RESEARCH   **Open Access**

An architecture for integrated intelligence in urban management using cloud computing

Zaheer Khan[1*], David Ludlow[1], Richard McClatchey[1] and Ashiq Anjum[2]

Abstract

With the emergence of new methodologies and technologies it has now become possible to manage large amounts of environmental sensing data and apply new integrated computing models to acquire information intelligence. This paper advocates the application of cloud technologies to support the information, communication and decision making needs of a wide variety of stakeholders in the complex business of the management of urban and regional development. The complexity is evident in the socio-economic and environmental interactions and impacts embodied in the concept of the urban-ecosystem. This highlights the need for more effective integrated environmental management systems. A key to understanding the nature of integrated environmental management systems is the identification of the need for horizontal integration of information across sectoral inter-agency boundaries at the local level, and the need for vertical coordination between levels of governance. This paper offers a user-oriented approach to the specification of requirements for the effective management of urban areas and the potential contributions that can be supported by cloud computing. The commonality of the influence of the drivers of change at the urban level offers the opportunity for the cloud computing community to develop generic solutions that can serve the needs of hundreds of cities throughout Europe and indeed globally. In this respect, different cloud based architecture scenarios are presented which utilise capabilities compliant to various standards in generating information and intelligence for urban governance.

Keywords: Information intelligence, Environmental monitoring, Data harmonisation, Service integration, Cloud and computing standards

Introduction

Integrated environmental intelligence can be described as the capability of a system to access, process, visualise and share data (spatial and non-spatial), metadata and models from various domains relating to land-use/cover, biodiversity, atmosphere as well as socio-economic aspects of the city for various purposes. Examples include the identification of environmental change and causality, future environmental trend analysis, socio-economic development, policy development and collaborative decision-making. Even if, there is a profusion and explosion in the growth of environmental data, this data mostly is fragmented and un-harmonised. Furthermore, such data exist in proprietary and open systems, less compliant to

standards and sometimes requires extensive computing capacity, constraining its use across different platforms. This suggests that integrated environmental monitoring requires compliance to standards, data harmonisation and service interoperability together with extensive on-demand processing and storage capacities in order to answer scientific and policy related questions.

In the above context, this research is an attempt to bridge the gaps between fragmented cross-thematic environmental information, with a particular focus on the urban environments. In this regard, it aims to develop a cloud-based framework that facilitates data accessibility and storage across platforms, and provides necessary on-demand computational resources for the required processing, simulation and visualisation tasks. This framework also aims to provide mechanisms for integrated information intelligence necessary to resolve the highly complex cross thematic multiple impact challenges associated with

* Correspondence: Zaheer2.Khan@uwe.ac.uk
[1]Faculty of Environment and Technology, University of the West of England, Bristol, UK
Full list of author information is available at the end of the article

climate change, as well as measures to ameliorate these impacts associated with climate change mitigation and adaptation. More specifically, from a user perspective, the aims include:

- informing complex policy and science related questions which cannot be addressed directly from non-integrated information sources;

- studying different environmental phenomena by investigating the causality relationship between environmental variables;

- enabling access to cross-thematic environmental information for climate change mitigation and/or adaptation;

- facilitating various stakeholders in generating environmental indicators to support policy interventions, assessment of the state of the urban environment, decision-making and policy implementation;

- empowering citizens in participatory environmental data collection and collaborative decision making processes and

- designing and utilising state of the art user-centred information technology tools and techniques in relation to all of the above.

From a technological perspective, the above aims help in identifying the required system capabilities which primarily necessitate acquisition and harmonisation of cross-thematic spatial and non-spatial data (using standards) in order to provide uniform access to interoperable information for various purposes. Further, it indicates that in order to deliver integrated intelligence derived from an increasing volume of environmental data, significant storage capacity is needed to preserve the environmental data and additional computational power is necessary to process this data. Furthermore, it suggests that specialised processing and visualisation services are needed to analyse the data in order to support enhanced decision-making. For instance, when the output from processed data e.g. information maps, scale from a local (i.e. City) to national or regional levels for comparative environmental indicator analyses, the amounts of data and the number of processing requirements typically increase. In this case, the cloud computing paradigm offers an excellent means of fulfilling the above requirements. It can provide an on-demand framework to make systematic the information flow from data collection to analysis, monitoring and assessment based on standard guidelines such as ISO 19100 Series, the Open Geospatial Consortium (OGC) [1] standards and the INfrastructure for SPatial InfoRmation in Europe (INSPIRE) [2] specifications. Figure 1 depicts both user and technology perspectives to achieve integrated information intelligence for environmental monitoring.
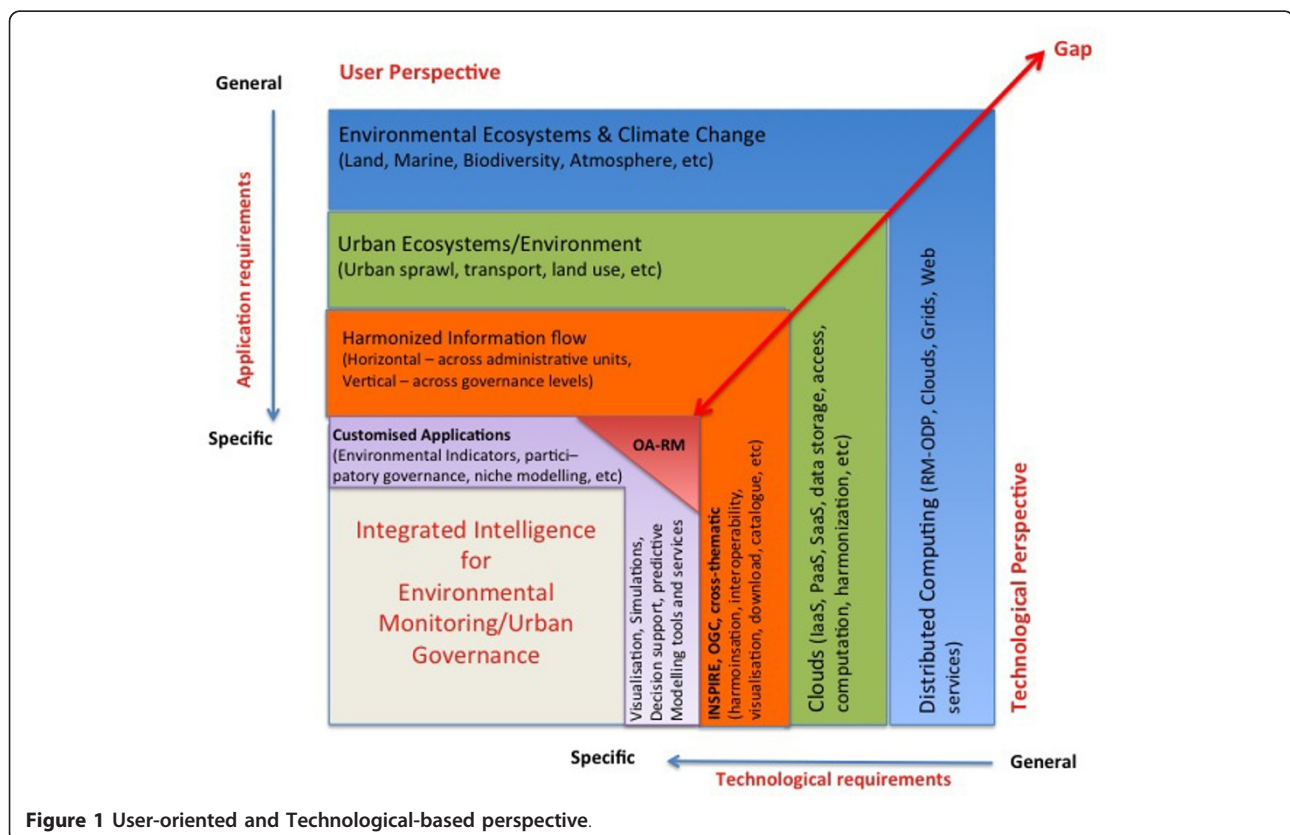

Figure 1 User-oriented and Technological-based perspective.

The challenge here is to introduce appropriate mapping, harmonisation, integration and utilisation of suitable tools and technologies in a cloud environment in order to fulfil stakeholder requirements.

