Research Incentive Fund (RIF)
Grant Proposal APPLICATION FORM

Please read the RIF Faculty Handbook before c
All fields will expand as needed. Submit completed form to research@zu.ac.ae
Please note: When saving this application form, please use the naming convention, 'your surname'_RIF_2011_proposal.doc, e.g.: Martin_RIF_2011_proposal.doc

OFFICE USE ONLY

SECTION A: PROPOSAL SUMMARY

Project Title:	Secure and Privacy-Preserving Querying of Personal Health Records on the Cloud			
Principal Investigator (PI):	Name:	Dima Alhadidi	College/Dept:	College of Technological Innovation
	Title:	Assistant Professor	Highest Degree:	PhD
	Tel:	0569540445	Year Awarded:	2010
	Email:	Dima.Alhadidi@zu.ac.ae		

List all Co-Investigators below, including those from other institutions:

Name	Email	Highest Degree	University/ College
Benjamin Fung	Ben.fung@mcgill.ca	PhD	McGill University
Farkhund Iqbal	Farkhund.Iqbal@zu.ac.ae	PhD	Zayed University

1. Abstract - Please provide a concise summary of the proposed research in plain language (max. 150 words).

To better understand what caused a disease, health organizations need as much data as possible about the infected patents. Personal Health Records (PHR) are user-friendly, online solutions that give patients a way of managing their own health information. Recently, architectures of storing PHRs in cloud have been proposed. However, privacy remains an issue for many patients. For example, medical centers in UAE are tackling the decision of what to move to the cloud despite being constrained by security fears. In this research, we will propose a promising approach that outsources PHRs in a form that does not reveal any private information about any patient to the cloud server. Furthermore, it allows health organizations producing statistical information about PHRs stored on a cloud while preserving the patient privacy. Moreover, the approach will not enable patients to infer about what health organizations are concerned about; not to create panic between patients.

2. Time Period – are you requesting a one-year or two-year grant?

☒ one-year grant ☐ two-year grant

3. Students – Does the project contribute to student research training?

☒ Yes, describe briefly Two students from MSIT CTI will be hired for this project. They will benefit from the research experience of the investigators of this proposal. Dr. Fung has more than 60 refereed publications that span across the prestigious research forums of data mining, privacy protection, cyber forensics, web services, and building engineering.

☐ No

4. Budget - What is the total budget requested for this proposal over the entire research period?

AED 99,160

5. Facilities and Equipment - Are you requesting space for a research assistant(s) and/or special equipment? If so, please describe briefly:

A research assistant will be hired from McGill University. A high-performance laptop is needed to conduct the experiments.

6. Ethical Clearance - Does this proposed research involve human or animal subjects?

☐ **Yes –** you will be required to apply for ethical clearance through the Research Ethics Committee if your proposal is successful

☒ **No**

SECTION B: DESCRIPTION OF RESEARCH PROJECT

Please provide a detailed project description using the six (6) section headings below.
Please write in plain language, limiting the use of jargon and acronyms.

1. Statement of the research problem(s)
(maximum 400 words)

There have been many fatal and highly contagious diseases throughout history. The Black Death was one of the most devastating pandemics in human history, peaking in Europe in the 14th century and killing between 75 and 200 million people. There are many organizations that work on studying epidemiology to prevent them from spreading around the world; the World Health Organization (WHO) is one of them. To better understand what caused a disease, health organizations need as much data as possible about the infected patients.

In the real world, electronic health records are usually managed by many different healthcare providers including primary care physicians, hospitals and pharmacies. Consequently, it is difficult to get a single patient's history due to the fact that it is spread between multiple providers. Combining these records together is even more difficult because of the cost of creating and enforcing the use of a standard that can be trusted by all parties. It has become a recent trend for patients to take these

matters into their own hands by managing their own records using a Personal Health Record (PHR) system. In the past few years, many providers have created platforms to manage PHRs with features including flexible access control and mobile access. These providers include Microsoft HealthVault and Dossia.

Recently, architectures for storing PHRs in the cloud have been proposed. However, privacy remains an issue for many patients. Since these records are stored on cloud servers, it means that these servers have the ability to read any medical record in the system. In addition, if an attacker were able to compromise a cloud server, then all the PHRs would be exposed. For these reasons, researchers have begun searching for an approach to allow patients storing their medical records on the cloud while preserving their privacy. Since health organizations need to generate statistics about health records outsourced to the cloud, the approach should support this process while also protecting the patient privacy.

2. Literature review – a concise and current review of scholarly research or important information relating to your research topic
(maximum 600 words)

Whereas an electronic health record (EHR) is a computer record that originates with and is controlled by doctors, a personal health record (PHR) can be generated by physicians, patients, hospitals, pharmacies, and other sources but is controlled by the patient. In the field of Electronic Health Records (EHR), which are administered by health care professionals, there have been many research proposals [Choe and Yoo 2008,Hembroff and Muftic 2010,Hupperich et al. 2012,Jafari et al. 2010,Neubauer and Heurix 2011,Ueckert 2003]. Hembroff and Muftic [Hembroff and Muftic 2010] and Neubauer and Heurix [Neubauer and Heurix 2011] have proposed the use of smart cards to better protect the patient's data.

