mesh

Multimedia sEmantic Syndication for enHanced news services

Information Society Technologies

...A view to the future of news

THE NEWS TODAY

In our days we are confronted with vast amounts of information commonly referred to as *news*. News about all aspects of our everyday lives are nowadays accessible to all corners of the world. But how easy is it for anyone to navigate this flood of information and what opportunities are there to get an objective view of controversial events, at national or international level?

Was the latest war an invasion or a liberation?

Were the latest elections a grand victory or the result of an unfair election system?

Our era of knowledge should provide for methods of understanding the meaning of *news*. Contemporary methods should be able to organise news in a semantic way that would allow the reader to have a complete overview of all similar and conflicting views, being also able to filter information according to personal preferences and interests. How can this be made possible?

news have to be understood by fully or semi-automatic mechanisms. This means that a news item (in any multimedia form; text, image or video) has to be analysed and categorised (i.e. annotated) according to its contents.

news consumers also have to be understood. This involves profiling of individuals in a structured manner and constantly updating these profiles through personally provided preferences but also through automatic understanding of their needs and interests by monitoring their requests and habits.

news items will then have to be matched to the readers interests and requests, by reasoning what news would be preferred by which reader and in which way.

news have to be delivered in an effective way. Personalised multimedia summaries can be a basic means of navigation into the full set of information, while items referring to the same subject will have to be shown in parallel to make critical reviews possible.

the source of information will also have to be understood and profiled if the reader is to be assisted in forming an objective view of actual events. Structured information will have to be provided for the source helping the reader to understand whether the news provided could be biased and towards which side. Credibility will also have to be measured in such a way.

In our era **mobility** is also a significant aspect to be taken into account. With mobile devices being used more and more both for the production and consumption of news, special technologies need to be advanced into a framework that allows effective inclusion of mobile prosumers.

The MESH project was initiated with the vision to integrate the above needs into a setting that would bring the world of news closer to knowledge-enabled services. Twelve different organisations with expertise in all these diverse fields have joined forces to make ***personal navigation in the world of news*** a reality.

A VIEW TO THE FUTURE OF NEWS

A number of application scenarios have been authored in the format of storylines which envision the future access to news through technologies planned to be developed in the framework of the MESH project.

Three broad categories of users have been identified as the targets for the planned MESH platform:

Personal user: An individual everyday news consumer who may belong to any profession or business activity and wishes to use the MESH platform to have access to everyday breaking or archived news in an easy and effective way.

Professional user: A professional who works in the area of Media and would like to use the MESH platform to make his/her work easier and more effective. This category may include journalists, photographers, reporters or any other employee in the sector of Media.

Business user: A professional associated to the world of Media in need to build a business using the MESH platform. This business will relate to models of electronic news services as a focused news agency or a thematic portal which will make use of the MESH technologies to provide intelligent content services.

Six prototype persons were identified in this process to represent the needs of these groups of users, in terms of their goals and personal characteristics. They act as 'stand-ins' for real users and help guide decisions about functionality and design. Through these persons, the motivations, expectations and goals are becoming more familiar so that the MESH consortium can reach a common understanding of the targeted product.

These persons were brought to life by giving them names and personalities. Even though fictitious, they are based on knowledge of real users and significant involvement of the consortium Media partners ensured that they represent end users rather than the opinion of the person authoring the storyline.

The following scenarios formed the basis for the project requirements recognition process. Through these stories the relevant system capabilities were identified and utilised to build questionnaires and guide group discussions in order to determine the importance of each feature, so that priorities could be set and the most important use cases could be described.

However, these storylines can also be seen as the vision of MESH for the world of media in the future. Technologies developed in the framework of this project and probably its successors aim to pave the way leading towards the days when these stories will be seen as everyday life.

So lets meet the future users of MESH...

Ilias Bonn – a travelling journalist

For three years Ilias Bonn has been travelling back and forth between Thessaloniki in Greece and Cologne in Germany. His Greek parents moved to Germany in the 1960s, where he was born in 1965. In early 2006 his parents returned to Greece to retire.

Ilias is at home in both languages and cultures, Greek and German. Speaking these two languages (Ilias also speaks English and a bit of Spanish) is an asset in many respects. Having a fluent command of Greek helps Ilias understand and stay up to date with local Greek news and information, as this forms an important part of his conversations with his extended Greek family. German, on the other hand, is the language he uses at home in Cologne with his numerous friends, and in everyday life. English, in turn, is the language in which some of Ilias' work is undertaken.