One of the basic elements for Integrated Environmental Monitoring (IEM) is the processing of 'data', which in the case of urban environments is mostly geo-tagged (e.g. in space and time) and combined with socio-economic and demographic elements. In this respect the OGC provides a sound foundation of standards in the distributed geo-processing domain. Recently OGC highlighted the significance of cloud computing for the geospatial computing world by indicating a number of application areas where cloud computing is well suited, including the modelling of complex environmental applications, the fusion of existing preserved data and new sensor data for decision-making, enterprises, scientific and governmental storage and compute-intensive applications [3]. Furthermore, it is highlighted in [3] that exchanging geospatial data between systems (clouds or user applications) is not sufficient rather it is also necessary to understand the meaning of data. For example the existence of thousands of spatial reference systems; various representations for spatial data including raster, vector, points and others; semantic heterogeneity; proprietary interfaces and vendor specific encodings also add to the complexities. All this requires integrated solutions for information intelligence for environmental monitoring.

The remainder of this paper is structured as follows. In next section background and related work are presented. Then a user perspective is presented briefly in order to derive technological capabilities in clouds. After this we present a proposed architecture for IEM followed by different implementation scenarios using cloud computing. Then we briefly discuss integrated intelligence and decision-making approaches followed by an assessment of the proposed architecture by identifying its benefits, potential research and technological challenges and an overall reflection on the expected research outcomes is presented. Finally, we conclude and identify future research possibilities.

Background and related work

The development of data harmonisation tools, decision-support systems, visualisation and simulation tools for specific environmental thematic domains is challenging. But the development and implementation of an integrated intelligence framework that can accommodate cross-thematic data harmonisation for various application domains using a variety of tools, components and services is even more challenging.

In the above context, the Global Earth Observation System of Systems (GEOSS) [4] and Ground European Network for Earth Science Interoperations-Digital Earth Communities (GENESI-DEC) [5] are examples of ongoing initiatives for integrating domain specific environmental data from various sources. More specifically, the recent development of new methodologies for information acquisition (for example, the Shared Environmental Information System (SEIS) [6], Global Monitoring for Environment and Security (GMES) [7], and information provided by member states to European Environment Agency (EEA) arising from various Directives including as Air quality, Noise and Water) and the increase of computing capacities (e.g. sensor nets, smart phones, Web 2.0, grids and clouds) opens new fields for monitoring and assessment with high information flow and the capacity to manipulate it.

From a technical perspective, managing increasing data volumes, the varying needs of scalable computing resources, harmonisation and interoperability processes, tools and services, visualisation and compliance to standards are the major drivers in developing an integrated system. In this regard, the cloud computing paradigm is a suitable cyber-infrastructure to fulfil most of the above requirements due to its characteristics such as elasticity, on-demand processing, a pay-as-you-use cost model, lower IT management costs etc. [8]. For example, recently it has been indicated that £21bn - the annual UK government IT budget - could be reduced by £3.2bn out of a central budget of £16bn by putting government applications and services in the cloud such as the proposed G-cloud [9]. The scope of cloud computing may vary from a private (e.g. for a single organisation use) to public (e.g. access provided beyond an organisation's use) or a combination of both (also called hybrid clouds). The scale of service provision also varies from Infrastructure as a Service (IaaS) to Platform as a Service (PaaS) to Software as a Service (SaaS). The level of flexibility, service development and management in using customised solutions varies from IaaS (i.e. more flexible, more development work and high management) to SaaS. There are many commercial vendors already providing public cloud services: for example, Microsoft (Azure), Amazon (EC2/S3), Google (App Engine) and Rackspace etc. Similarly, some well-known research based initiatives are also under development such as Eucalyptus, OpenNebula, Nimbus, and OpenStack. The cloud user community also demands standards and preferably open interfaces and encoding to avoid lock-in and to facilitate easy migration from one provider to another. In this regard, the Open Cloud Computing Interface (OCCI) [10] is an OGF working group that aims to develop interoperability, portability and integration standards to serve many cloud computing models such as IaaS, PaaS and SaaS. OCCI specifications provide a RESTful protocol and API for many kinds of management tasks. OCCI is work in progress in many cloud projects such as Eucalyptus, Openstack, SLA@SOI, RESERVOIR, jclouds etc.

There exist cloud (and in particular SaaS) enabled tools for different environmental applications such as land, transport and urban planning, e.g. ESS [11]. Similarly, the UK Natural Environmental Research Council has funded a research project called Environmental Virtual Observatory pilot (EVOp) [12] to use clouds to achieve similar objectives for the soil and water domains. However, further steps are needed to realise cross-thematic integrated intelligence for environmental management, using standard harmonisation processes in order to facilitate information requirements at all levels of governance i.e. local, national, and EU.

An example: user perspective for an urban environment

Urban management aims to respond effectively to the key political concerns of European cities today including the socio-economic and environmental impacts of climate change, deteriorating public health and biodiversity loss etc. All of these urban impacts are associated with a dysfunctional urban model in which urban transportation systems, in the context of urban sprawl, generate excess greenhouse gas (GHG) emissions promoting climate change. Furthermore, motorised transportation as the prime generator of GHG emissions also degrades the environment in respect of noise and air quality, impacting adversely on human health. These political concerns arising from the direct socio-economic and environmental impacts of the drivers of change are fundamentally related at the urban level in the land-use-transport-environment nexus, in which the degree of compactness of the city determines the potential for public transportation solutions which in general offer multiple socio-economic and environmental benefits.

Given the interrelated, interconnected and complex nature of drivers and impacts in the urban environment, policy integration and integrated urban governance are viewed as the "sine qua non" of effective responses. Urban governance is thereby promoted via integrated approaches to territorial impact assessment, policy formulation and policy implementation.

Despite these aspirations for the development of integrated approaches, the reality remains that urban intelligence at the local level is generally poorly developed in the critical areas of territorial impact assessment and urban planning. Agencies addressing the management of urban regions have a critical need to integrate large sets of data across the sectoral domains of land-use, transport, health, etc. in order to respond to the political demands for sustainable cities, but this is rarely achieved effectively. Furthermore, the complexities surrounding data integration to secure reliable intelligence increase further when the typical framework conditions for inter-agency collaboration on territorial planning issues requires intercity or intra-regional comparisons and

assessments. Failure to secure policy integration can be attributed to a variety of factors including notably organisational and procedural barriers associated with the sectoral responsibilities for land-use, transport and environment, primarily in a horizontal perspective at the local level, and in a vertical perspective between government agencies responsible for policy development from local, to regional, national and EU levels.

These failures further reinforce the demands of the policy-making community for the development of integrated urban management frameworks, reflected in the demand for definition and development of new assessment methodologies for urban and regional development. For instance, the European Environment Agency's (EEA) Integrated Urban Monitoring for Europe (IUME) initiative [13] connects the political drives for climate change mitigation and adaptation, healthy and economically viable urban communities, that characterise the European ideals of good governance and sustainable development. Further initiatives are focused around the policy model (see Figure 2 below), which identifies specific data requirements of planning agencies according to a generic agency driven policy cycle for the management of cities. Furthermore, the definition of these integrated management frameworks must also respond to the evolution of the governance model and new demands, external to the government process, arising from the requirement for effective stakeholder and citizen engagement. Such demands necessitate new mechanisms for building capacity and supporting the ability of a community to create sustainable futures, as advocated, for example, by Yigitcanlar T [14], by introducing citizen science, participatory governance and collaborative decision-making.

The user-defined generic city management policy cycle (see Figure 2 below) structures policy related information requirements within a framework that proceeds from issue identification and description, to the definition of response and engagement with the political authority, to implementation and monitoring of the effectiveness of the actions proposed. Information required to support planning activities within such a management framework permits the identification of precisely the nature of the information that is required by which agency, for which purpose, and at which stage in the policy cycle. These user-defined integration principles permit the specification of policy defined requirements which in turn form the foundations for the development of ICT related solutions, specified in the framework of the data and process model as described below. Figure 2 demonstrates the basis for an appropriate user and policy defined model of information needs according to a model of the horizontal integration of data flowing between sectoral responsibilities at the local level and vertical coordination of data flowing between agencies of governance from the local to EU levels.

