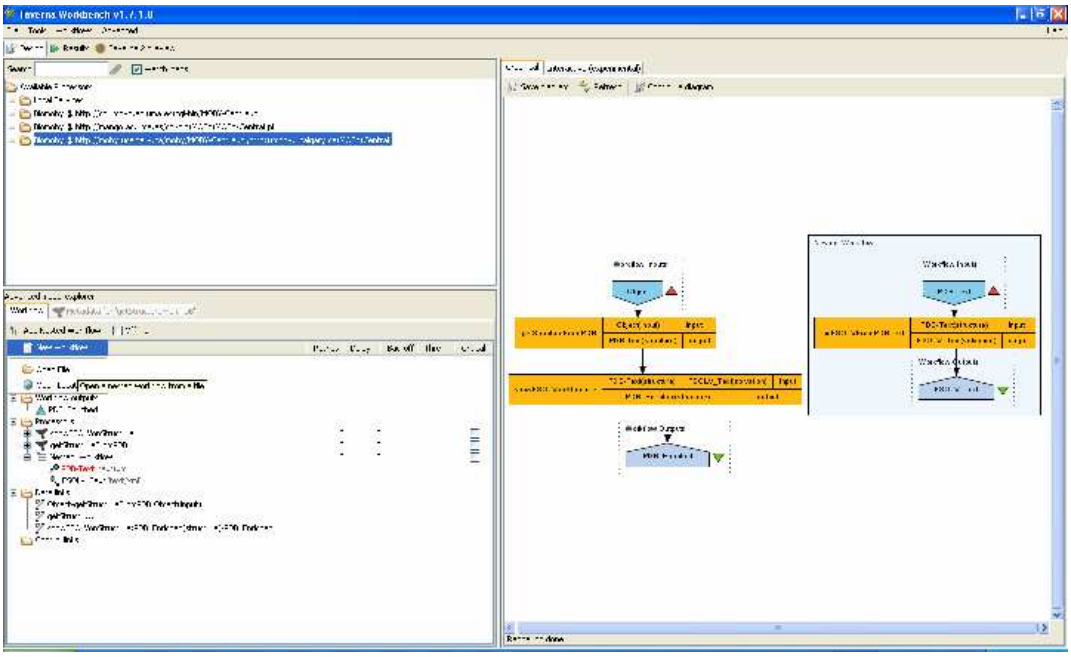


Figure 58: Taverna view (3)

To complete the Workflow we just link the original workflow with the nested one. You can do it by right clicking over the outputs on the left side tree and selecting the input with you want link it. You must link the getStructureFromPDB's output with the nested workflow's input.

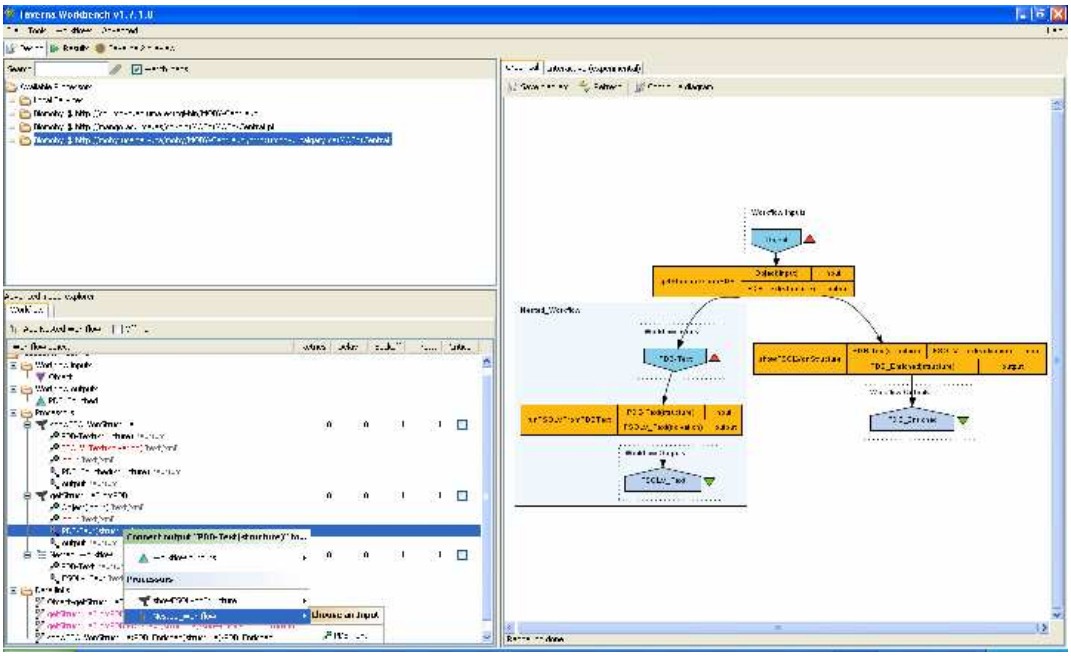


Figure 59: Taverna view (4)

You must link also the nested workflow's output with the showFSOLVonStructure's FSOLV-Text input.

