

LEOpatra: A Mobile Application for Smart Fertilization Based on Linked Data

Stefan Burgstaller¹, Wolfgang Angermair¹, Fabian Niggemann², Silke Migdall²,
Heike Bach², Ioannis Vlahopoulos³, Dimitrianos Savva³, Panayiotis Smeros⁴, George
Stamoulis³, Konstantina Bereta³, Manolis Koubarakis³

¹ FarmFacts GmbH, Pfarrkirchen, Germany
{Burgstaller, Angermair}@farmfacts.de

² VISTA GmbH, Munich, Germany
{Migdall, Bach, niggemann}@vista-geo.de

³ National and Kapodistrian University of Athens, Greece
{johnvl, dimis, gstam, konstantina.bereta, koubarak}@di.uoa.gr

⁴ EPFL, Lausanne, Switzerland
panayiotis.smeros@epfl.ch

Abstract. We present LEOpatra, a mobile application for smart fertilization. The focus of LEOpatra is the variable rate application of fertilizers in a field based on information obtained from satellite images and other geospatial data. LEOpatra is able to identify restricted zones where no or only a decreased amount of fertilizers is allowed. LEOpatra has been developed in FP7 project LEO and uses linked data as the basic technology for integrating the various relevant data sources. LEOpatra has been evaluated by German farmers in trials that took place in Bavaria.

1 Introduction

In this paper, we describe LEOpatra (Linked Earth Observation precision agriculture tractor application), a mobile smart fertilization application that was developed in the context of the EU FP7 project LEO¹. The focus of LEOpatra is the variable rate application of fertilizers in a field based on information coming from satellite images and other geospatial data. LEOpatra is able to identify restricted zones where no or only a decreased amount of fertilizers is allowed. LEOpatra offers field management support for farmers in a low-cost way, utilizing an easy-to-use smart farming solution, which can run on any mobile device (tablet or smartphone) with Android OS.

The main technology for integrating the various geospatial data sources in LEOpatra is linked data. *Linked data* is a new data paradigm which studies how one can make RDF data available on the Web, and interconnect it with other data with the aim of increasing its value [HB11]. In the last few years, linked *geospatial* data has received attention as researchers and practitioners have started tapping the wealth of geospatial information available on the Web [KKK+12]. As a result, the *linked open data (LOD) cloud* has been rapidly populated with geospatial data some (e.g., Open-

¹ <http://linkedgeodata.org/>

StreetMap data, administrative boundaries of various countries, land use/land cover data etc.).

Another data paradigm exploited by LEOPatra is *open Earth Observation (EO) data*. Petabytes of open EO data has recently become freely available due to the Copernicus program of the European Union and the Landsat program of the US. LEOPatra employs and extends EO data processing methods that were developed within the ESA project **talkingfields**². The processed satellite data forms the basis of the prescription maps produced by LEOPatra and get correlated with datasets that contain other useful information, such as restricted zones in order to provide variable rate application of fertilizers.

Open EO data that are currently made available by the Copernicus and Landsat programs are *not* following the linked data paradigm. Therefore, from the perspective of a user, the EO data and other kinds of geospatial data necessary to satisfy his or her information need can only be found in different *data silos*, where each silo may contain only part of the needed data. Opening up these silos by publishing their contents as RDF and interlinking them with semantic connections will allow the development of data analytics applications with great environmental and financial value.

The European project TELEIOS³ was the first project internationally that has introduced the linked data paradigm to the EO domain, and developed prototype applications that are based on transforming EO products into RDF, and combining them with linked geospatial data. Examples of such applications include wildfire monitoring and burnt scar mapping, semantic catalogues for EO archives, and rapid mapping. The wildfire monitoring application is the basis of the FIREHUB service at the National Observatory of Athens⁴ which has been used operationally by government agencies in Greece since the summer of 2012. FIREHUB was awarded the Best Service award in the Copernicus Master's competition of 2014.

TELEIOS concentrated on developing data models, query languages, scalable query evaluation techniques, and efficient data management systems that can be used to prototype applications of linked EO data. However, developing a methodology and related software tools that support the whole lifecycle of linked open EO data (e.g., publishing, interlinking etc.) has not been tackled by TELEIOS. The main objective of the subsequent European project “Linked Open Earth Observation Data for Precision Farming” (LEO) was to go beyond TELEIOS by designing and implementing software supporting the whole life cycle of linked open EO data and its combination with linked geospatial data, and by developing a smart farming application that heavily utilizes such data. This application is LEOPatra which we present on this paper.