Figure 2 Integrated Environmental Monitoring: Needs and Challenges.
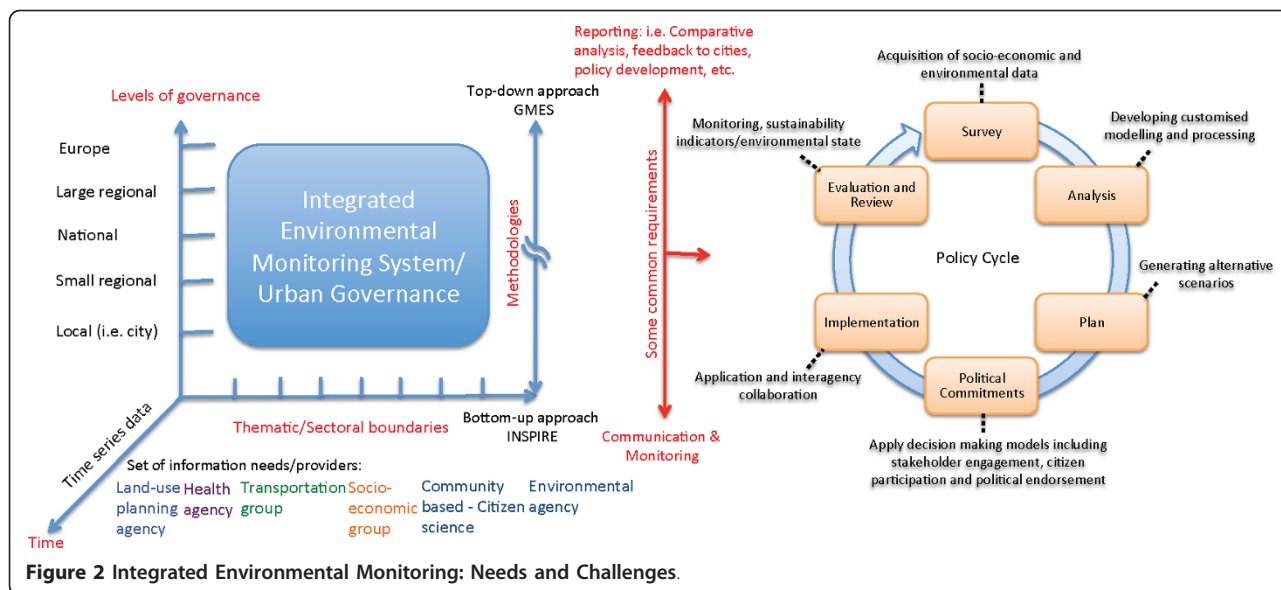
Mapping the user-oriented perspective to technological perspective

Using the above user perspective we now attempt to present a common process to be used for environmental management with the objective of identifying technical capabilities required for the implementation of an IEM System (IEMS).

A generic process for environmental management using ICT
In the following we attempt to present a general monitoring process with the objective of mapping process activities to specific methods, technologies and tools for integrated information intelligence. The main steps of the process are:

Step 1 ***Data Acquisition:*** several approaches can be considered for this step including automated discovery and retrieval from distributed repositories, remote sensing, participatory sensing and citizens' science, etc. For example, EEA requires member-states to produce domain specific (i.e. air quality, noise, water) harmonised catalogues, a meta-database including data sources, data sets and other information such as reports and published meta information via EEA guideline compliant national web sites. EEA also use a Clearing House Mechanism (CHM) to get access to data and exchange of information between various stakeholders. Local governments are also active in local environmental monitoring initiatives e.g. SI@M in Vitoria-Gasteiz [15] and use different sensors (air quality, noise), LIght Detection And Ranging (LIDAR) and other methods such as population surveys etc. for data collection.

Step 2 ***Data Assimilation and Preliminary Analysis:*** this step is mainly required to ensure that captured data meets standard or common application schema structures,

and may require the harmonisation and transformation of the data according to required target system specifications. This step also determines whether there is need for additional data for required processing services.

Step 3 ***Data Integration and Processing:*** after performing discovery and integration of the required datasets, this step includes processing for various environmental models and visualisation of results for decision-making, etc.

Step 4 ***Data Export:*** this step is needed for exchange of information for further processing, storage and use in different tools and applications such as different visualisation needs by different stakeholders.

Required technical capabilities for an IEM system
The most effective environment for accomplishing integrated information intelligence for urban management can be a service oriented environment based on common core standardised data formats, services and applications. Such an environment enables the collection, storage, processing, visualisation and dissemination of information between various stakeholders. An important element in a Publish-Find-Bind model (Service producer/Registry/Service consumer) is to achieve interoperability between services by supporting various services, data harmonisation and transformation, and encoding standards in order to facilitate information flow between different services. In general, the above generic process indicates that the following capabilities should be considered as technical requirements for information intelligence in an integrated environmental monitoring system:

1) ***Data Acquisition:*** the ability to collect data from various sources including databases, flat files, web services, sensors networks and web portals (e.g. OGC's Sensor Web Enablement and Sensor Observation Service and Web

Feature Service (WFS), Web Map Service - WMS), participatory sensing and citizens' observations (e.g. using smart phones). This would further require:

a) ***Metadata:*** the ability to get additional information about the data such as ownership, usage restrictions, cost models, etc. and compliance to metadata standards such as Dublin Core or ISO 19115. Furthermore the use of common vocabularies or ontologies to mask semantic heterogeneity can be beneficial.

b) ***Discovery and Metadata Catalogues:*** the ability to discover relevant (meta-)data from various sources that match specific criteria (e.g. OGC Web Coverage Processing Service (WCS) with filter encoding). For instance, the OGC discovery service standard (OGC Catalogue Service for Web - CSW) provides standard interfaces that can be used on catalogues and includes geospatial extensions. Also, the Universal Description Discovery and Integration (UDDI) framework or OASIS's Electronic business Registry Information Model (ebRIM) models can be used for registry mechanism (publish/find/bind pattern).

c) ***Data quality:*** the ability to assess the quality of retrieved data, e.g. missing data, resolution support, etc.

d) ***Retrieval:*** the ability to retrieve the required data from a specified source using various methods such as the OGC download service standard, File Transfer Protocol (FTP), Hyper Text Transfer Protocol (HTTP), eXtensible Markup Language-Remote Procedure Call (XML-RPC), Simple Object Access Protocol (SOAP) services, etc.

e) ***Storage:*** the ability to store transition and processed data.

2) ***Schema mapping and transformations:*** the ability to perform schema mapping to reference datasets, to harmonise spatial schema based on ISO 19100 series (and INSPIRE specifications for Europe), for instance, to coordinate transformation using different coordinate reference systems.

3) ***Service interoperability:*** the adoption of standards such as W3C's web standards, OASIS's RM-ODP [16] and OGC's view (WMS), download (WFS), discovery (CSW) services.

4) ***Data fusion, processing and synthesis:*** the ability to integrate data and apply computational processing steps e.g. using OGC Web Processing Service (WPS) standard, in order to generate desired results and build synthesis around gaps in data coverage (e.g. OGC's sensor, feature and decision fusion models can be utilised).

5) ***Workflow management:*** the ability to design, compose and execute workflows, e.g. in a distributed environment using the ActiveBPEL Engine, Kepler, Taverna, LONI. Standard compliant services enable developers to formulate service chains (orchestration) for a specific business process models (or workflow) enactment (e.g. BPMSOA Framework [17]). For example, an OGC White paper [3] suggests that OGC standards compliant services at SaaS level enable service chaining in a workflow.

6) ***Provenance:*** the ability to preserve and track information about sources of data and processes completed in the past, e.g. CRISTAL [18].

7) ***Visualisation:*** the ability to generate data and process outputs in a user-friendly i.e. understandable from a user perspective by using various Graphical User Interface (GUI) tools and techniques such as 2D/3D models (e.g. OGC Web 3D Service - W3DS proposal), simulations and maps for various platforms such as web and smartphones using OGC Web Map Service standard.

8) ***Decision-making:*** the ability to enable users to take decisions based on the recommended 'best-fit' output from various scenarios using artificial intelligence (predicate, description, fuzzy logic), expert systems tools and techniques such as OGC standards (OWS Context standard, KML standard and Decision Fusion Model etc.), DROOLS and JESS rule-based engines.

9) ***Social-networking:*** the ability to enable users to interact with each other and share experiences (e.g. Web 2.0, forums).

10) ***Information Sharing and Feedback mechanisms:*** the ability to share information (e.g. OGC Context documents, Web services, etc.) between users and to enable them to provide feedback/comments on results, annotate data and processing outputs (e.g. Web 2.0, Wikis, Issue Tracking).

11) ***Security and reliability:*** the ability to implement authentication, authorisation, encryption, decryption, auditing and backup mechanisms to enable use of data and services by legitimate users and avoid loss of data. Secure service standards like ISO/IEC 10181-4:1997 - Security Framework for Open Systems can be utilised.

12) ***Scalability and Extensibility:*** the ability to add new users, new data sources and new application-specific models without compromising Quality of Service (QoS) requirements such as query response time (INSPIRE requirement) etc.