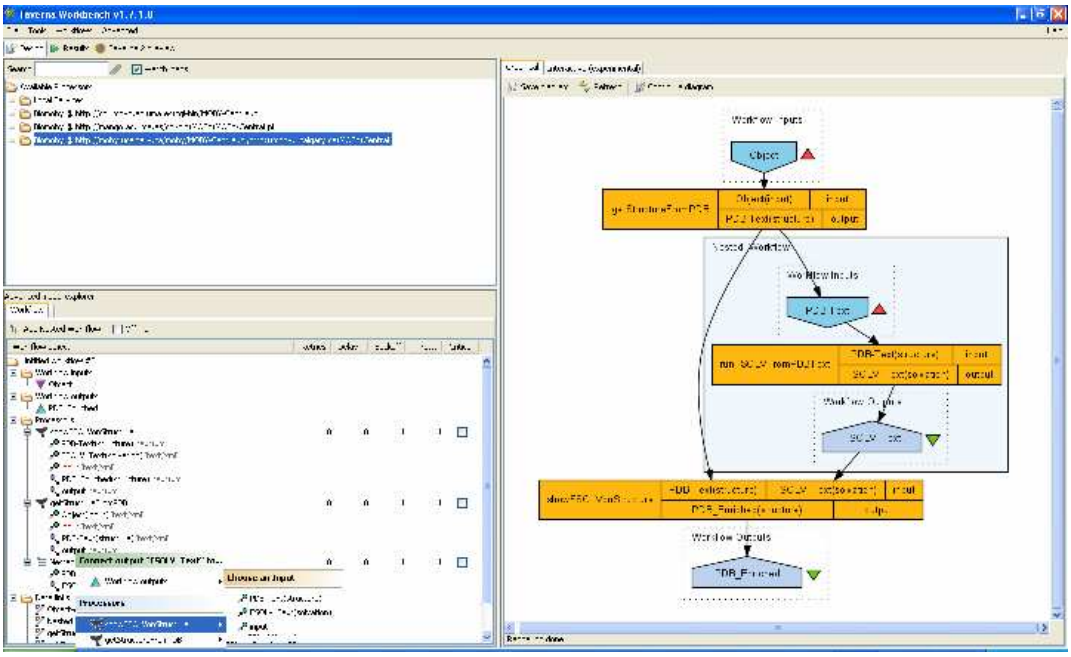


Figure 60: Taverna view (5)

Now you have the workflow and you can execute it. If you want to retrieve also a PropertySequence as the original workflow you only need to query Magallanes how to get a PropertySequence from FSOLV_Text and include the workflow as the previous.

12 jORCA tutorials

jORCA is a desktop client able to efficiently integrate different type of web-services repositories mapping metadata over a general definition to support scalable service discovery and to achieve flexible inter-communication between tools. jORCA manages repositories heterogeneity supported by the Modular-API that provides a uniform view of metadata (e.g. GRID-based, pure WSDL–services, BioMoby and others), making the integration of bioinformatics Web-Services easier.

jORCA allows the execution of different web-services invocation protocols and includes features to cover a broad range of user skills: drag-and-drop edition styles, easy incorporation of viewers for data types, custom favorite's section that allows the user to freely define a customized environment, on-line updated documentation, etc and it is fully configurable by means of default parameters in text-files (e.g. initial-links, look & feel, etc). Several methods are provided to help user in discovering tasks to identify the appropriated data-types or tools to solve specific problems: text boxes for fast keyword based searches, compatible web-services related to a type of data, advanced keyword searching including 'did you mean' Google-like methods to manage misspelling words and endowed with user profile learning capabilities. An automatic interface builder provides a uniform look for service parameters and invocation tools, making easier the pipelining of tools by I/O compatibility.

Advanced features such as mirroring or manual scheduling, asynchronous calls and execution tracing (editable log record of events) are also provided. jORCA is extensible by a programmatic Java API to incorporate new repositories and web-service invocation methods.

jORCA is present in the ACGT Portal within a series of video tutorials, which are accessible from the home page of the portal.