It has become a recent trend for patients to manage their own records using a Personal Health Record (PHR) system. In this domain, Microsoft HealthVault and Dossia encrypt personal health records using symmetric keys, while all communication is encrypted using Secure Sockets Layer (SSL). Recently, the executive office of President President's Council of Advisors on Science and Technology (PCAST) has reported that more extensive cloud-based solutions in the future should support public health reporting and basic clinical research [PCAST 2010]. In the domain of storing health records in the cloud, Narayan et al. [Narayan et al. 2010] have shown how Attribute-Based Encryption (ABE) scheme can be used to construct a secure and a privacy-preserving Electronic Health Record (EHR) system that enables patients sharing data among healthcare providers in a flexible manner. The EHR are stored on untrusted cloud storage and encrypted using symmetric key cryptography. The

proposed keyword search functionality allows the patient choosing what terms may be searched, and who may be able to access the search terms. For example, a patient may allow a certain hospital searching the terms "Diabetes" and "male". Narayan et al. have assumed also that there is a trusted authority who generates keys for users of the system. Similarly, Li et al. [Li et al. 2010] and Akinyele et al. [Akinyele et al. 2011] have leveraged ABE techniques to encrypt health records under the assumption of the existence of a trusted authority. Li et al. [Li et al. 2010] have suggested Attribute-Based Encryption (ABE) as a solution to secure the stored medical records. ABE is utilized to encrypt and store the PHR data on semi-trusted servers, so that patients, as well as various users from public domains with different professional roles, can have access to PHRs. To produce statistical information about health records, patients can give access to health organizations using ABE. According to a report from the consulting firm PwC [PwC 2012], health organizations are falling short in protecting the privacy and security of patient information [Miami 2012]. Additionally, according to the same report [PwC 2012], more than half of health organizations have at least one issue with information security and privacy since 2009 and the most frequently observed issue is the improper use of protected health information by an employee in the organization.

Comparing the proposed approach with the aforementioned research proposals, the proposed approach allows executing various types of SQL queries while preserving the privacy of the patient and the health organization as well. The health records will be stored using schemes, which are semantically secure instead of using the symmetric key cryptography. The symmetric key cryptography has historically been susceptible to known plaintext attacks, chosen plaintext attacks, differential cryptanalysis and linear cryptanalysis. Additionally, the proposed approach will not have a trusted authority that represents a single point of failure.

[References cited in this document are listed in a separate file, Dima_RIF_2014_ReferenceList]

3. Goals of the research – anticipated outcomes including potential problem solutions, or contribution to knowledge or understanding of issues
(maximum 400 words)

Long-Term Objectives: The long-term goal of the proposed research is to take a leading role in producing new knowledge and innovative technologies in the domain of outsourcing health data to the cloud. This helps patients to benefit from the cloud advantages while protecting their privacy and generating the required medical statistics.

Cloud computing is the future. There is a need to define a sustainable environment that allows

storing health records in cloud servers and at the same time protects the patient privacy. The proposed approach in this proposal is a general one that affects all the people everywhere. Patients need to manage their medical history. They need to store their medical records in the cloud. Patients and cloud service providers will get benefit from the results of this research regardless of their locations.

Short-Term Objectives: The short-term goal of this proposed research project is to address the following vital issues:

• Outsourcing health records. Proposing a format that allows patients storing their health records on the cloud server and at the same time protecting their privacy. Cloud servers should not have the ability to read the health records. Moreover, PHRs should not be exposed in the case of external attacks.

• Producing statistics. Detailing a protocol that allows health organizations producing statistical information about PHRs stored in the cloud. The main characteristics of the proposed protocol are the following:

o It supports the outsourcing format of the health records in the cloud server.

o It protects the privacy of both health organizations and patients.

o It supports aggregate queries such count, sum, max and min.

• Executing efficiently. Health organization can execute queries and generate statistics efficiently. At the same time, patients can access and modify their health records easily and without any unreasonable overhead.

• Implementing a prototype. A prototype can evaluate the performance of the proposed protocol and accordingly report on the results of the implementation. The prototype will be simulated using a real cloud server.

4. Research methodology – anticipated methods to be used in your research process, including main research questions, data gathering, documentation or analysis planned
(maximum 400 words)

To reach the stated objectives, we will follow a methodology that is based on three lines of

research:

• Secure Outsourcing. Our goal is to protect the privacy of the outsourced health records such that the cloud server cannot read them. To this end, semantically-secure encryption schemes are adopted to encrypt PHRs before outsourcing. Using these schemes, it must be infeasible for a computationally bounded adversary to derive significant information about a message when given only its ciphertext and the corresponding public key.