13) ***Alerts and Events:*** the ability to facilitate publish-subscribe patterns for event notifications using e.g. WS-Notification.

Proposed architecture for IEM system

Figure 3 depicts a proposed architecture from a technological perspective. It consists primarily of five horizontal and two vertical layers. The output from the first
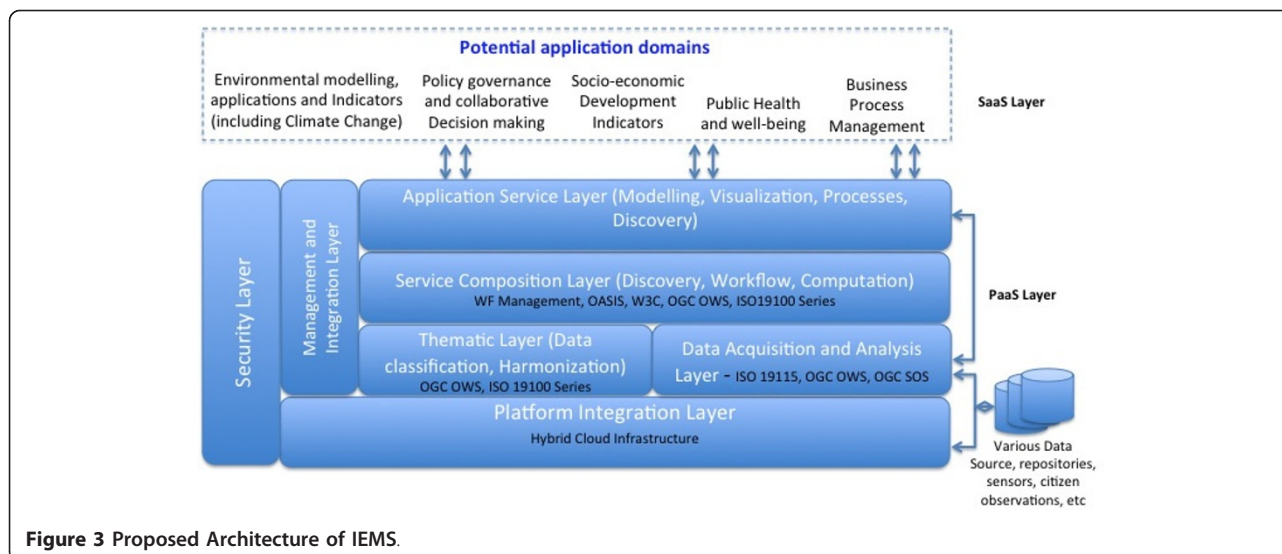
Figure 3 Proposed Architecture of IEMS.

two bottom layers is generic which can be tailored to specific application needs in the three layers.

• The *platform integration* layer depicts a cyber-infrastructure based on a hybrid cloud environment that ensures cross-platform accessibility of environmental data.

• The *data acquisition and analysis* layer is used to access environmental data from various sources including remote database repositories, sensor nets, and citizens' observations in the cloud environment. This layer also ensures the quality of data acquired and identifies the need for necessary data harmonisation and data cleansing.

• The *thematic* layer classifies the acquired data into application specific thematic categories and performs data harmonisation and updates the data/service catalogues for further use of the data.

• The *service composition* layer is required to design workflows, identify data sources, and link necessary processing components to enact the workflows. Furthermore, necessary analytical analysis of the workflow outputs can be performed in this layer. This layer also ensures that the provenance of data and specific processes is maintained that can be utilised for analysis by different expert systems in the *application* layer.

• The *application service* layer uses the outcomes from the service composition layer in application domain specific tools such as simulations and visual maps to perform an analytical analysis for decision-making. Further, this layer enables stakeholders to use existing tools and develop new application domain specific components and services (at SaaS level).

• The *management and integration* layer is used to automate the flow of data and information between the horizontal layers. It ensures that processed outputs from one layer to another are syntactically correct. It also aims

to handle change management that occurs at different layers and to reduce the extent to which the layered architecture requires management overhead.

• The *security* layer ensures the necessary authentication, authorization and auditing for the use of data and services by legitimate users.

One of the requirements in achieving the aims of this research is to understand the incompatibilities between various standards and *processes*. For example, interoperability between OGC and standard W3C web service standards needs additional consideration for wider coverage of spatial and non-spatial application aspects. For example, Lee et al. [19] highlight the need for standards and in particular collaborative efforts between OGC and OGF in order to identify possible areas for initial integration efforts between geospatial and grid computing domains and as a result some projects use standards from both domains such as GDI-Grid [20], use of grid for geospatial data [21], European Space Agency's G-PoD [22], etc.

The service oriented architecture (SOA, publish-find-bind) style has been adopted by general web services which use WSDL and SOAP messages for interaction between web services. Such an approach is platform independent, suitable for composing service oriented workflows and easy to implement. However, it may not be suitable for substantial data transfer. Similarly, OGC OWS (OGC Web Services) are based on a service oriented approach using XML/SOAP and HTTP mechanisms and provide a readymade solution for cloud computing to perform geo-processing activities. However, most implementations of OGC OWS are HTTP based which requires additional wrappers/transformations for cross-platform service compatibility and utilisation.

In the above context, the OGC WS-Phase 5 interoperability programme [23] developed and demonstrated

SOAP/WSDL binding patterns for WMS, WCS-T, WFS-T and WPS. Also, a conflation workflow process and a BPEL script was designed and implemented to demonstrate service chaining and workflow, web processing services, and service interoperability using a variety of OGC service standards. Furthermore, substantial efforts have been identified in the literature concerning the compatibility aspects of web/grid and OGC services which are promising for IEM. For example, Hobona et al. [24] investigated the integration of grid services and geospatial web services into workflows (using BPEL and SCUFL) for geoscientific processing. They indicated that both OGSA and ISO 19119:2005 (Geographic information - services) follow a SOA based publish-find-bind pattern however mechanisms differ and can restrict interoperability between the two approaches. These same authors [24] proposed SOAP-enabled interfaces for OGC XML schemas as wrappers that on the one hand facilitate interoperability between SOAP clients and geo-services and enable relatively easy workflow composition but, on the other hand, introduce performance overheads. Similarly, in [25], researchers investigated a WS-I + based approach for geospatial grid with the objective of utilising continuous data streaming (e.g. data from sensors). They also demonstrated the use of WSDL/SOAP based interfaces for OGC web services (with HTTP approach) and introduced mechanisms to handle performance issues with large datasets. Aydin, et. al. [26] attempted to develop OGC compliant grid services and introduced WSDL based wrappers to transform input/output from OGC (e.g. GML document) to grid services and vice versa. Also, authors of [26] emphasised the need for a decentralised dynamic discovery architecture (like UDDI) to define services, interfaces, operations and protocol bindings for the discovery of WFSs using WS-Discovery specifications. In [27], researchers investigated the possibility of utilising an UDDI based metadata oriented service information model for geospatial services publishing and discovery. All the above literature indicates that the interoperability between general web service standards and OGC standards exists to some extent and can be used in a cloud environment to provide services for different types of clients.

With respect to cloud computing and OGC standards, Blower JD [28] demonstrated the deployment of OGC WMS using Google App Engine. Brananski et al. [29] presented geospatial services in a cloud environment using Amazon EC2. Schaffer et al. [30] investigated the use of SDI components as value-added services into Amazon Web Services (AWS - EC2) cloud and mapped a publish-find-bind pattern from SDI to Cloud computing. They demonstrated the integration of OGC WPS at the SaaS layer and performed scalability tests for Wildfire use case

with the objective of investigating interoperability of geospatial web services and cloud computing. Similarly, Foerster, et. al [31] demonstrated the effectiveness of a hybrid cloud (Open Nebula - private cloud and EC2 - public cloud) together with OGC a compliant 52° north based Web Processing Service. Baranski et al. [32] demonstrated that INSPIRE-related challenging QoS requirements are not easy to achieve without a scalable high performance IT infrastructure, especially during peak times (many clients accessing services concurrently). They achieved promising results by using the hybrid cloud (Eucalyptus - private cloud and Amazon EC2 - public cloud) architecture with OGC compliant 52° north based Web Processing Service.