It furthermore has the role of the lingua franca, especially in working environments that involve people from different nationalities.

Travel is a significant part of Ilias' life. It is of particular importance for his job as a freelance business and economics journalist and consultant. Hardly a week goes by without Ilias travelling to places like Brussels, Berlin or Greece, to name but a few. At the European CEDEFOP Institute (European Institute for the Advancement of Professional Skills) Ilias is working as a reviewer and external advisor. CEDEFOP has always been of interest to Ilias in the past when he was working as a consultant for asset management systems. However, in the past, the physical distance has always kept him from working for CEDEFOP. This has all changed, thanks to advancements in technology, and thanks to MESH! The institute only requires Ilias' physical presence for two days per month as long as he can guarantee being "always-on", i.e. always reachable.

Consequently, it is vital for Ilias Bonn professional life to have access to news content in various formats coming from various sources. Due to his travelling habits it is of equal importance to him to have similar access capabilities through portable end devices that work in all locations world-wide and let him access news he requires when he requires it. MESH is to provide for most of the needs of Ilias Bonn and the following sections will highlight this in greater detail.

Ilias Bonn writes as a specialist journalist about business and financial issues. His expertise lies in the area of telecommunications and how these influence journalistic working practices. He regularly reports about international conferences for the European Journalism Centre (EJC) in Maastricht and is a lecturer / trainer for the Centre's e-Learning website. He also regularly publishes articles in various journals, magazines and newspapers, as well as online media.

Switzerland's Neue Zuercher Zeitung is the only medium that publishes his works in the "traditional" way, namely in printed form in the physical newspaper. However, Switzerland is very advanced when it comes to asset and content management. In 2009 the publishing house became one of the first commercial clients of the MESH platform. It soon became obvious that use of the MESH platform and system would have a fundamental impact on the journalistic working practices and the way news are

consumed.

All tools that form part of the MESH platform are online-based. Connectivity has been ensured as a result of the EU's i2010 strategy. As a result, "always-on" connectivity is guaranteed in all EU member states for all its networked citizens. Any disruption for more than three minutes results in hefty fines for network providers. As a result, there is hardly a place left in which one cannot connect online, should it be desired. One of the few places that still experiences some difficulties in this respect is on the Thalys train that connects Cologne with Brussels. Although the trip can now be made in 70 minutes, and although satellite-based connections have been in operation on the train since 2007, the repeaters in some of the tunnels seem not to cope with speeds above 320 km/h. The situation is altogether different in Switzerland though. Connections seem to even improve and stabilise the longer tunnels.

Travelling by plane, too, is no longer pure relaxation or offline time. Some passengers disconnect the VoIP functionalities of their devices though, in order not to receive calls and disturb fellow passengers constantly. Switching off is no problem as calls are being transcribed using the speech to text MESH functionality in real-time. Passengers who get dizzy reading however, can use the external speech synthesis modules which plug onto the MESH system and allow reading of documents by one' s favourite voice. Time high up above the clouds is also ideally suited to view video content, as this is much easier done than reading or typing. Ever since MESH has become fully in line (synchronised) with Ilias Bonn's profile, even short journeys are ideal for effective video searches as MESH automatically refines results to clips which could be interesting to Ilias for watching or reusing in his articles. Requesting licences for the material is dealt automatically by the MESH platform. New licence requests only occur when new sources are accessed and used, or when content is accessed on a basis different to usage models that are in operation already between Ilias and the various content suppliers he uses.

Security is an overriding concern and of utmost priority. This is not only the case in trains, at airports or when passing through toll stations, but in numerous other everyday life situation such as accessing a physical work space and the like. Here, MESH is at the forefront of developments, too. It has to compete with other strong products, mainly originating in the US and Asia. Contrary to its competitor products, independent organisations are conducting a monthly data security audit of the MESH system and the data stored in it. This is a pre-requisite in the EU as data security and data privacy has become a citizen's right and forms part of the European Citizen Charta. Ever since IPv6 had been introduced, leading scientists have frequently pointed out the dangers that go with these advancements in technology. Especially, the permanent tracking of movement and activities of individuals has been highlighted again and again. The same holds true for the misuse and abuse of personal profiles and data.