The rest of the paper is organized as follows. Section 2 describes the main components of LEOPatra and Section 3 presents its functionalities. Section 4 describes related work in the area of precision farming. Finally, Section 5 concludes the paper.

² <https://artes-apps.esa.int/projects/talkingfields>

³ <http://www.earthobservatory.eu/>

⁴ <http://195.251.203.238/seviri/>

2 EO Data processing and LOD for Precision Farming

LEOpatra is implemented as a software solution which integrates satellite data with linked geospatial data and mobile technologies in order to provide assistance to farmers during the fertilization of their fields. LEOpatra is offered as a low-cost tool for small and medium farmers in order to boost rural development.

Figure 1 shows an overview of the architecture of the smart fertilization application implemented by LEOpatra.

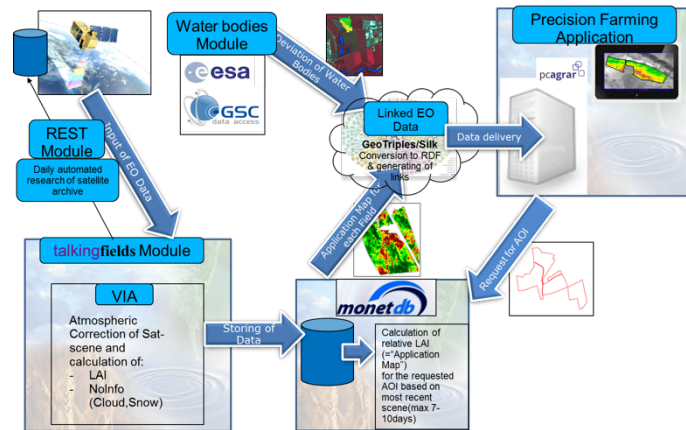


Fig. 1. The EO-based processing chain for smart fertilization

Starting with the REST (Representational State Transfer) modules shown in the figure, satellite data is downloaded automatically as soon as it is available for some pre-defined search criteria (covering the area of interest). Then, all downloaded satellite scenes are automatically processed by the **talkingfields** module for further use.

Processing includes the following tasks: atmospheric correction of the data, classification of different land uses and calculation of Leaf Area Index (LAI) for vegetation areas as an indicator of relative biomass variation within a certain area. The storing, querying and post processing of all EO-based data have been implemented in the well-known DBMS MonetDB.⁵

After a request for a specific area (e.g. given field boundaries) by a user of LEOpatra, the available latest data of LAI within the specific field is retrieved. The results are converted to RDF by using the tool GeoTriples and get interlinked with other open data sources using the tool Silk. The results are available to the users via a dedicated server operated by FarmFacts. In the following sections we describe in more details the components of the processing chain that that realizes LEOpatra.

⁵ <https://www.monetdb.org/Home>

The REST module

The REST module guarantees an up to date satellite data basis, via an automated way of downloading and processing satellite scenes. The first step in this process is to combine two ways of automated satellite data download. The first way is BlackBridge, a private geospatial data provider, which is offering an interface to search and download RapidEye satellite data automatically. Additionally, a tool for automated command line download of Landsat data has been adapted to address the needs of the application. This tool⁶ has been adapted with the ability to set temporal and spatial restrictions, allowing users to retrieve and automatically download, for example, if new satellite scenes are available in a given interval for a specific area. The downloaded scenes can optionally (i.e., if selected so in the parameter file) be processed automatically by the *talkingfields* Module. Landsat data haven been used instead of Sentinel data since the Sentinel satellites had not been launched when the application was developed. The current version of the module uses Sentinel 2 data from the Copernicus Open Access Hub.

