

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

**Μπαριτάκης Παύλος
Μεταπτυχιακός Φοιτητής**

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

Επόπτης Μεταπτυχιακής Εργασίας: Καθηγητής, Δ. Πλεξουσάκης

Παρασκευή, 11 Ιουνίου 2021 , ώρα 10:00 π.μ.

Join Zoom Meeting

<https://zoom.us/j/95999519600>

“Διαλειτουργικότητα σε αποθήκες μεγάλου όγκου μετεωρολογικών και χωρικών δεδομένων”

Περίληψη

Στις μέρες μας με την κλιματική αλλαγή να αποτελεί μείζων ζήτημα για την Επιστημονική Κοινότητα αλλά και για τους απλούς πολίτες επίσης, η Επιστημονική Κοινότητα αγωνίζεται για να βρει αιτίες, αποτελέσματα και προσπαθεί να κάνει προβλέψεις σχετικά με τις μελλοντικές συνέπειες που θα επιφέρει στις ζωές μας.

Έχοντας το Δίκτυο των Πραγμάτων (IoT), όπως τους αισθητήρες Μετεωρολογικών δεδομένων, σαν βέλος στη φαρέτρα, βοηθά την Επιστημονική Κοινότητα να συλλέγει σημαντικές πληροφορίες σχετικά με τα Μετεωρολογικά και Χωρικά δεδομένα ώστε με τη χρήση των δεδομένων αυτών να υλοποιούν στατιστικά και μαθηματικά μοντέλα, τα οποία και θα τους οδηγήσουν σε ακριβέστερα αποτελέσματα, με ταχύτερο ρυθμό από τα ήδη υπάρχοντα.

Η συχνότητα συλλογής των Μετεωρολογικών και Χωρικών δεδομένων από διαφορετικές πηγές πληροφόρησης, οδηγεί σε μεγάλα τμήματα δεδομένων, τα οποία πρέπει να αποθηκευτούν και να είναι διαχειρίσιμα αποδοτικά, με την έννοια του να μπορούν να είναι χρήσιμα προς τους χρήστες μετατρέποντας ακατέργαστα δεδομένα σε γνώση.

Η λύση για την αποδοτική αποθήκευση και διαχείριση αυτών των μεγάλου όγκου δεδομένων δόθηκε αρχικά με την κατασκευή μιας «Αποθήκης Δεδομένων» με συλλογή από διαφορετικές βάσεις δεδομένων, σχετικές με τις πηγές πληροφόρησης από όπου προέρχονται και δευτερευόντως με τη χρήση επιπέδου «Γνωσιακής Αναπαράστασης» , πάνω από το επίπεδο της «Αποθήκης Δεδομένων».

Σκοπός της εργασίας αυτής είναι η υλοποίηση μιας Διαδικτυακής Εφαρμογής Διαχειριστικού Πληροφοριακού Συστήματος, το οποίο χρησιμοποιεί Μη Σχεσιακές NoSQL βάσεις δεδομένων για την κατασκευή του Επιπέδου Αποθήκευσης, όπως επίσης και του επιπέδου «Γνωσιακής Αναπαράστασης». Η βάση δεδομένων «Apache Cassandra DB» χρησιμοποιήθηκε για το επίπεδο αποθήκευσης και το επίπεδο Γνωσιακής Αναπαράστασης υλοποιήθηκε με τη χρήση γράφων δεδομένων και πιο συγκεκριμένα με «Neo4j Graph DB». Ο συνδυασμός αυτών των δύο Μη Σχεσιακών βάσεων δεδομένων οδήγησε σε ένα δυναμικό Διαχειριστικό Πληροφοριακό Σύστημα το οποίο είναι ικανό να χειρίζεται τον όγκο δεδομένων από τους αισθητήρες. Η διαδικτυακή εφαρμογή μπορεί να χρησιμοποιηθεί από αρχάριους ως και έμπειρους χρήστες για να συλλέγουν και να διαχειρίζονται τα δεδομένα, να δημιουργούν στατιστικά και πίνακες δεδομένων, όπως και να εκτελούν δυναμικές επερωτήσεις στην «Αποθήκη Δεδομένων» κατά απαίτηση.

University of Crete

Computer Science Department

M.Sc. Thesis presentation / examination

Pavlos Baritakis

Master's Thesis Supervisor: Professor, D. Pleksousakis

Friday, 11 June 2021, 10:00 a.m.

Join Zoom Meeting

<https://zoom.us/j/95999519600>

“Interoperability over meteorological and spatial big data warehouse”

Abstract

Nowadays with climate change being a major issue in Scientific Community and to citizens too, the scientific community struggles to find causes, results and make predictions about future impacts to our lives.

Having Internet Of Things, like Meteorological Sensors, like an arrow in the quiver, helps the Scientific Community collect important meteorological and spatial data in a manner of using this kind of data to build statistical and mathematical models to guide them to more accurate results, plus faster than the existing models.

The frequent collection of the meteorological and spatial data from heterogeneous sources of information, drives to huge portions of data that have to be stored and managed efficiently in a sense of being useful to users by converting raw data format into knowledge.

The solution to the efficient storage and management of these big portions of data was given firstly by building a data warehouse with a collection of different databases, regarding the sources of information and secondly by using a Knowledge Base Layer over the data warehouse. With this approach, we create interoperability over the data warehouse.

The approach of the thesis is a Web-Based Management Information System that uses NoSQL databases to build the Storage Layer so as the Knowledge Base Representation Layer. The Apache Cassandra DB is used as the Storage Layer and the Knowledge Base Layer implemented with the usage of Neo4j Graph DB. The combination of these two NoSQL Databases leads to a dynamic M.I.S. Web-Based Application that can handle the load of data from sensors. The Web App can be used easily from novice to more advanced users to gather and manage the data, create statistics, views and execute dynamic queries to the database warehouse to have results on demand.