The "content auctions" (S27) (content sales of his copyrighted material) which Ilias conducts for international re-use generate further income for him. Through semantic analysis MESH understands the content of Ilias submitted material and suggests purchasing to specific MESH users according to their preferences as described in their profiles. When a user expresses an interest in acquiring a news item purchasing is negotiated with the owner who is notified of all interested parties and can decide on his pricing policy.

Every three hours MESH provides Ilias with a personalised summary of Greek news

including textual articles, videos and images from several sources. Pictures and video material from the early days of the Internet and streaming media are available free of charge from German providers. A high quality video of Germany's first true broadband on-demand video portal "Qurt", in turn, costs seven MESHros. The MESHro is the currency that is used to make transactions on a non-commercial basis (no re-selling allowed) among users of the MESH platform. This fee may increase to up to 20 MESHros if the item is accessed more than 2000 times. If individual items are requested more than 5000 times, the system and its agent propose a commercialisation of the item. It also suggests paying a certain percentage to the operators of MESH (e.g. 10%) while, in turn, reducing Ilias' annual subscription fee to the MESH service by another percentage per cent - depending on popularity of the item made available). The MESH system / architecture "knows" what it is doing as the score of its users regarding individual content items are put into correlation with each other, providing content suppliers with valuable information about quality and quantity of the content that is being offered.

Items that have been produced in the past and are stored or referenced in the MESH system remain in the MESH archive, accessible at all times. Every time new and related material is added, this is referenced, linked and associated with existing content. This process can be based on available metadata or on the automatic extraction of meaning, done by individual MESH tools and components. Apart from using various automatic annotation mechanisms, Ilias is still also using a manual annotation tool for tagging his work according to his personal judgement or for checking on the automatic annotation provided by MESH. Furthermore, there are a number of specialists from his respective fields of expertise in various MESH peer groups. While travelling, and in the course of conferences and physical meetings, they all greatly enjoy testing the limitations of the MESH platform and system. One of their favourite tests: finding a famous quote, a particular tune or a well-known movie scene… Often, the winner is not MESH nor Google nor SMART-WEB, but one of Ilias' friends and colleagues!

The video search function also provides quick and quite reliable results as both real-time and off-line analysis are used for the recognition of content based on the understanding of images, speech and text. Even older material is always taken into consideration with every query, either because the personal agent has "learned" from past experiences or because other MESH users have fine-tuned and further improved existing data.

MESH has become a vital part in the life of Ilias Bonn. He can no longer imagine his life without MESH!

Martha Jong – a retired doctor

In the late 1960s Martha Jong went to the Afghan-Pakistani border region for the first time and worked as a doctor in a temporary (military) hospital in a peaceful part of this region. For the next 35 years she spent her annual vacations as an eye specialist in that same region and, during the last two years of her professional life, worked permanently for an NGO in Pakistan. Today, hundreds of diaries, thousands of photographs and a video archive

with tapes in various formats such as VHS, miniDV, and HDV100 are piled up in the cellar of her London house. She had learned to appreciate the advantages of the internet early-on and corresponded online with her colleagues, patients, friends and the administration in Pakistan.

Since 2007, almost nothing has been published about this region, although many lives fell victim to terror there (especially in the last few years where there was a proliferation of violence.) Martha sometimes wonders whether it was all worth it. And things did not really improve much to this very day. Rather the contrary: Now, in 2010, Martha can no longer travel to the area because of safety reasons. Even so, she spends long hours trying to find the truth as to what is really going on in Pakistan … for herself, and for others.

Usually, she has to find and collect all the desired information herself. She does so by accessing various sources through the MESH platform, which provides her with an overview of most news in multimedia summaries. However, she can not always really trust what is presented in such a way. Only a few broadcasting organisations and agencies still have correspondents in the region. A lot of information is being sent out via blogs, video-blogs, newsletters, e-mail communication and the like. Pakistani broadcasting stations and the news of the allied occupying powers can only partially be trusted.