Talkingfields Module & Vista Image Analysis (VIA) chain

After satellite data is made available by the REST module, the talkingfields module with its processing chain named starts performing the following steps automatically:

- All satellite scenes (Level 1b) are pre-processed including an advanced atmospheric and geometrical correction delivering bottom-of-atmosphere spectral reflectance values (Level 2b). To get bottom-of-atmosphere measurements needed for comparability of different scenes, the atmospheric transfer model MODTRAN (MODerate resolution atmospheric TRANsmis-sion) is used. The results of MODTRAN are used to correct the image to bot-tom-of-atmosphere values.
- To identify relevant land use classes, an unsupervised land use classification is performed using the unique spectral characteristics of each surface. The seven land uses are classified resulting in a so-called Level 3-product. This step is vitally important to control the quality of produced information (e.g. no delivery of false information from areas with cloud shadows).
- By using the radiative transfer model SLC (Soil-Leaf-Canopy) with an in-verse modelling technique, precise information about plant physiological pa-rameters for all relevant land uses is then derived. SLC is simulating typical spectral response curves for given physiological plant parameters (e.g. water content, nitrogen content, Leaf Area etc.). Adapting the simulated to satel-lite-observed spectral response curves is resulting in the needed information about the land use class “vegetation”. For LEOpatra, LAI (Leaf Area Index) values produced by SLC, for the relevant crop areas are extracted.

⁶ The tool for downloading Landsat data is developed by Olivier Hagolle and it is available on the USGS-server under the GNU general public license.

MonetDB

The column store database MonetDB is used to store and query the LAI-values for LEOpatra. Using a database management system like MonetDB offers advantages in performance of querying the needed products as follows:

- Having the most recent LAI values for each satellite image at hand, the next step is the request by the farmer for one (or several) specific fields (more details of the information exchange will be described in the next chapter).
- Receiving the request on the Vista side by the FarmFacts server results in the reading of the field boundaries from MonetDB and extraction of the relevant LAI values.
- In a next step, the relative LAI values (compared to the mean value of the field as deviation in percent) are calculated for each field. This serves as biomass indicator and basis for the prescription map.

MonetDB has an multidimensional array management module, i.e., it offers functionalities that allow developers to handle satellite images and EO products as arrays, and query them using the query language SciQL [ZKM13]. SciQL is an extension of the query language SQL. SciQL queries are typically used to encode image processing steps such as the ones presented above, instead of using complicated, customized code in a programming language.

Tools for linked geospatial data

In order to combine EO data with other geospatial sources, all data need to follow a common format. For this purpose, we use the RDF data model, which is a machine readable format for describing resources that are available on the Web as linked data. These resources can be queried in a uniform way using the query language SPARQL. Since the framework of RDF and SPARQL does not provide geospatial support, extensions of it were proposed recently. One of these extensions is the framework of stRDF and stSPARQL, an extension of RDF and SPARQL with spatial and temporal support proposed by our group [KKK12, BSK13]. Another extension of the query language SPARQL with geospatial support is the OGC standard GeoSPARQL. In the following, we briefly describe the tools that we use for the conversion and processing of EO and geospatial data as linked data, and then we describe how this is achieved.

GeoTriples. GeoTriples is a tool for transforming vector data and their metadata into RDF, and to support natively many popular geospatial data formats (e.g., shapefiles, spatially enabled DBMS, KML, GeoJSON, etc.). GeoTriples produces an RDF dump of the data in accordance with a certain ontology, and a set of relevant mappings. The mapping generator employs and extends the mapping languages R2RML and RML to create mappings that dictate the method of conversion of the raw data into RDF. y. GeoTriples is a free and open source tool⁷.

⁷ <https://github.com/LinkedEOData/GeoTriples>

Strabon

Strabon [KKK12, BSK13] is a spatiotemporal RDF store. It enables users to store geospatially-enhanced RDF files and query them using the query languages stSPARQL and GeoSPARQL. Strabon is freely available as open source software (<http://strabon.di.uoa.gr>). According to recent benchmarks [GKK13, BSK13], Strabon is the most efficient and rich in functionalities spatiotemporal RDF store.

Silk

We developed a spatiotemporal extension of the interlinking tool Silk [SK16] that addresses the problem of discovering other kinds of geospatial or temporal semantic links. For example, in linked EO datasets, it is often useful to discover links involving topological relationships e.g., A geo:sfContains F where A is the area covered by a remotely sensed multispectral image I, F is a geographical feature of interest (field, lake, city etc.) and geo:sfContains is a topological relationship from the topology vocabulary extension of GeoSPARQL. The existence of this link might indicate that I is an appropriate image for studying certain properties of F. Our work extends the link discovery tool Silk to be able to discover precise geospatial and temporal links among geospatial RDF data [SK16 and it is now integrated into the default branch of Silk⁸ which is available as open source.