13 Data mining tools tutorial

This tutorial was produced by SIB as a training/evaluation mixt tutorial and was for the first time applied to internal end-users during the Vienna Consortium Meeting, in January 2009. This tutorial was developed by SIVECO into a video, which is available through the ACGT Portal.

The tutorial presents a step-by-step approach to using the ACGT platform for data mining, starting from the creation of a user account to the creation of a useful result.

13.1 User account creation

The ACGT environment supports a mechanism which allows the identity of users belonging to specific virtual organizations (VOs) to be verified. This is done via so-called registration and certification authorities.

Summarizing the procedure:

1. The new user requests the right to use ACGT by filling a form.
2. The identify of the user is verified by the registration authority (usually a person in the same institution as the user). The registration authority also makes sure that all legal documents are signed by the user.
3. The user is allowed to generate a key and a certificate which will be used to prove his/her identity. The key and certificate are "signed" (with a key) by the registration authority and protected by a password that the user chooses.
4. The user then creates an ACGT portal account (a username/password pair).
5. The user associates his/her portal account with his/her certificate (which proves to the ACGT environment that the portal account is truly associated to the person that owns the certificate). This is done through the delegation process.[2]

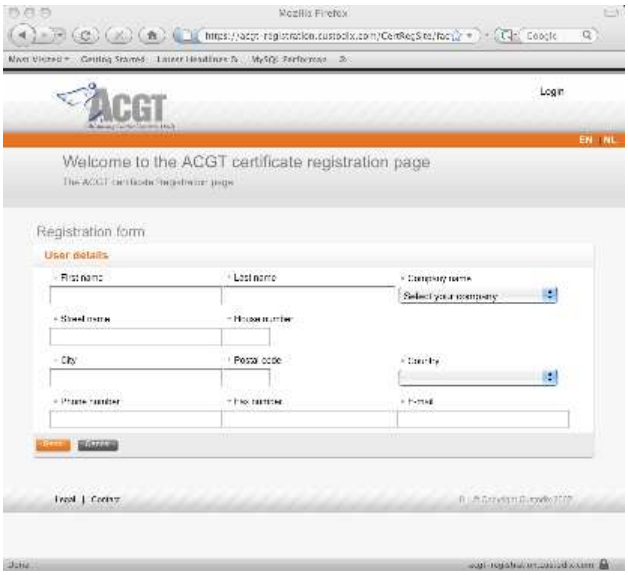
In practice the steps are as follow:

Step 1:

The user registers on the certificate registration page:

https://acgt-registration.custodix.com/CertRegSite/

[2] More precisely, in the delegation process, the user delegates to the portal the right to use his/her credentials (stored securely on a third-party server) to act on his/her behalf. Without the delegation process, each component of the ACGT infrastructure would repeatedly ask the user to prove his/her identity by providing his/her credentials, which would be a significant annoyance.



Step 2:

The appropriate registration authority (RA) reviews the application, contacts the user for signing documents (if necessary) and accepts/denies the registration. Upon acceptance, the RA sends the user an **e-mail with a link**, which contains the RA key to use to generate the certificate.

Step 3:

Clicking on the link brings you to an applet that creates a key and a certificate, which are protected by a **password** chosen by the user.

The key and certificate are stored in the home directory of the user in the directory *.globus*.

Under Windows: C:\Documents and Settings\username\.globus

Step 4:

The user creates an account on the portal home page (**http://rd.siveco.ro/acgt/portal**) by clicking on the **Create new account** link (check that the "ACGT portal" option is selected in the upper left corner). The account creation will be confirmed by **e-mail**.

Figure 61: User account creation in the ACGT Portal

Step 5:

The last step is to create the delegation rule between the portal account and the user credentials. Go to the portal home page (**http://rd.siveco.ro/acgt/portal**) and select the "ACGT MyProxy tool" on the left (blue arrow on the figure above). This will start an applet which can be used for registration.

Figure 62: ACGT MyProxy tool

With the configure option of the applet, advanced parameters can be setup, for instance if the user certificate and key have not be stored in the default location. Besides, the delegation must be done to a different server (Proxy Hostname) if the user wishes to access the development environment of the ACGT project.

Stable ACGT environment: **myproxy.custodix.com**

Development ACGT: acgt-node2.custodix.com

Clicking then on the "**Delegate credentials**" button opens a window in which the user name used to create the credentials has to be provided, as well as a new password[3] which is used

[3] For simplicity, this password can be identical to the one used to create the portal account.

by the machine which stores the mapping between portal account and credentials. The last password is the one used on Step 3 to protect the credentials.