Martha often receives photos of people and landscapes and information which seems out of date or even manipulated. She has meanwhile digitized the diaries and her entire audiovisual material and published it online. Many of her former colleagues in Pakistan have done this, too. Some have gone so far to add newspaper articles to her blog about Pakistan. Martha gives the MESH platform access to her notes via a web service and permission to also use her audiovisual material as reference material. As a trusted source of information for the situation in Pakistan, she has reached a point of valued credibility. Some of her former colleagues, especially the engineers who helped in building the infrastructure, also possess lots of material. In the case of questionable news from agencies, Martha instructs MESH to carry out a deep-linking with the blogs and sources. In doing so, especially photo and video material can usually be shown to have been manipulated. Martha's photographs of buildings, people or landscapes often appear as search results. In such cases, Martha examines whether the material is genuine or has been manipulated. The research results are added to the metadata of pictures and videos.

Broadcasting organisations such as the BBC, DEUTSCHE WELLE, as well as various publishing houses have already worked with Martha’s archive material. MESH organizes and regulates this by means of usage rights. Simultaneously, usage fees are being paid. Big media companies often use an interface connected to the MESH-system or are part of the MESH platform, so that a personalised and automated deep linking can be effectively used in the comments on the contributions that have been added.
In the beginning, MESH experienced some minor problems with news summaries and creating links between them. At that time, Martha resided in Spain and looked after her daughter's grandchildren, leaving her with little time for her Pakistani interests. In order to resolve this problem and to assist Martha further, MESH is now searching for more Spanish sources related to the subject. Highly important information is directly sent to Martha ’s PDA and while she is travelling around by plane, she is provided with relevant

summaries. These are usually long enough to last the duration of a four hour flight. Summaries are read out to her, of course, as Martha had never been able to read in an airplane without becoming travel-sick.

John Clark – a news correspondent

John Clark is an independent journalist writing articles for several news organisations. To maintain an adequate contact with past and current news events he is subscribed through the MESH platform to several news libraries. Today he is sitting in his office at home trying to put together a review article on past developments in the middle-east through the last decade. Before engaging in any work, he submits a query in his natural language through his home internet connection to the MESH platform. The system queries all connected news organisations and a distributed semantic search is initiated in all multimedia (text, image or video) material residing in the remote libraries. Although the query is submitted in English, the search engine adapts this to the language of the material queried and returns all available results in their original languages. The platform assembles all available multimedia material that satisfies the query and categorises content identifying the fragments that refer to the same event, and comparing them to deliver the results in a structure identifying similarities in content. All material is assessed according to the credibility of its source at the time of the article composition, as the platform utilises an intelligent source credibility measure which depends on past user satisfaction as well as on expert opinions.

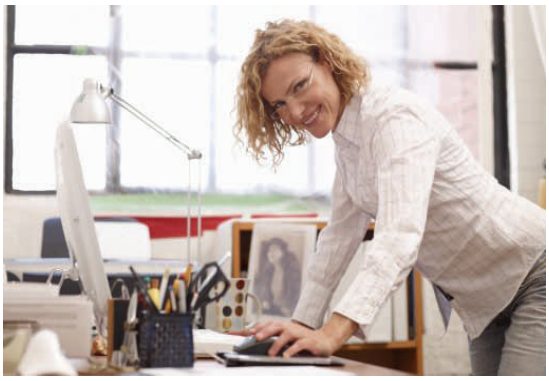
John is then given the result of his query with links to all relevant material. To help accessing the content, the query results are clustered according to thematic content; for each cluster a core overall multimedia summary is presented, together with a relation of the sources it has been produced from, information about the differences between those sources and a short text summary of each source. Organisations that John is subscribed to offer also expanded multimedia summaries and access to the original sources; in other cases he is offered a highlights excerpt and has the option to either subscribe to the organisation, or purchase access only for this article. All payments are automatically handled by the platform upon John's approval, or cleared through his monthly subscription. John marks the material he selects to download, and the platform returns it in the preferred format according to his profile.

When John finishes his article he considers several alternative ways for marketing his work. He can easily pass it to his preferred newspaper with the usual price he has agreed by uploading it through the MESH platform within his folder in their private space. However, since this is a special piece of work, he decides to make it available through the platform to the highest bidding organisation. So, he sets the time deadline until which bids are to be collected and either uploads his own summary which will be viewable to all interested organisations or instructs the platform to automatically compose a short summary. At the end of the deadline he will log in, view the bids and decide, considering the price offered and the prestige of each organisation. He

could also decide at that point to make it available to all public with a small fee per reader, if he considers this scheme to be more beneficial.