The above tools are used in the processing chain of LEOpatra as follows. First, an ontology is constructed for all the data sources that we want to integrate. An *ontology* is a conceptual way of encoding the schema of the data. Based on the ontology, the data is transformed into RDF using the tool GeoTriples. Then, all datasets are stored in the spatiotemporal RDF store Strabon⁹.

Table 1. GeoSPARQL query to present all the fields of a farm and their fertilization measures

```
SELECT ?RasterCell ?wkt ?hasCV ?hasFertV
WHERE {
  ?Field rdf:type tf:Field .
  ?RasterCell rdf:type tf:RasterCell .
  ?RasterCell tf:belongsToField ?Field .
  ?RasterCell geos:hasGeometry ?geom .
  ?geom geos:asWKT ?wkt .
  ?Field tf:belongsToFarm tf:002 .
  OPTIONAL { ?RasterCell tf:hasCV ?hasCV.} .
  OPTIONAL { ?RasterCell tf:hasFertV ?hasFertValue.} } }
```

Table 2. GeoSPARQL query to present all the fields of a farm with their fertilization measures

```
SELECT ?wktField ?labelField ?nameField
?waterBody ?nameOSM
WHERE {
  ?waterBody geos:close ?Field .
  ?Field rdf:type tf:FieldB .
  ?Field geos:hasGeometry ?geom .
  ?geom geos:asWKT ?wktField .
  ?Field rdfs:label ?labelField .
  ?Field tf:hasName ?nameField .
  ?waterBody rdf:type osm:OSMFeature .
  OPTIONAL {
    ?waterBody osm:hasOSMId ?OSMId. } .
  OPTIONAL { ?waterBody osm:hasName
    ?nameOSM } }
```

⁸ <http://silkframework.org/>

⁹ <https://goo.gl/BCGN8P>

Next we perform interlinking using the tool Silk as shown in the queries of Table 1 and Table 2 respectively.

Water Bodies Module

The smart fertilization application takes into account information about the field, but also about the field's surroundings, the geospatial context of the field. With the help of LOD-tools developed within the LEO project (e.g., Strabon and Silk), information from different sources is integrated and is made available to LEOpatra. For example, for the identification of restricted zones, water bodies play a significant role. The extension of water bodies can be highly dynamic for various reasons, e.g., extensive rain, snowmelt or human water management activities. For this reason, it would be desirable to offer always the most recent and valid data for supporting the management activities of the farmer. We achieve this by correlating water areas derived by satellite data and the water bodies layer of OpenStreetMap (OSM).

High-resolution satellite images offer a way to map water bodies on up to date data basis. Through the Core Datasets of the GMES Space Component Data Access (GSCDA) system of ESA, a first attempt to use available datasets for deviation of water bodies was made in LEO. These datasets are not up to date but can be used to estimate the necessary spatial and spectral resolution for the purpose of water body deviation with the needed accuracy.



As can be seen in **Figure 2**, in general, bigger water bodies can be derived pretty well by using the SPOT scene with a Maximum Likelihood Classification. Taking a closer look reveals areas where the classification has accuracy problems (marked with bright green). Especially narrow riverbeds bordered by trees and bushes show gaps in the classification of the course of the river. Other accuracy problems can be found in shallow water areas. Nevertheless, SPOT Data can be used to classify water bodies in most parts in good accuracy and can be used as additional information source, especially for temporally varying water levels.

Although sometimes OSM data may be inaccurate due to the nature of crowd sourcing, a major advantage of this data source is that it is free and publicly available, also in RDF format provided by the project LinkedGeoData¹⁰.

Table 1 is comparing the OSM data in the area of a demo user with the deviation of water bodies from the SPOT5 scenes described in above. A summary of the advantages and disadvantages of the two analysed datasets is also presented.

¹⁰ <http://linkedgeo.org/>

Table 1. Comparison of EO data and OSM data for the identification of water bodies.