• Secure Querying. The main intent of this research thread is to protect the privacy of the patients when health organizations execute queries over the outsourced health data to generate the required medical statistics. To this end, we will adopt the research results that has been recently published by Barouti et al. [Barouti et al 2013]. In this research, Barouti et al. has proposed protocols for keyword search and aggregate SQL queries that preserve the privacy of both the client and the database owner. Our proposed protocol in the health domain will depend on homomorphic properties of some semantically secure encryption schemes. Homomorphic encryption is a form of encryption where a specific algebraic operation performed on the plaintext is equivalent to another (possibly different) algebraic operation performed on the ciphertext. For example, the Paillier's scheme is an additive homomorphic public key encryption. Using Paillier's scheme, given two ciphertexts E(x) and E(y) of two plaintexts x and y respectively, an encryption of their sum E(x+y) can be efficiently computed by multiplying the ciphertexts modulo a public key N, i.e., E(x+y)=E(x).E(y) mod N. Threshold schemes can be considered because of the multi-owner setting. Multiple owners (patients) should encrypt their records such that they cannot decrypt each other records and at the same time they should reduce the key distribution complexity.

• Efficient execution. Tree structures can be considered because they considerably reduce communication and computation overhead.

• Analysis and Validation. This research thread will discuss the correctness, the privacy and the efficiency of the proposed protocol. We will provide the required proofs in this domain depending on the definitions of the Secure Multiparty Computation (SMC) together with the complexity analysis. Moreover, we will estimate the computation and the communication costs of the proposed protocol. We will try to get real health data from real hospitals and perform the practical experiments in a real cloud. If we cannot access real health data, we can employ some of the publicly available datasets, e.g., Breast Cancer [BC 2012]. Publicly available datasets are used by researchers to validate their research proposals and to compare the results.

5. Research schedule and deliverables – what are the major phases of your research anticipated, and what do you

realistically plan to accomplish at what stage

Major phases and milestones of the proposed project are described below along with the required timeframe to accomplish each task.

T1. Data collection and preparation (January 2014): Collecting medical data is a very challenging task for researchers. Under the Health Insurance Portability and Accountability Act (HIPAA) in United States, health care enterprises must guard protected health information and implement policies and procedures to safeguard it. If we cannot collect real medical data, we can resort to publicly available medical datasets or generate random ones.

T2. Literature review (February - March, 2014): Research proposals that are related to the proposed protocol will be reviewed. Actually, we will focus on the research proposals that target the following two subjects: (1) secure storage and access of health records and (2) private data outsourcing.

T3. Choosing a suitable format to store the medical data in the cloud (April, 2013): Many encryption schemes and anonymaization techniques will be studied carefully to decide about the storage format. The format should protect the patient privacy from the cloud server and attackers and allow producing the required statistics in a secure way.

T4. Proposing a protocol to execute queries and generate statistics (May-July, 2014). The protocol should be designed such that no more information than the query result should be revealed to the health organization. The proposed protocol also does not enable patients to infer about what health organizations are concerned about; not to create panic about epidemics in the community.

T5. Analyzing the proposed protocol (August 2014)

T6. Implementing, testing and conducting experimental results (September- October 2014)

T7. Evaluation and dissemination of results (Novemebr - December, 2014): In addition to share the findings and experimental results with the students and the faculty of Zayed University and McGill University, we expect to publish the new results in a journal and a conference paper.

6. Budget narrative – describe and justify your main budget items. An itemized budget spreadsheet will also be attached to this proposal

The estimated total budget for the proposed project is (AED 99,160), which is distributed as follows:

- Hiring two students from MSIT CTI for 224 hours at the rate of 95 AED/hour (AED 42,560).

- Hiring one student as a research assistant from McGill University for 480 hours at the rate of 95

AED/hour (AED 45,600)

- Buying a laptop (AED 11000).

SECTION C: RIF BUDGET SPREADSHEET

Please attach your completed RIF Budget Spreadsheet.
Please use the naming convention 'your surname'_RIF_budget.xlsx eg Martin_RIF_budget.xlsx

SECTION D: SEDONA CV

Please attach your updated SEDONA CV
Please use the naming convention 'your surname'_SEDONA_CV.doc eg Martin_SEDONA_CV.doc

SECTION E: DEAN'S EVALUATION

You must obtain the physical signature of your Dean before submitting this application form. Applications without signatures will not be accepted.
If you do not have access to digital signatures, it is recommended to:
- print the completed form
- obtain the Dean's signature
- sign the application yourself
- scan and email to research@zu.ac.ae Please note: When saving this application form, please use the naming convention, 'your surname'_RIF_2011_proposal.doc eg Martin_RIF_2011_proposal.doc

Dean's Name:		College/Department	
I endorse that this project is appropriate for the unit to undertake as part of its educational, service or research programs; that appropriate and sufficient staff are available and willing to supervise; and that adequate space and facilities are available.			☐Yes ☐No
I approve the request for facilities and equipment. (Arrangements will be made directly with the Principal Investigator)			☐Yes ☐No ☐ NA
Comments:			
Dean's Signature:		Date:	

I certify that all information provided is true and correct at the time of submission.

Submit to research@zu.ac.ae

PI's Signature:		Date:	