Actually, last week when he was in Latin America covering a hot story he preferred to make the article available to the public with just a small fee. He realised he made much more money this way, as an amazing number of readers wanted to read the story as soon as it happened. He managed to be the first journalist producing a full article as he wrote the story on his mobile device which used his favourite template and made the article immediately available through the MESH platform. A simultaneous search for relevant pictures on the platform gave him the capability to purchase and include in his article a couple of photographs taken by a photographer who was covering the same event. He had taken the photographs just a few minutes ago, annotated them with his mobile phone and made them available through the platform. John did not even have to upload the full resolution copies. Browsing through the thumbnails he decided which he wanted to include and instructed the platform to insert them into his article according to his selected template, which had an optional placeholder for multimedia material.

Claudia – a business analyst



Claudia switched on her PDA as she entered the train going to work in the morning. Claudia is an investor and the first thing she checks every day is the latest news on the stock market along with any developments in the financial sector. Her profile is described along these lines on the MESH platform and the first page she sees on her browser is a compiled table of viewing recommendations extracted from among the news material generated since her last login, together with automatically generated expanded versions of material she accessed in previous sessions and considered interesting to track. While watching content pieces, she can skim over uninteresting pieces by dynamically selecting ‘summary-mode’, in which the material is excerpted on the fly. From time to time references to related information appear; the system adds the ones Claudia feels worth checking to a navigation folder for later review (if a link catches her attention, Claudia may follow immediately and later return to the segment she was viewing). Some of the longest content pieces are extracted so that only the portions that are relevant to Claudia’s interest are shown. She also has a number of favourite financial analysts. Every time one of them publishes an article either directly on the MESH platform, or through a news organisation, Claudia is given the link to this on her browser’s first page. Some columns she has defined as very critical and the moment they are published, she is also sent an alert with a reminder. Although Claudia described her profile in very general terms when she first subscribed to MESH, the platform has specialised her preferences, after watching Claudia’s reading and her assessments on articles, which she fills in from time to time. However, if she can afford the time, after going through the platform suggested access she also does a more thorough search, in case something else happens to interest her as well. Quite often she also selects to purchase specific pictures or video clips to use in her reports and presentations. In some

cases they are even free of charge for non-commercial uses. Claudia repeats her reading several times through the day, either from her office on her desktop computer, or remotely through her mobile device. The MESH platform immediately understands her terminal device and adapts to the proper preferences.

Bill Jones – building a news agency

Mr. Jones has been in the area of news publishing for quite a few years, working for several organisations, usually on the editorial board. However, he was never satisfied with his financial benefits even though he worked very hard and had developed a great intuition on how to market a good story. He has been looking into starting his own business in the area for quite some time, but the necessary investment was prohibiting and the associated risk too high for him to take.

Recently he decided to try building a virtual news agency through the *MESH* platform following the example of a past colleague. Using acquaintances he had made with several journalists, photographers and cameramen over the past few years he drew some deals with several of them covering diverse topics so that they will provide stories on their area of expertise in a close to daily basis. It was quite easy to build his own secure space on the *MESH* platform, which his associates will use to find suggested templates and upload material. Of course, he will also be able to purchase material published every day on the platform by independent journalists.

Having defined his suppliers, the next step was to set his potential consumers. Using again his contacts in the world of media he made key people aware that he will be providing through the *MESH* platform daily material on a number of interesting to them areas. He also announced the kick-off of his portal to relevant MESH societies and managed to attract several subscriptions for immediate access to published material in their areas of interest. More promised to keep an eye on his portal on the MESH platform and maybe purchase particular news items or decide on a subscription that suits them later on. Advertising his virtual news-agency on the home page of *MESH* he may also attract the general public to either purchase access to particular stories, or build their own electronic newspaper based on their preferences. All mechanisms are centrally provided by the platform and have been built to be easily mastered by any individual news reader. Special press review services are also available using the technology provided by the MESH platform through specific modules at an extra cost which is however mitigated to the agency clients. These reviews may involve a Market Analysis for a specific area which may include new products, entertainment news (such as new films, books, etc.), or even an analysis of market competitors (based on published news material).