Step 6:

The VO manager can add a user to the appropriate VO.

Additional configuration step

Some elements of the ACGT interface are coded as Java "applets" instead of "portlets". Applets require an other step of certification to make sure that they are not unduly using the resources of the local machine.

The user can authorize the execution of ACGT applets by adding the following two certificates to the browser:

https://acgt.custodix.com/certificates/acgtca.crt

https://acgt.custodix.com/certificates/ca.cer

In Firefox, the certificates can be added through the menu Tools, then Options…, Advanced tab, View Certificates, Authorities, then using the Import… button to load the two files above.

Under Windows, double-clicking the certificates should install them automatically.

At this point the user can start using the ACGT environment.

Figure 63: The ACGT Portal before logging in

13.2 User account creation

There are two main mechanisms to feed data into the tools which are part of the ACGT data mining environment: user files and databases.

The ACGT DMS

The ACGT data management system (DMS) is a grid-based system which allows storing data (files) in a secure and remotely accessible way.

Database access through mediator queries

A tutorial for this mechanism remains to be setup.

Steps to describe:

1. Database integration (i.e. I have a MySQL/PostgreSQL/xxxQL database on my computer and want to be able to use them in a workflow, what should I do?)
2. Semantic mapping
3. Query building

Uploading a file to the ACGT DMS

We want to upload a clinical-trial demographics file to the ACGT DMS.

The file to be used for the exercise can be retrieved here:

http://www.isrec.isb-sib.ch/~sengstag/MCMP_PartB_public/demo_public.csv

Download the file to your local file system.

To upload the file, click on the “Data” tab of the ACGT portal. (Alternatively, click on “My Data” under the MyACGT tab.) The DMS editor should appear.

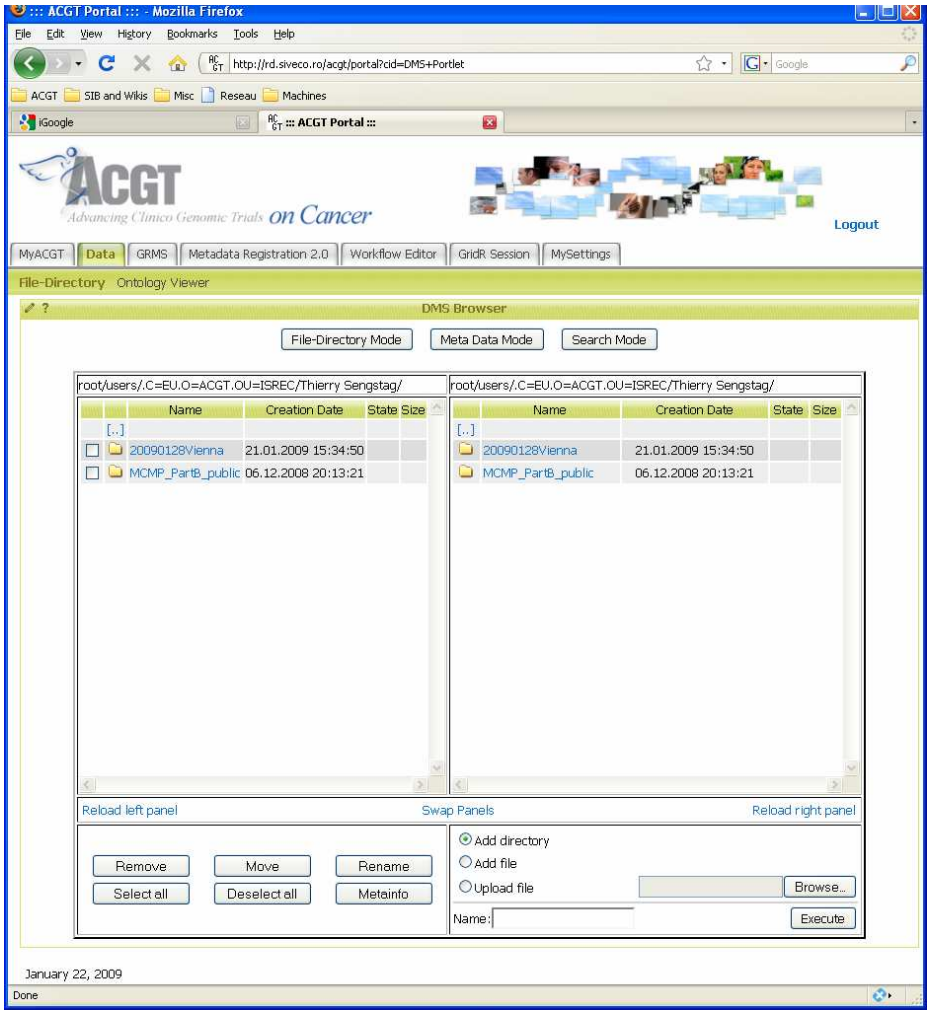


Figure 64: The ACGT DMS portlet

The left panel of the DMS browser can be used to navigate the directory structure, while the right panel will display the contents of files, depending on the context.