Water bodies from EO-data (SPOT5)	Water bodies from OSM-data
 <p data-bbox="331 898 786 976">Figure 2. Maximum Likelihood Classification of water bodies using a pan-sharpened SPOT5 scene</p>	 <p data-bbox="810 887 1265 943">Figure 3. Comparison with OSM Waterbodies layer</p>
Advantage	
<ul style="list-style-type: none"> • Identification of water bodies is actual and up to date: it can be used to identify the current extent of temporarily variable water bodies. • Usage of EO data: available information all over the world. It can be derived where other data sources are not available. 	<ul style="list-style-type: none"> • More accurate in certain areas, mainly rivers and shallow water. • Free and open availability of information. • Already provided in RDF-format.
Disadvantage	
<ul style="list-style-type: none"> • Identification of the extension of water bodies fail in accuracy. • Cost of high resolution EO data makes it not practicable to use this data source for the desired low-cost mobile application. 	<ul style="list-style-type: none"> • Some areas are falsely classified as water bodies: the OSM data can only be seen as indication for management decision. The information has to be checked in the “real world”.

As a conclusion for LEOpatra, the preferred solution is combining of the advantages of both datasets. To achieve this, we perform the following steps:

1. The water body information derived by EO data is converted into RDF using the tool GeoTriples.
2. The resulting RDF file is stored in the spatiotemporal RDF store Strabon, together with the RDF version of the OSM dataset that contains information about water bodies.
3. Then, we pose queries combining the two datasets.

4. In order to materialize the links between these two datasets, i.e., the information that encodes the topological relations that hold between the entities, we use the tool Silk.

3 The Mobile Application LEOpatra

LEOpatra was developed as a low-cost accessible tool, which supports the variable rate application of fertilizers in a field. The application can run on any mobile device with Android OS and it integrates linked open data sources in its functionality. These features make LEOpatra attractive and helpful for every farmer or consultant with focus on the site-specific fertilization.

To understand how LEOpatra works, it is necessary to know the current situation of the methods to apply fertilizers and plant protection products. The techniques with agricultural machinery are mainly dependent on the following parameters:

- **Machine Parameters:** The set of tractor and mounted implements plays a big role in the process. The tractor supplies the power to the power take-off (PTO) shaft to activate the distribution mechanism and the dosing system in the implement. The implement has a fixed working width to cover a specific surface during the application.
- **Product Parameters:** Each product has a specific set of properties like density, fluidity and granulometry. These properties depend on the nature of the product, but they remain constant if the product will not be changed during the application. The farmer can only set a range for the applied quantity according to the technical recommendations, but she does not have a way to control the target quantity on the field without sophisticated equipment.

Since interfaces between mobile devices (smartphones and tablets with iOS and Android operating system) and agricultural machinery for a direct control of the machinery do not exist until now, LEOpatra is only able to act supportively by providing appropriate decision support information to the farmer.

To make this possible, the mobile device has to obtain the necessary information about the non-adjustable parameters and can then visualize the adjustable ones. The driving speed is predetermined to be used as adjustable control variable, since the majority of the devices include a GPS (Global Positioning System) receiver on which the current speed can be determined. A power shift gear and a continuously variable transmission (CVT) offer tractors the possibility to adjust the speed without changing the PTO speed, so changing the driving speed will not influence the PTO speed. For these reasons, the technical conditions exist to regulate the applied fertilizer quantity exactly based on the driving speed while the other factors remain constant.

Next to this functionality, LEOpatra has fulfilled additional requirements so that it can be used optimally:

1. Usability/Design:
 - Users do not work daily with the application, so the application must be easy to use and intuitive.
 - Due to the strong acceleration forces in a tractor moving over rough terrain, the controls must be sufficiently large.

- The readability of the display by sunlight is difficult, which is why the contrasts within the application for the delimitation of the functional areas must be as large as possible.
 - Sources for linked open data are available online. Therefore, the exchange of data via mobile internet must be possible.
2. Versatility/Flexibility:
- Offline mode must be ensured because in rural areas mobile internet is not fast enough to be able to use web-based applications.
 - For software applications, latency times should not exceed 4-5 seconds. As mobile internet in rural areas cannot ensure the necessary transfer speed, we have to work with techniques that allow working without constant internet access.