Mr. Jones also decided to appeal to younger ages by providing personalised e-papers to students and young researchers who define their interests around particular thematic areas like science, or environmental issues. This will also function as a dynamic

personalised encyclopaedia, delivering film clips, news items or documentaries in an organised manner according to the readers interests. Strict legislation however imposes limitations on the content MESH may deliver according to age groups. This is handled by the *rate-it* MESH plugin which rates multimedia extracts into predefined categories according to the involved content (e.g. violence or adult scenes).

Teodora - A young periodist and blogger

Teodora is a young periodist, daughter of the director of a known newspaper. She is well aware of the changes that are happening in the information sector, and thus has setup a personal *teodoranews.com* portal with the help of the MESH platform.

Her site is quickly growing in number of users, currently nearly 16000 visits average per day, by means of an intelligent use of the MESH blog integration module, the links to many other small independent sites, easily located in a navigation bar in a side of the web, the RSS feeds, and many activities for syndicating and aggregating information. For instance, joining different information sources in her single portal.

Teodora earns some money from the publicity in her site, but she also works in a traditional TV channel, so she works full-time in the news sector. She would not be able to do any of her duties without the internet:

She start her day consulting the MESH platform, which uses feeds from news.google.es, but highly personalized to her own requirements, as her main area for the TV channel is technology related news, with strong focus on social changes as a consequence of these new technologies.

As she reviews and filters these news items, she pays more attention to the most interesting ones. MESH captures her attention patterns and makes them available to those relatives (and customers) of her who consider she has some authority in her fields of interest.

She then consults the news produced by her colleagues and friends by using MESH, which retrieves summaries, and filters lots of information to get a quick view, so that she can post news on her personal website before she leaves for her work in the TV channel. She downloads the summaries, with all the attached information that allows her to filter and personalize it, to a PDA.

Very often, she is in charge of a particular news program for the TV station, and she has to prepare all the information, under immense time pressure. Main problem nowadays is to find the exact information you are looking for, the more difficult the less structured or less known theme you are searching for. With this purpose, she uses MESH as a search mechanism with highly optimized algorithms to suit her preferences.

So that she can pay attention to the most relevant news items related to the topic she has to cover, MESH lets her benefit from the expertise of her community of correspondents by letting their own attention patterns filter out those news items that they did not consider

worth being paid attention to.

When she is travelling towards the TV channel studios, she consults the MESH summaries downloaded to the PDA. And also whenever she does not find the expected information using her home computer, she uses the MESH mobile service, which is less powerful, but more efficient on the move.

The MESH system is at the centre of her work and she has a lot of control on many of the parameters, which allow her to optimize her use of the system:

She can set her preferences manually, and change them at any time. These preferences configure rules, but she has control on the way the preferences are translated into filtering rules. She can also modify or update filtering using examples. For instance, if she finds that the filtering has taken out important news for her, she can instruct the system so that the filtering rules are changed accordingly.

The MESH platform also provides suggestions for particular themes based on previous search history. MESH stores the search history but using all the privacy mechanisms also used for maintaining the user preferences within a secure service.

MESH allows for control on the recall / precision trade-off depending on the scope of the search made. She has control on this trade-off by using a single measure of retrieval effectiveness.

When on the move, the MESH system detects this context and adapts its algorithms to present more precise information, using more strict rules for filtering. This way, the results provided by MESH are better suited to less than optimum context.

As an additional aspect for the mobility condition, MESH applies screening mechanisms, instead of full-blown filtered and optimized search, so that good quality results can be provided quickly. This is especially effective in conditions in which the user has little time to consult too much information, and thus requires very quick input of good quality, if not optimum.

You –in 2030!

You are in your kitchen on a rainy Monday morning in the year 2030. You pour a cup of coffee and turn to the kitchen wall. “Give me the news!,” you say, and the wall, changes into a gorgeous full-color map of the world. Headlines, pictures pinpoint the locations of news stories from a variety of sources around the world.

You ask for each story in the order you prefer. You are saving time by getting only what you want, when you want it.

You grab your personal digital assistant, a combination of cellular telephone, computer, television, radio, and fax machine. Plugged in all night to the multimedia center at your home, this tiny device has been receiving updated versions of customized news. You leave for work..

Is this science fiction? Maybe today. But we are passing through truly exciting times and the better prepared will be the ones to lead media in the Semantic Age of journalism…

The research leading to this document has received funding from the European Community’s Sixth Framework Programme. However, it reflects only the authors’ views, and the European Community is not liable for any use that may be made of the information contained therein.