Create a directory:

- In the lower right panel, make sure the “Add directory” button is selected
- Write the name of the directory to be created in the “Name” text field
- Click on “Execute”

The directory should now be displayed in the left panel. Click on its name (in the left panel) to open it and see its contents.

Uploading a file will store it in the directory currently open in the left panel.

Uploading a file:

- Navigate to the directory in which the file has to be stored
- In the lower right panel select the “Upload file” button
- In the “Name” text field, write down the name of the file as it should be stored in the DMS

- Click on the Browse button to select the file to upload.
- Then click on “Execute”.

The file is then transferred to the ACGT DMS.

Once uploaded, the file contents can be visualized (in the right panel) by clicking on its name in the left panel.

Files and directories stored on the DMS can be deleted by checking the box near the elements to be removed on the left panel, then by clicking on the “Remove” button in the lower left panel. Directories must be emptied before they can be removed.

13.3 The GridR portlet

The API documentation for the GridR environment is given on the ACGT wiki:

http://wiki.healthgrid.org/ACGT:GridR

In the present exercise, the user wishes to plot the survival curves for the patients listed in the file we uploaded in the previous section, stratified by treatment.

After logging into the ACGT portal, the user can create an interactive GridR session as follows:

- Open GridRSession portlet
- Click on Connect
  This creates the communication with the specified GridR server.
- Click on New Session
  This creates a GridR session in which the user can work. The user can have several parallel GridR sessions.
- Import a file from the DMS to the local session (using one of the GridR API functions):

```
demo = grid.AcgtReadFromDms(18571, "demo_public.csv")
```

(The number here should be the identifier of the file on the ACGT DMS. As this identifier is currently not available from the DMS interface, a small hack using the workflow editor is needed. This will be fixed in future releases of the DMS interface.)

- Click on Reload (top of the tree-view for the working dir), the file will appear in the GridR environment after a short while.
- Paste script commands in the command filed:

```
rownames(demo) = as.character(demo[,2])

library(survival)

treat = demo[,"treatment"]

time = demo[,"t.rfs"]

event = demo[,"e.rfs"]

my.col = ifelse(treat == "none", 1,2)

fit = survfit(Surv(time,event) ~ treat)
```

```
plot(fit, col=my.col)
```

- Click on CreatePlot
- Click on Reload - a png file should appear in the directory listing
- Click on the png file – the picture will be displayed in the applet.

Tip: Sometimes you have to enter a "blank" command to see the output of an R statement. We are already working on this.