All the above literature demonstrates that significant efforts are taking place to utilise cloud computing for the geospatial domain whereas Foerster, et al. [31] and Baranski et al. [32] extended the use of a hybrid cloud i.e. a mixture of a private cloud for efficient utilisation of existing IT infrastructure and security reasons and a public cloud for QoS i.e. high performance, reliability and availability at peak service demands. The above also reveals that the use of OGC services in a cloud environment does not change the way OGC services are utilised without cloud computing. Rather use of cloud computing bring QoS i.e. high performance, scalability and low-up front IT infrastructure costs.

Implementation scenarios

This section discusses different possibilities for implementing the IEM using cloud computing. Figures 4, 5 and 6 depict how IEM can be implemented using different cloud architectures. In Figure 4, it mainly represents customer and service provider perspectives. The *Gateway Proxy* is a client side stub to access services in a cloud environment and use according to application needs. The *Registry* component is used to register services provided by a specific cloud platform to enable inter-cloud service discovery and provision. The *Common Access Gateway* refers to a controlled entry point to a cloud environment. It also performs several functions to ensure QoS delivery of information and services to a client. It consists of a *Service Request/Response Handler* ensuring that service agnostic interface i.e. necessary transformation between OGC WS (e.g. HTTP based) and W3C (WSDL and SOAP) based service requests can be handled. The *Resource Pool* consists of a pool of IPs and VMs for private clouds (such as OpenNebula or Eucalyptus based VMs) and public clouds (e.g. AMI for Amazon clouds). The *Application Handler* deals with specific thematic applications and their harmonisation aspects. The *Workflow Handler* manages pipelining of different services. The *Federated Registry* is an instance of a UDDI or ebRIM like structure that keeps track of active cloud platform
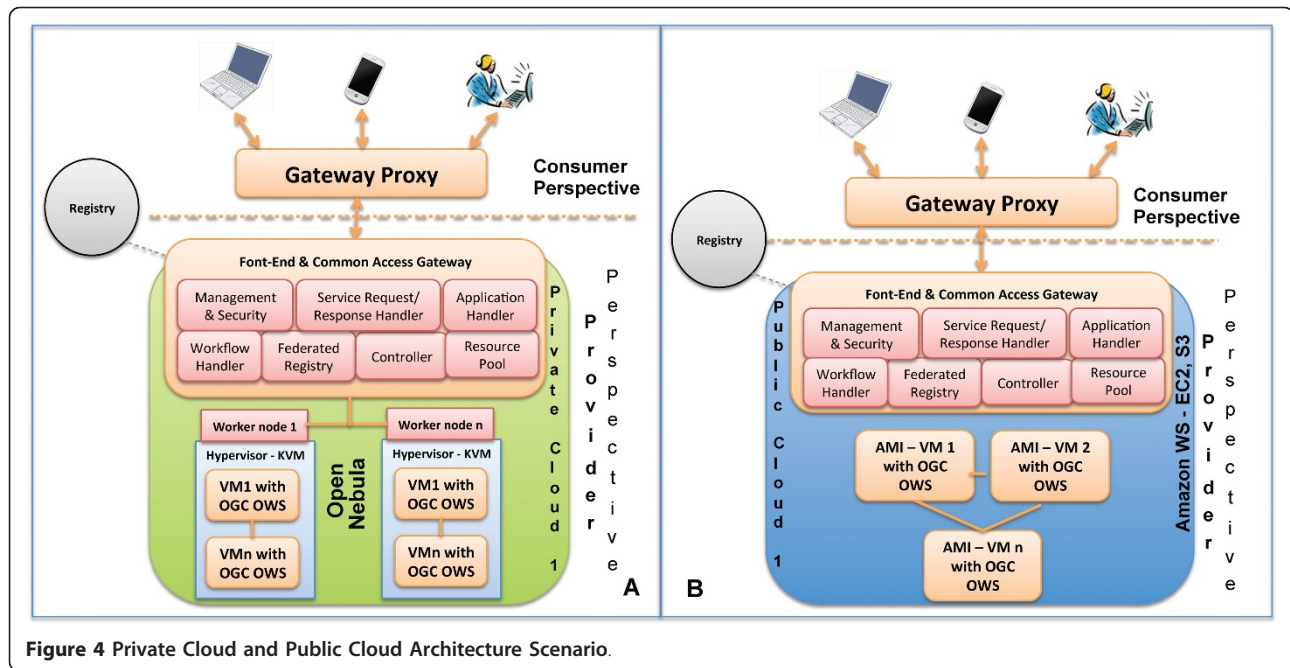
Figure 4 Private Cloud and Public Cloud Architecture Scenario.

resources such as specific thematic services. The *Controller* schedules and manages the lifecycle of Virtual Machines (VMs) on a cloud platform and establishes communication between VMs with clients through the *Gateway Proxy*. The *Management and Security* ensures the provenance of application execution with authenticated and authorised access to services in a cloud platform. Each worker node is running KVM as a hypervisor and each VM instance will consist of open source HTTP server/ Apache Axis, JVM run-time etc. that enables deployment of OGC Web Services and standard web services on VMs. Additional open source software can be used based on specific application demand.

Figure 5 depicts a multiple cloud scenario where two or more organisations can exchange data and services using a private and/or public cloud infrastructure. The *Inter-cloud communication* component ensures compatible communication between two private cloud architectures. The *Registry* component can realise specific service/data discovery and bindings from multi-cloud platforms. Such a scenario emphasises the need for inter-cloud compatibility such as the OCCI [10] demonstrated in different cloud platforms [33].

Figure 6 depicts a hybrid cloud architecture where resources can be deployed in both private and public clouds but have compatible infrastructures with a single *Common Access Gateway* for transparent access to computational and storage resources. As an example, in Figure 6, OpenNebula is used as a private cloud and Amazon EC2 and S3 are used as public cloud infrastructures. In the private cloud, the *Front End* node is an

instance of the *Controller* component that can also use discovery service(s) such as OGC CSW. At peak demand time, additional resources can be acquired from public clouds like Amazon EC2, S3. This hybrid cloud infrastructure facilitates both efficient utilisation of existing organisation IT resources in a secure environment and outsources additional computation/storage requirement to a public cloud infrastructure. Amazon specific VMs (i.e. AMIs) with specific software platforms must be available in the *Resource Pool* in order to deploy dynamically when required. However, such interaction also requires interoperability between cloud infrastructures. In the above example both Open Nebula and Eucalyptus are compatible with Amazon AWS.

Information intelligence and decision making model
Yigitcanlar T [14] emphasised the need for a virtual decision making environment using web-based Geographical Information System (GIS) integrated with other technologies such as multimedia, virtual reality or visualisation. The same author [14] also argued that such an environment can facilitate wider public participation in decision-making processes and development of a sustainable and smart urban environment. In this paper we propose suitable technologies and compliance to various standards in order to achieve similar objectives. The main aim here is to use data from various sources, to apply processing models, to visualise results, to build a knowledge-base and to use intelligence for decision-making and exchange outcomes. Using a knowledge pyramid Figure 7 depicts the main value-added elements
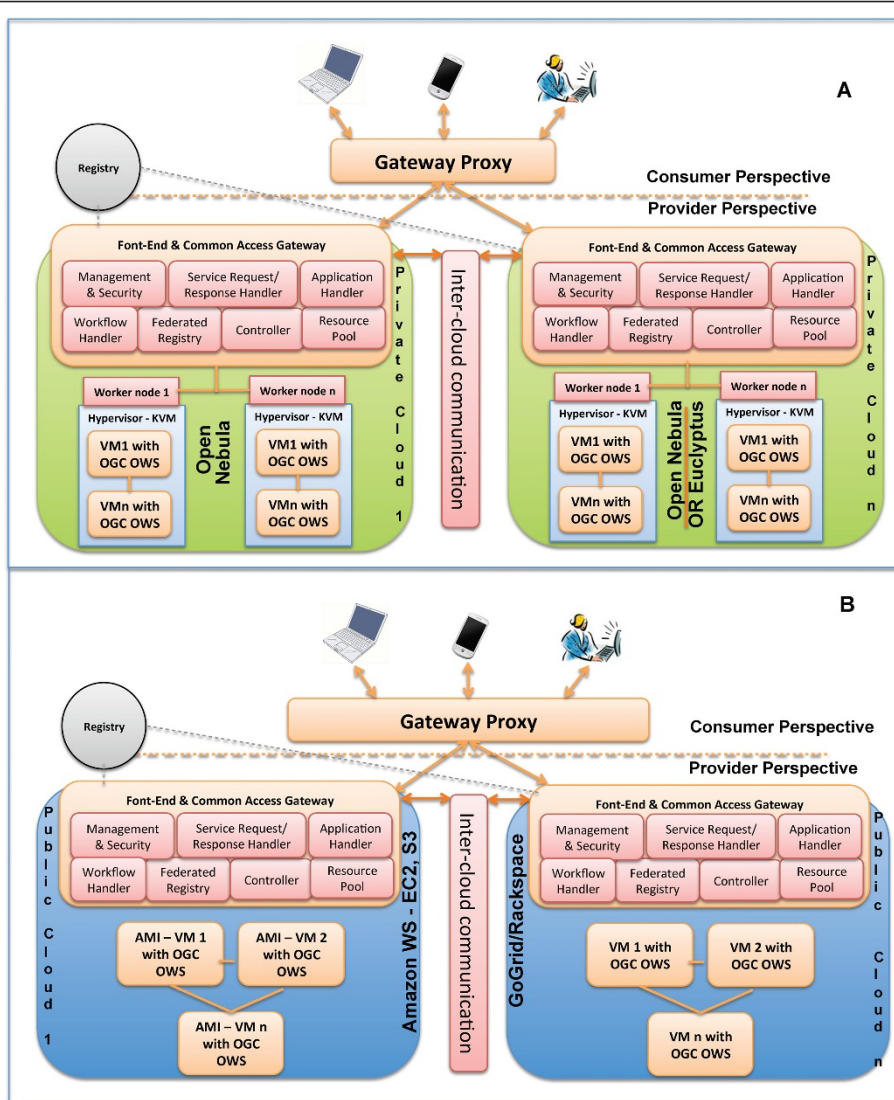
Figure 5 Private Multi-Cloud and Public Multi-Cloud Architecture Scenarios.