The supporting function for smart fertilization is based on information for driving speed to achieve the desired result. All other factors influencing the quantities must be kept constant. The following information must be available to make a valid calculation concerning the optimum speed range: (i) target quantity, (ii) currently applied quantity (Calibration test), (iii) nutrient content and (iv) working width. The target quantity comes from a prescription map, which defines the application quantity in defined areas modeled as raster cells. As an additional service, a farmer can enter the pure nutrient content for the product she wants to apply, so that the correct product quantity can be calculated. The current quantity to apply will be calculated by a calibration test. For this, the fertilizer distributed from the machine in a minute at a defined PTO is collected and weighed.

To start the fertilization, the user taps on the field marked with a tractor. This mark means that the field has an assigned task. The next screen performs a zoom in the selected field and the user can click the “Play” button in the menu bar of LEOpatra, when this appears to start the process. Finally, the speed predictor appears and the support function runs during the fertilization, as shown in Fig. 4.

4 Related work

The ESA project TalkingFields¹¹ was a demonstration project that developed four different smart farming services based on EO data and integrated into the Farm Management Software of the farmers. The services developed within TalkingFields focus on site-characteristics for soil probing and basal dressing as well as on the monitoring of current biomass and yield. The TalkingFields services are operational since 2014 and provided commercially by VISTA and FarmFacts. Stimulus (N-Smart) was another project funded by ESA to further the use of cloud computing technologies within different Earth Observation fields. The ICT solution of Stimulus is based on a multi-mission EO web portal developed within APPS4GMES¹².

¹¹ www.talkingfields.de

¹² www.apps4gmes.com



Fig. 4. Example of the Speed Predictor (Smart Fertilization Functionality)

4 Conclusions

In this paper we presented a smart fertilization application named LEOpatra. The aim of LEOpatra is to enable the variable range application of fertilizers in a field using a low-cost mobile device. LEOpatra is based on state of the art technologies for processing satellite data to detect information relevant for fertilization, and linked data as the core data integration technology.

References

- [BFC+14] P. Blanchart, M. Ferecatu, S. Cui, and M. Datcu. *Pattern retrieval in large image databases using multiscale coarse-to-fine cascaded active learning*. In *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1127–1141, April 2014.
- [BSK13] Konstantina Bereta, Panayiotis Smeros, and Manolis Koubarakis. *Representation and querying of valid time of triples in linked geospatial data*. In the 10th Extended Semantic Web Conference (ESWC), 2013
- [GKK13] George Garbis, Kostis Kyzirakos, and Manolis Koubarakis. *Geographica: A benchmark for geospatial RDF stores (long version)*. In *The SemanticWeb – ISWC 2013*, volume 8219 of *Lecture Notes in Computer Science*, pages 343–359. Springer Berlin Hei-

- delberg, 2013.
- [HB11] Tom Heath and Christian Bizer (2011) *Linked Data: Evolving the Web into a Global Data Space* (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool.
- [KKK12] K. Kyzirakos, M. Karpathiotakis, and M. Koubarakis. *Strabon: A Semantic Geospatial DBMS*. In ISWC, 2012.
- [KKK+12] Manolis Koubarakis, Manos Karpathiotakis, Kostis Kyzirakos, Charalampos Nikolaou, and Michael Sioutis: "Data Models and Query Languages for Linked Geospatial Data". In 8th Reasoning Web Summer School (RW2012), Vienna, Austria, September 2012.
- [KKN+16] Manolis Koubarakis, Kostis Kyzirakos, et al. *Managing Big, Linked, and Open Earth-Observation Data Using the TELEIOS/LEO software stack*. In IEEE Geoscience and Remote Sensing Magazine, Vol. 4, Issue 3, p. 23-37, September 2016.
- [NDB+15] Charalampos Nikolaou, Kallirroi Dogani, Konstantina Bereta, George Garbis, Manos Karpathiotakis, Kostis Kyzirakos, Manolis Koubarakis. *Sextant: Visualizing time-evolving linked geospatial data*. In Journal of Web Semantics, Vol. 35, No 01.
- [SK16] Panayiotis Smeros and Manolis Koubarakis. *Discovering Spatial and Temporal Links among RDF Data*. Linked Data on the Web Workshop (LDOW2016), 12 April 2016.
- [ZKM13] Y. Zhang, M. Kersten, S. Manegold. SciQL: Array Data Processing Inside an RDBMS. In ACM SIGMOD International Conference on Management of Data (SIGMOD 2013), New York, NY, USA, 22-27 June 2013.