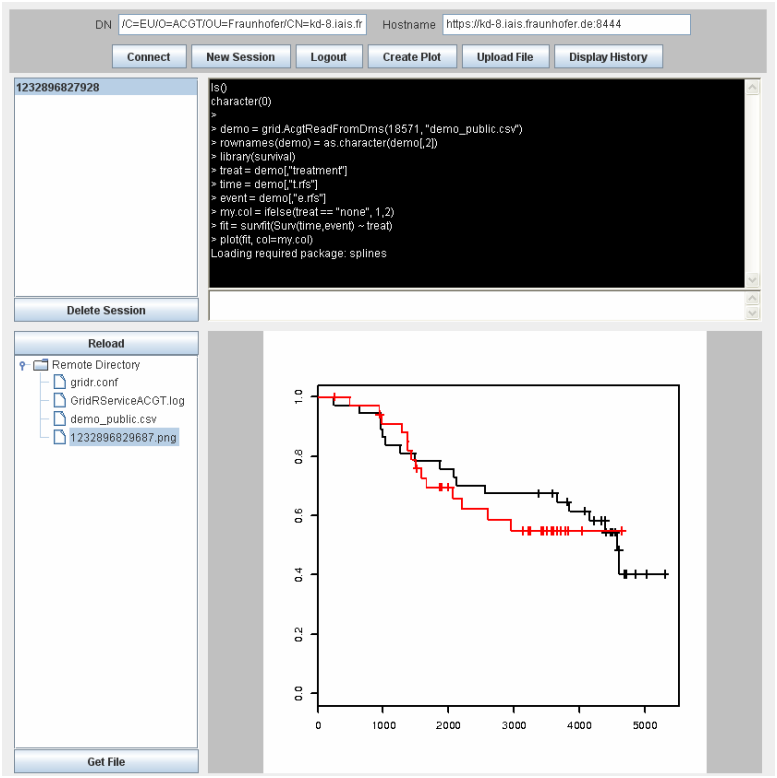


Figure 65: The ACGT GridR portlet

13.4 Creation of a GridR-based data mining workflow

Register a GridR script as a service usable in a workflow

Definition: A *GridR script* is an R script which *may* call the GridR functions defined to execute R on a grid (e.g. the ACGT Grid). A plain normal R script which is *not* using the GridR extensions is also a *valid* GridR script.

Background information

In a workflow environment, GridR scripts are embedded in “GridR components” (also known as “GridR modules”). The latter provide an input and output mechanism which lets the GridR scripts interact with other workflow components.

A GridR component is thus a normal GridR script, with inputs and outputs; for instance, the following GridR component has two inputs and one output.

Figure 66: A GridR component

The number and type of inputs and outputs are defined when the script is registered as a service (see next section of this tutorial).

Internally, workflow components use files to transfer information as input and output.

A GridR script registered as a component usable in a workflow can access its inputs and outputs through two R objects (lists), ***gridr.input*** and ***gridr.output***, which are automatically created in the GridR session that will execute the user script.

Inputs

Each element of the list gridr.input is a data frame resulting from reading the file at the corresponding input position, more precisely gridr.input[[i]] contains the data read with the R command read.csv() (with all default parameters) on the file attached to the i^{th} input connector.

Outputs

The names attached to the output connectors of GridR components (which are defined at the time of registration of the component, see below) are accessible through the elements of the list gridr.output; thus gridr.output[[i]] is a string which can be used as a file name in a R *write* command. A typical use of this mechanism is thus as follows:

```
… R script defining the contents of the data frame x …

write.csv(x, file=gridr.output[[i]])
```

Scenario script

In this simple scenario the user wishes to assess the effect of treatment on patient survival using the demographics data contained in the file we have uploaded to the DMS on the previous step.

The GridR script is as follows:

```
demo = gridr.input[[1]]

pdf(file = gridr.output[[1]], paper="a4r", width=28./2.54,
    height=18./2.54) # an alternative usage of gridr.output[[]]

rownames(demo) = as.character(demo[,2])

library(survival)

treat = demo[,"treatment"]
```

```
time  = demo[,"t.rfs"]

event = demo[,"e.rfs"]

my.col = ifelse(treat == "none", 1,2)

fit = survfit(Surv(time,event) ~ treat)

plot(fit, col=my.col)

dev.off()
```

Verify the column headers with the DMS portal.

Registration of script as GridR service

The script above can be registered as a GridR service in the ACGT service repository (the "repo"). The latter can be accessed through the ACGT portal under the tab "Metadata Registration 2.0".

The services registered into the ACGT infrastructure are visible in the "Services" sub-tab of the "Metadata Registration 2.0" tab.

The services themselves are grouped in categories which can with see by expanding the "Service" tree (left panel).