for decision-making models. The key to transform data (structured and unstructured) into information and intelligence is to use WPS with specific knowledge discovery algorithms. Such information can be used as an input to a decision making process in order to consider alternative scenarios and discover the best-fit outcomes. Typical steps in the decision making process include i) information collection; ii) analysis and decision alternatives; iii) impact assessment; iv) decision-making and communication; and v) process refinement.

OGC uses three fusion categories which collectively contribute towards decision-making processes. These categories are: i) *Sensor fusion*, which mainly deals with image change detection and feature extraction e.g. same location different time, shape identification; ii) *Feature/Object fusion* which aims for conflation of attributes in

geometries, e.g. entity aggregation (tabular data and GIS map) and relational analysis for change detection; and iii) *decision fusion* that uses data e.g. geo-data such maps and demographics with the aim to identify meanings and purposes in a specific context [34]. OGC's decision fusion model enables a user to acquire processed information based on different types of required data (structured e.g. relational, object oriented data models using OGC City GML, KML etc. and/or unstructured e.g. wikis, blogs, YouTube videos using GeoRSS, GeoJSON etc.) from various sources (databases, sensors, social-networking, WWW, etc.) using specific interoperable IT tools and services. Furthermore, using OGC standards different fusion process models can be adopted such as the Joint Directors of Laboratories (JDL) Fusion model, the Observe Orient Decide Act (OODA) loop and Multi-INT [34].
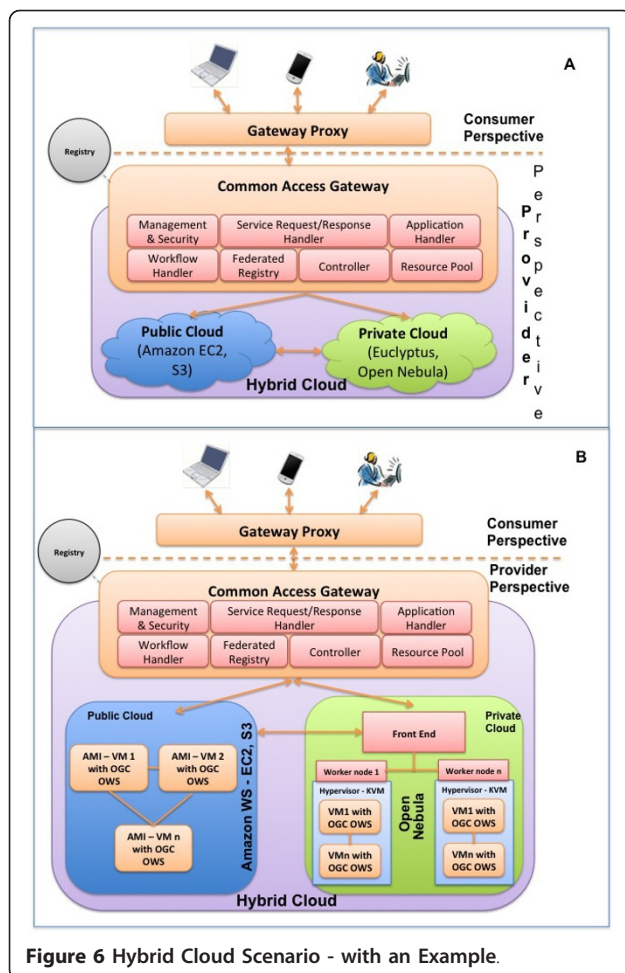
Figure 6 Hybrid Cloud Scenario - with an Example.

Figure 7 Integrated information for decision-making.

OGC's vision for decision fusion fits nicely with the acquisition of information intelligence for urban management in a cloud environment. Pravia MA et al. [35] provide a comprehensive multi-INT model for an urban situation divided into soft and hard sources. A service oriented environment with necessary capabilities as mentioned above (mapping user to technological perspective section) are in line with OGC's service platform for decision fusion [34]. A cloud environment can provide a decision-making platform at the PaaS layer that can compile non-spatial, spatial and temporal data from various sources in an urban environment (e.g. hard sources: data coordinates, signal frequency and location, video, imagery, buildings, land-use, event occurrence, socio-economic, demographic etc. and soft sources: social-networks on web, etc.) for a specific context. Furthermore, such a decision-making platform can be utilised for domain specific information intelligence at the SaaS layer.

Assessment and reflection on the IEMS architecture
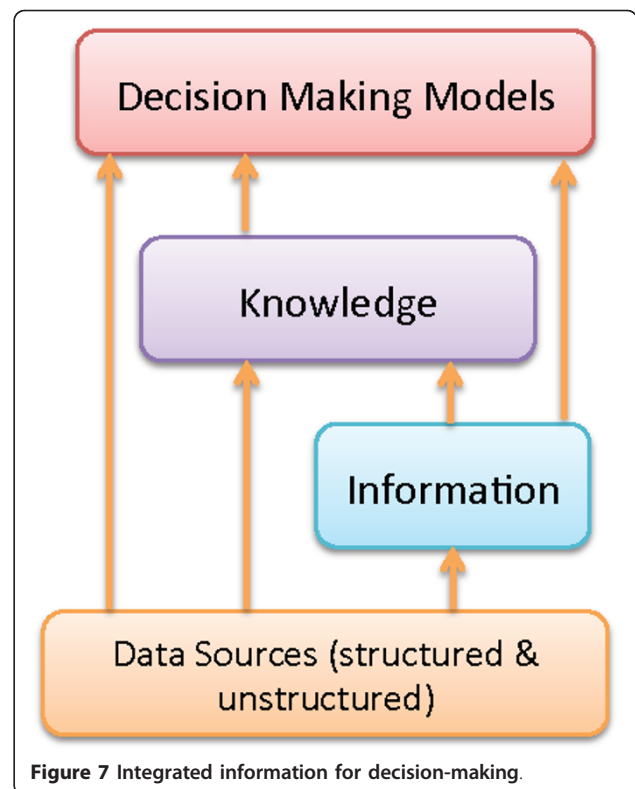The IEMS layered architecture has the following benefits, but also reveals several core challenges which require

further investigation. In addition, we critically reflect on the extent to which this IEMS architecture is suitable for the objectives stated in the introduction section.

Benefits of the layered IEMS architecture
The layered design of the architecture adopts a 'separation of concerns' as a core design principle in order to avoid reinventing the wheel and to capitalise on the experience and knowledge from existing projects, tools and services. The layered architecture also brings extensibility and flexibility since new components can be added at various layers. Likewise outputs from different layers can be tailored to the specification of various tools and utilised in different applications. Furthermore, a service oriented model in a cloud environment provides flexibility in using both standard web services for wider adoptability and workflow composition as well as OGC standards to manage substantial amounts of spatial data and geoprocessing services.