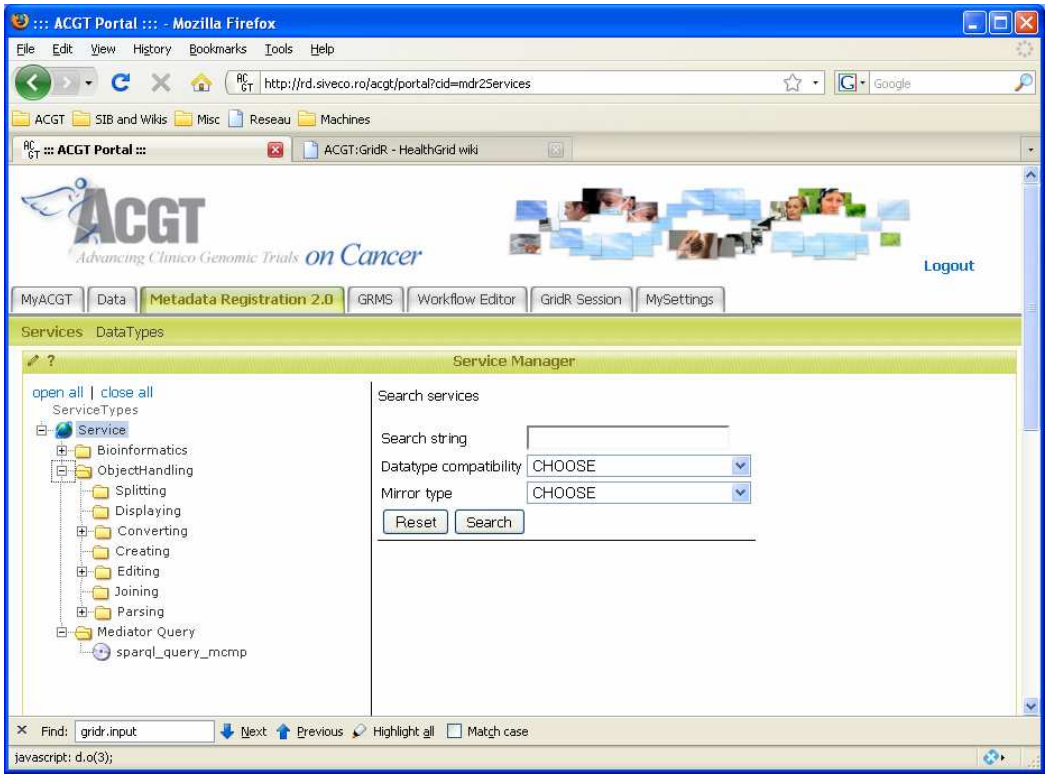


Figure 67: ACGT services tree

We are going to add our script to the “Bioinformatics” category. Expand it to view the currently existing services.

In this interface:

- Selecting a specific service shows the details associated to the service, including the specification of its interface.
- Selecting a category (clicking on the folder containing a set of related services) provides link to four broad categories of services:
    - Service type (“Add service type”)
    - Generic service (“Add Service”)
    - GridR service (“Add RScript”)
    - Mediator queries (“Add mediator query”)

After expanding the “Bioinformatics” category in the left panel, locate the “MCMP_illumina_filter_annot” service and click on it. The right panel, then shows how the interface of that GridR service was defined (in the section “RScript parameters”), while the script itself is described in the “Script data” section at the bottom.

To register the script defined above, click on the “Add RScript” button.

Choose a name for the script and add some documentation.

Then define the inputs and outputs for the scripts:

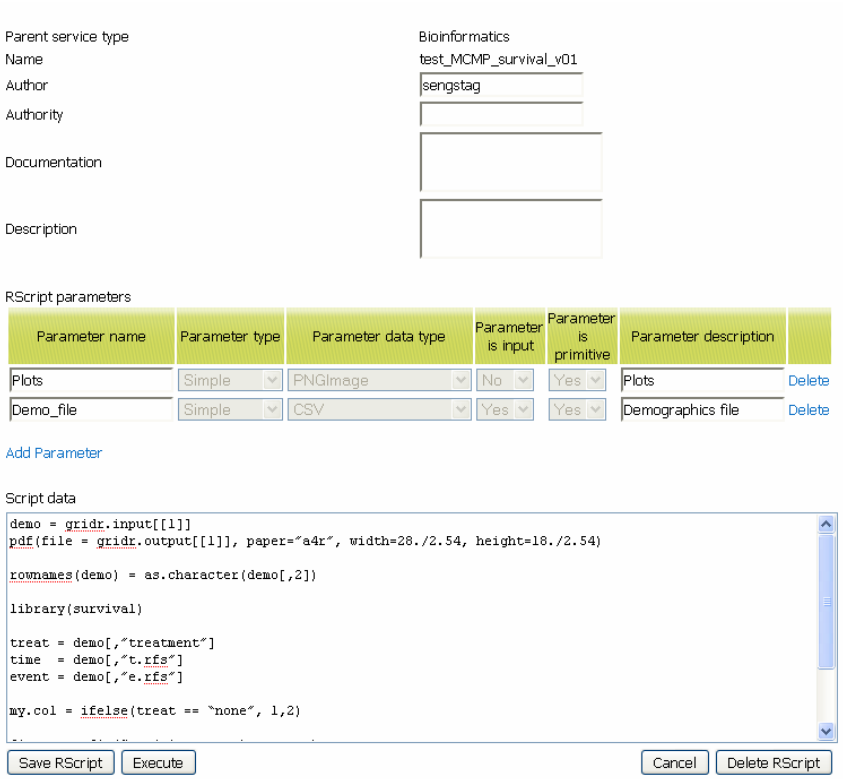


Figure 68: ACGT RScript definition

Parameter names are arbitrary strings which are used internally to exchange data.