For example, the architectural design of the IEMS can adopt the ORCHESTRA Architecture Reference Model (OA-RM) [36] based approach in a cloud environment as it can provide a useful specification for data and service interoperability in a distributed environment using both ISO, OGC and web service standards. Similarly, for spatial data harmonisation there exist several tools and services which can be reused for the above purpose, e.g., the HUMBOLDT project [37] provides various INSPIRE

and OGC compliant state-of-the-art harmonisation tools and services. However, such tools require investigation to be utilised in a cloud environment.

The use of cloud computing provides high performance and scalability, transparent access to on-demand resources and low up front investment costs for integrated information intelligence for urban management that can be used for decision-making and other purposes.

Technological challenges

The implementation of the IEMS architecture is not straightforward as it poses the following challenges:

1) The layered architecture provides flexibility in managing functionalities at individual layers. However, managing the complexity of various layers and integrating outputs from one layer to other in a (semi-)automated process chain is not straightforward and requires rigorous management of tasks at various layers. This complexity increases further when multi-cloud architectures are implemented (Figure 5) and hence a management and integration service must be introduced to handle this complexity.

2) To provide a virtual environment for various stakeholders e.g. citizens, research, scientific and policy making communities, enabling them to select multiple variables across various environmental attributes (e.g. urban, transport, biodiversity and health) and to extract corresponding harmonised information from a cloud infrastructure which could be used for various purposes such as modelling, simulations, environmental indicators and decision making support for policy development. In this regard, access to environmental data from different distributed repositories is required. However, these distributed data sources may have different technological infrastructures including proprietary systems, desktop based systems, web services, grids, clouds, and non-standardised data models, which makes this task more challenging and requires rigorous analysis for data accessibility, harmonisation and transformation and compliance to policy requirements.

3) To support the uptake of cross-thematic harmonisation of selected environmental thematic areas such as land-use, transport and health according to existing standards such as the INSPIRE directive. The implementation of a thematic layer (Figure 3) in a cloud infrastructure based on INSPIRE components including implementation rules, metadata, data specification, services, monitoring and reporting will be challenging due to the recent specification of INSPIRE annexes.

4) To enable various stakeholders to visually design workflows (e.g. using BPMN and BPEL), to compose and bind services for execution, to identify data sources and to execute these workflows in a (multi-)cloud environment. In addition, this requires the development of provenance mechanisms to keep track of data and service sources and intelligent processing performed for future references.

5) To develop case studies around citizen science i.e. to enable citizens to participate in environmental monitoring in order to raise their awareness and responsiveness towards the management of the urban environment using tools such as smart phone applications (for participatory sensing) and web 2.0 (for collaborative decision-making). This would also require the acquisition of data using standard interfaces such as OGC's sensor web enablement e.g. sensor observation service (SOS) and OGC's WFS and visualisation using OGC's WMS in a cloud environment.

6) As indicated above, data may originate from two different cloud infrastructures. However, each cloud infrastructure is unique and mostly incompatible with each other i.e. the underlying cloud architecture, data models and access mechanisms, and services vary from one cloud infrastructure to another. Such an incompatibility introduces a gap in the acquisition of integrated intelligence across the platforms not only for applications from the same domain but also for cross-disciplinary applications such as environment, health and economic development. Similar to OGC, one of the aims of the open cloud consortium (OCC) [38] is to support the development of standards for cloud computing and frameworks for interoperability between clouds. This will enable cross platform interoperability, consistent security mechanisms and content sharing. Using such interoperability initiatives, the inter-cloud cross-thematic urban management dimension can bring novelty and a unique perspective in cloud computing research and, at the same time, will benefit both scientific, policy making and business communities in exploiting the full potential of clouds.

7) As indicated in other studies, the hybrid clouds can be used for OGC standards and in particular WPS. Further investigation is needed to test the behaviour of OGC and standard web services integration in a hybrid cloud environment.

Reflection and critical analysis

The real significance of integrated information intelligence is to be able to answer science and policy questions which cannot be answered directly from fragmented information sources. For example, in relation to urban and regional governance, integrated intelligence offers the opportunity to genuinely support evidence based practice and to provide urban planners with the tools and methodologies that will enable them to effectively manage the complexity of the city. This will enable different stakeholders (environmentalists, urban planners, policy makers and citizens), who cannot deal with the complexity of increasing data, harmonisation, service interoperability and processing needs, to gain benefit from this work.

The proposed IEMS architecture and implementation scenarios attempt to provide the necessary generic capabilities in order to fulfil the requirements for the development of an integrated intelligence system for monitoring

of the urban environment. Despite the fact that there are several tools which can provide specific technical capability, integrating and using these tools in a cloud environment in compliance to different standards is not straightforward and leads to several technological challenges as discussed in this section.

The benefit of a cloud environment is twofold. Firstly the use of OGC and standard web services for data acquisition, processing and sharing at the PaaS level provides the opportunity for various stakeholders to develop context based domain specific applications at the SaaS level. For example, these services at the PaaS level in a cloud environment encapsulate the complexity of data acquisition, cross-thematic harmonised transformations, computer intensive processing, multi-dimensional modelling and visualisation and collaborative decision support mechanisms for various stakeholders. Secondly the extensibility and scalability characteristics of cloud platforms will accommodate continuously increasing data volumes and caching for visualisation and user groups who can be involved for citizen science-based participatory environmental monitoring.

One of the critical aspects that the above research attempts to address is building the capacity and ability of a community in creating sustainable futures, and to be involved in community driven environmental decision-making processes. For example, citizens can have access to informed environmental intelligence affecting citizens' daily lives and regional economic prosperity which is directly or indirectly affected by environmental well being. Similarly, urban planners can obtain improved predictions of impacts of urban planning policy by incorporating multiple variables such as land-use, socio-economic, transportation and public health. Furthermore, the processed output can be presented in various forms that could be used to indicate the effectiveness of environmental changes at various levels (e.g., local, national and European) with benefits and consequences to stakeholders, policy and decision-makers and the general public. However, rigorous development efforts are required from technological peers to develop solutions at the SaaS level using the identified technical capabilities and a proposed IEM architecture to achieve the above aims.

Conclusions

In this paper, we have argued that the integration of fragmented environmental information can result in better environmental monitoring in order to mitigate various environmental challenges such as issues related to climate change. Furthermore, we argued that cloud computing can play a major role in achieving environmental information integration by providing on-demand processing and storage capabilities. Our urban management based example assists in identifying a generic set of

technical capabilities for information intelligence and proposes a layered architecture for IEMS using different cloud based implementation scenarios. However, based on our preliminary assessment we argue that it is not a simple matter to develop integrated intelligence to its full potential as a result of certain technological challenges. These challenges require rigorous investigation in future research in realising the true potential of the proposed cloud based architecture for integrated information intelligence and decision making in environmental domain.

Author details
[1]Faculty of Environment and Technology, University of the West of England, Bristol, UK [2]School of Computing and Mathematics, University of Derby, Derby, UK

Authors' contributions
ZK contributed to the ideas of applying cloud computing for integrated information intelligence for urban management. He participated in identifying user information needs; carried out mapping of user needs to technical capabilities; proposed cloud implementation scenarios using web and OGC standards; and critically assessed benefits and challenges for integrated monitoring and decision making models using cloud environment. DL contributed to elaborate user's perspective in an urban environment and linked it with policy context in order to determine the required technical capabilities in the integrated urban management. RM contributed in defining the proposed architecture of the IEM based on the required technical capabilities and participated in the assessment of IEMS. AA participated in mapping the user needs into technical processes, tools and identified capabilities required for the integrated information intelligence in cloud environment. He also contributed to the assessment of the IEM. All authors read and approved the final manuscript.