Finally save the script with the button at the bottom of the input form.

The GridR component is now registered and ready to be used in a workflow.

It should be noted that with the present version of the service repository, all services are public.

Create a workflow using a GridR component

Open the workflow editor by clicking on the “Workflow Editor” tab.

By default, the workflow editor shows an empty area which is ready to accept a new workflow.

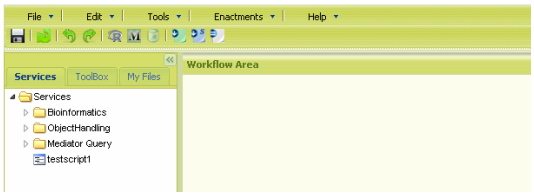


Figure 69: ACGT Workflow editor commands

The workflow editor commands are in the upper-left part of the window.

We want to connect the demographics file we have uploaded to the script we have just registered.

File insertion

Click on the “My Files” tab in the workflow editor toolbox in the left panel. The tree shows the files available in the DMS. Locate the demographics file, and drag and drop it to the work area (central panel).

GridR script insertion

Regular services are inserted in the workflow by drag-and-dropping them to the work area from the “Services” tab in the left panel.

However GridR-based services are internally treated differently from regular services (the same is true for Mediator queries and access to BASE databases). To insert a GridR service click on the corresponding icon from the upper toolbar:

Figure 70: GridR icon from the workflow editor toolbar

Select the script in the list that appears on the screen and “add” it to the workflow.

Then connect the file to the input of the workflow…

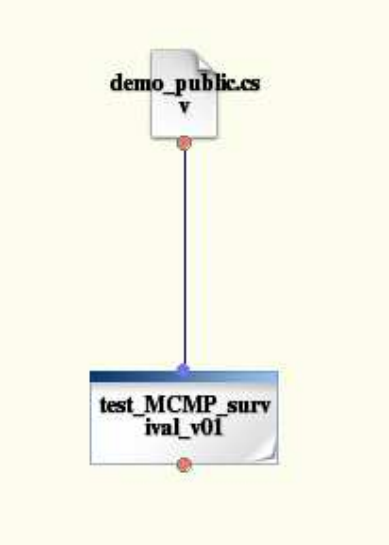


Figure 71: Service connection into a workflow

… and save the workflow (File / Save as…) with a name that describes its global functionality.

At this point we have finished defining the workflow. It must now be made executable, which is achieved by clicking on the “Deploy” command in the “Tools” menu.

A workflow is only visible to its owner unless it is explicitly *published*. This is achieved through to “Tools” menu.

13.5 Workflow execution

Execution from the workflow editor

The workflow can be executed directly from the workflow editor, after it has been deployed: Select the “Run” option of the “Tools” menu.

A window opens to let the user input the parameters of the run. With the workflow defined above no input parameters were required, so we just proceed.

The execution of the workflow is monitored in the “Status” panel at the bottom of the workflow editor. In addition active workflow elements can be seen blinking in the editor.

Once the execution of a workflow has started, it is running in background. The user can leave the session and the process continues. The status of a specific running workflow can be watched by selecting the corresponding execution “process” in from the list that pops-up after selecting “Open…” in the “Enactments” menu.

Execution from the service repository

A predefined workflow can be executed directly from the service repository:

Select the workflow from the service tree, then click on the “Execute” action in the “Available operations” section of the right panel. (If the selected service is not a workflow other options are available on the right panel; such services are usually not runnable in standalone mode.)

Once the “Execute” action is triggered a form appears, letting the user fill in the input fields for the workflow, which execution is actually started by clicking on the “Execute operation” button at the bottom.

Workflow execution output

Each GridR component in a workflow creates its own working directory in the DMS, in which temporary files are created. The results of a GridR-based workflow execution can thus be found in the GridR directories associated to the scripts of interest.

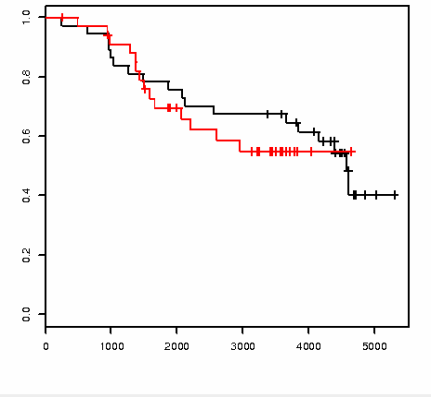


Figure 72: Service connection into a workflow