Authors' information
Dr. Zaheer Khan is a Postdoctoral Research Fellow in the Faculty of Environment and Technology of UWE, Bristol, UK. He has over 7 years of teaching and research experience in system analysis and design, databases, distributed systems, etc. He has been actively working in large-scale EU and non-EU collaborative computing research projects: FP6 Health-e-Child, FP6 HUMBOLDT, FP7 LifeWatch and SAGE. Currently he is working on the FP7 UrbanAPI project dealing with ICT applications for urban governance. His main research interests are in distributed computing in particular cloud computing, software engineering, business process management, data interoperability and management, software agents and geospatial information systems.
Mr. David Ludlow is an urban planner with extensive practice experience gained in UK land-use planning agencies, and at UWE as project manager for more than 30 EU funded projects focused on integrated urban management information systems, and sustainable urban development. He is a member of the EU Expert Group on the Urban Environment (European Sustainable Cities Project), and the Expert Group for the EU Thematic Strategy on the Urban Environment (6th Environment Action Programme). Currently he is UWE project manager for EU FP7 projects urbanAPI and urbanNexus, the ESPON SEGI project, and urban expert for the European Topic Centre Spatial Information Analysis (European Environment Agency). Professor Richard McClatchey currently leads the Centre for Complex Cooperative Systems at UWE and is active in collaborative computing projects at CERN, and with many international partners in other EC projects: Health-e-Child, SHARE, neuGRID and N4U. He has been research active for the past 30 years and has led many projects funded by industry and by the EC in the areas of large-scale distributed data and process management and database modelling and in systems design and integration.
Dr. Ashiq Anjum is a senior lecturer in the School of Computing and Mathematics at the University of Derby. His areas of expertise include Distributed, Parallel and Concurrent Systems (mainly Cloud and Grid Computing). He has worked on a number of scientific and industrial projects (notable contributions include FP7 funded TRASNFORM project, FP7 funded

neuGrid project, FP6 funded Health-e-Child project, NSF/DOE funded CAIGEE project and US Aid funded GAE project) in collaboration with leading industrial and academic partners in Europe and USA. He has 12 years of experience in academic and industrial research and has around 60 publications to his credit.

Competing interests
The authors declare that they have no competing interests.

References
1. OGC (2011) Open Geospatial Consortium. http://www.opengeospatial.org/. Accessed 21 Nov 2011
2. INSPIRE (2011) INfrastructure for SPatial InfoRmation in Europe. http://inspire.jrc.ec.europa.eu/. Accessed 21 Nov 2011
3. , OGC Standards and Cloud Computing, OGC White Paper
4. GEOSS (2011) Global earth observation system of systems. http://www.earthobservations.org/geoss.shtml. Accessed 21 Nov 2011
5. GENESI-DEC Project (2010) http://www.genesi-dec.eu/. Accessed 21 Nov 2011
6. SEIS (2011) Shared environmental information system. http://ec.europa.eu/environment/seis/. Accessed 21 Nov 2011
7. GMES (2011) Global monitoring for environment and security. http://www.gmes.info/. Last Accessed 21 Nov 2011
8. Furht B, Escalante A (eds) (2010) Handbook of Cloud Computing, ISBN 978-1-4419-6523-3. Springer-Verlag, New York Inc
9. Arthur C (2011) The Technology Waiting for Whitehall, Cloud Technology Government Policy. Media Guardian: The Guardian Newspaper, UK. p 6http://www.guardian.co.uk/cloud-technology/government-it-technology-spending-cloud-computing. Accessed 20 Feb 2012
10. OCCI (2011) Open cloud computing interface. http://occi-wg.org/. Accessed 10 November 2011
11. ESS (2011) Environmental software and services. http://www.ess.co.at/. Accessed 22 Nov 2011
12. EVOp Project (2010) Environmental virtual observatory pilot. http://www.environmentalvirtualobservatory.org/. Accessed 22 Nov 2011
13. IUME (2011) Integrated urban monitoring in Europe, EEA initiative. http://iume.ew.eea.europa.eu/. Accessed 22 Nov 2011
14. Yigitcanlar T (2009) Planing for smart urban ecosystems: information technology applications for capacity building in environmental decision making. Theor Empirical Res Urban Manage 3(12):5–21
15. Sl@M (2010) Environmental Atlas - Vitoria-Gasteiz. http://www.vitoria-gasteiz.org/w24/en/html/14/531.shtml. Accessed 10 Sept 2011
16. ISO/IEC 10746-1 (1998) Information Technology - Open Distributed Processing - Reference Model.
17. Khan Z, Odeh M, McClatchey R (2011) Bridging the gap between business process models and service oriented architectures with reference to grid environment. Int J Grid and Utility Comput 2(4):253–283. doi:10.1504/IJGUC.2011.042943.
18. Anjum A, Bloodsworth P, Branson A, Habib I, McClatchey R, Soomro K, Munir K "Provenance data management in a workflow application for studies of Alzheimer's Disease". In Press at the International Journal of Computers and their Applications (IJCA). Special Issue on Scientific Workflows, Provenance and Their Applications
19. Lee C, Percivall G (2008) Standards-based computing capabilities for distributed geospatial applications. IEEE Computer Computer Society 41(11):50–57
20. Maué P, Kiehle C (2009) Grid technologies for geospatial applications: an overview. GIS Sci 3:65–67
21. Di L, Chen A, Yang W, Zhao P (2003) The integration of grid technology with OGC Web Services (OWS) in NWGISS for NASA EOS data. Proceedings of Global Grid Forum (GGF) 8 and High Performance Distributed Computing (HPDC) 12. Seattle, Washington
22. G-POD (2011) Grid processing on demand, European space agency's earth observation platform. http://gpod.eo.esa.int/. Accessed: 14 Nov 2011
23. OGC WS - Phase 5 (2008) OGC's interoperability programme. http://www.opengeospatial.org/projects/initiatives/ows-5. 15 Nov 2011
24. Hobona G, Fairbairn D, Hiden H, James P (2012) Orchestration of grid-enabled geospatial web services in geoscientific workflows. IEEE Trans Autom Sci Eng 7(2):407–411
25. Aydin G, Sayar A, Gadgil H, Aktas MS, Fox GC, Ko S, Bulut H, Pierce ME building and applying geographical information system grids, special issue on geographical information systems and grids based on GGF15 workshop, concurrency and computation: practice and experience.http://grids.ucs.indiana.edu/ptliupages/publications/GISGrids_Concurrency_submitted.pdf. Last Accessed: 10/11/2011
26. Aydin G, Pierce M, Fox G, Aktas M, Sayar A (2004) "Implementing GIS Grid Services for the International Solid Earth Research Virtual Observatory", Extended abstract for Fourth ACES (APEC Cooperation for Earthquake Simulation) Workshop. Beijing China
27. Aktas MS, Aydin G, Fox G, Gadgil H, Pierce M, Sayar A (2004) "Information services for grid/web services oriented architecture (SOA) based geospatial applications, CGL technical report, June 2005. http://grids.ucs.indiana.edu/ptliupages/publications/UDDI_Aktas_Final_Fix.pdf. Accessed on: 10 November 2011
28. Blower JD (2010) GIS in the cloud: implementing a web map service on google app engine. Washington, D.C pp 1–4 In Proceedings of the First International Conference on Computing for Geospatial Research and Application
29. Baranski B, Schäffer B, Redweik R (2009) Geoprocessing in the clouds. Sydney, Australia In Proceedings of the Free and Open Source Software for Geospatial (FOSS4G) Conference
30. Schäffer B, Baranski B, Foerster T (2010) Towards spatial data infrastructures in the clouds. In Geospatial Thinking: Proceedings of the Thirteenth AGILE International Conference on Geographic Information Science, Guimarães, Portugal. Berlin: Springer Verlag Lecture Notes in Geoinformation and Cartography pp 399–418
31. Foerster T, Baranski B, Schaeffer B, Lange K (2010) Geoprocessing in hybrid clouds. In: Zipf A, Behncke K, Hillen F, Schaefermeyer J (ed) Die Welt im Netz. Kiel, Germany pp 13–19. Presented at the Geoinformatik 2010
32. Baranski B, Foerster T, Schäffer B, Lange K (2011) Matching INSPIRE quality of service requirements with hybrid clouds. Trans GIS 15(s1):125–142. doi:10.1111/j.1467-9671.2011.01265.x
33. Metsch T, Edmonds A, Bayon V (2010) Using cloud standards for interoperability of cloud frameworks, technical report. http://sla-at-soi.eu/wp-content/uploads/2010/04/RESERVOIR-SLA@SOI-interop-techReport.pdf. Accessed 21 Nov 2011
34. OGC Fusion Standards Study (2010) Phase 2 Engineering Report. George P. 2010
35. Pravia MA, Prasanth RK, Arambel PO, Sidner C, Chong CY (2008) Generation of a fundamental data set for hard/soft information fusion, fusion 2008: the 11th international conference on information fusion. Cologne: international society of information fusion http://isif.org/fusion/proceedings/fusion08CD/papers/1569108037.pdf. Accessed: 21 Nov 2011
36. ORCHESTRA project (2008) http://www.eu-orchestra.org/. Accessed 20 Sept 2011
37. HUMBOLDT Project (2010) http://www.esdi-humboldt.eu/home.html. Accessed 24 April 2011
38. OCC (2011) Open cloud consortium. http://opencloudconsortium.org/. Accessed 23 Nov 2011