

MOHAMED ADEL REZK M.Sc. B.Comm.

Senior Research Assistant, Insight Centre for Data Analytics

✉ moh.adelrezk@gmail.com ☎ 089-9676871 ✉ 6 Uillin, Moycullen 📍 Galway, IE

🌐 mohadelrezk 🔗 dblp/mohamedadelrezk 🌐 mohadelrezk

🐦 @MohamedAdelRezK 🔗 gscholar/mohamedadelrezk 🔗 insight-centre/mohamedadelrezk
🌐 eGov-Centre/mohamedadelrezk 🔗 NewsLab/mohamedadelrezk



EXPERIENCE

Senior Research Assistnat

DKE Unit - Insight Centre for Data Analytics - sfi funded

reporting to and mentored by: prof. Mathieu D'Aquin

📅 Jan 2017 – Ongoing 📍 Galway, IE

Project: Automating Data Science

- Conceptualized the problem of data science automation in context of enabling citizen data scientists, and designed a solution/model to answer it.
- Defined possible knowledge sources for this model to learn from, Scrapped structured knowledge sources like rminer, weka, orange, knime's workflows on github was not enough and datasets were not available.
- Scrapped unstructured but richer knowledge source (kaggle's kernels and datasets).
- Feature engineering of the unstructured kaggle kernels and datasets to determine its data science workflow.
- Applied frequent itemset and association mining using apriori algorithm, to find patterns in the data science workflow.
- Applied frequent sequence rules to test the sequqncinality of the workflows.
- Developed (.cxt) context file generator in python for FCA as the available FCA bedrock couldn't handle the diversity of workflows.
- Developed a python wrapper for the FCA tool inclose4 (c++).
- Applied formal concept analysis/lattice to extract concepts from the workflows to use it for taxonomy building and labeling processes.
- Developed a dynamic survey webapp to scientifically taxonomize data science tasks using discovered concepts and to avoid most of the unjustified bias.
- Next, we will build a classifier, test and optimize it and finally plug it to any data science or open data platform. Then designing KPIs for measuring the effect on citizen data science.
- Recently DARPA is funding 30 groups in 30 universities in the US to answer the same question.
- Code repos. for this project are internally accessible on gitlab@insight-centre.

Langs.	python, javascript, noSQL, bash script
Frameworks	django, mongodb, jupyter server, git
Libs.	selenium, bloodhound.js, jquery, bootstrap3 kaggle api, pymongo, geckodriver, pyvirtualdisplay mlxtend, spmfJARwrapper, FCA bedrock, inclose4, github api
Envs.	unix, windows
Tools	pycharm, studio3t, jupyter notebook, Atom jira, rapid miner, knime, weka, orange, nginx
Data types	csv, xml, json, cxt
Algorithms	Frequent Items sets, Frequent Sequence Rule Association Rule, Formal Concept Lattice

eGov Unit - Insight Centre for Data Analytics

reporting to and mentored by: Dr. Adegboyega Ojo

📅 Oct 2015 – Ongoing 📍 Galway, IE

Project: European Linked Open Statistical Data Hub

- Provided consultation for derilinx and adapt centre on how to utilize their linked data creator app (juma) to create linked statistics data according to rdf data cube vocabulary specification and constraints.

MOST PROUD OF

Awards

🏆 **CSS Egypt Award**
Enhancing Internet User's Experience using Semantic Web Technologies
Enhanca App

📈 **Distinction-class honours**
Top of 2010 Class

💖 **Enthusiastic & Innovative**
I like challenging tasks, never give up and always find a path.

EDUCATION

M.Sc. in Management Information Systems

Alexandria University & NUIG

📅 Sept 2015 – June 2017

Advanced Statistics Data Structures
Research Methods
Advanced Topics in Info. Sys

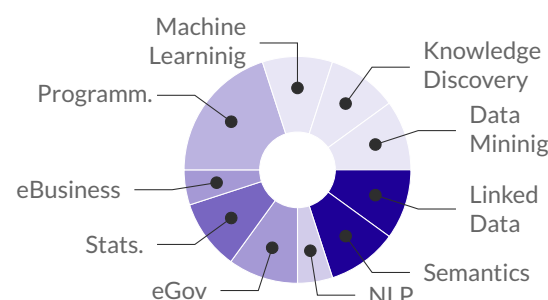
B.Comm. in Management Information Systems

Alexandria University

📅 Sept 2006 – June 2010

Statistics DB Design Networks
OOP in JAVA & C# e-Commerce

KNOWLEDGE



- Designed and co-implemented configuration files to better handle input csvs and create rdfs conforming to the integrity constraints of the vocabulary.
- Implemented and configured json-qb-api over data cubes created by adapt centre's juma app.
- Developed adaptations to rdf data cube explorations and visualization tools developed for open government intelligence project.
- Presented the pipeline and co-managed the testing hackathon attended by european statistics offices representatives at Insee, Paris, France.
- Developing python tool for bulk creation of rdf data cubes to replace juma tool.
- Code repos for this project are internally accessible on gitlab@insight-centre.

Langs.	python, java, sparql, javascript
Frameworks	django, pip, git
Libs.	rdflib, pivot.js
Envs.	unix
Tools	virtuoso, pycharm, trello, nginx, jupyter notebook
Data types	rdf, xml, json, json-ld, csv

Project: Open Government Intelligence

- Extended tarql library (java) and created [cube builder] by adding more interfacing features i.e webapp, web service and command line and enabled configurable rdf data cube vocabulary compliance and creation.
- Designed a pipeline [assisted schema creator] for data cube vocabulary mapping and linked data enrichment using openrefine, rdf-refine extension and excel sheets.
- Requirement analysis, deployment and bug reporting of json-qb-api, cubiq and table2qb tools.
- Developed a python wrapper for the clojure tool table2qb, with extra functionalities like data preprocessing, codelist creation, data chunking and decreasing data processing's big O.
- Developed rdf data cube python communicator to handle communications between js based visualization apps and cubiq api.
- Developing a python real time rdf data creation, publishing and visualization pipeline for the marine institute pilot [rescue, renewable energy and tourism] use cases.
- Developing (linked data - druid - superset) bridge tool in python.
- Pilots data preprocessing and restructuring to fit the pipeline and comply with data requirements.
- Co-authored the paper "Using Linked Statistical Data to Improve Marine Search and Rescue Operations in Ireland" Stasiewicz et al. 2018, to communicate the real time rdf data publishing tool details and use cases.
- Co-authored the project deliverables and presented our work to partners and pilots during plenary meetings in Dublin, Lithuania and Thessaloniki.
- Code repos [Cube builder, Schema creator, Real time rdf publishing pipeline, Table2qb python wrapper, Lithuanian pilot data restructuring]
- Other Code repos are internally accessible on gitlab@insight-centre.

Langs.	python, java, sparql, javascript, clojure, graphql
Frameworks	django, maven, pip, git, lein
Libs.	rdflib, tarql, pivot.js
Envs.	unix
Tools	virtuoso, eclipse, pycharm, nginx, jupyter notebook
	trello, jira
Data types	rdf, xml, json, json-ld, csv, xlx

Project: ROUTE-TO-PA

- Developed an anomaly detection service using LOF to point anomalies in published datasets on ckan instances.
- Developed a dataset Quality assessment service to assist in data preparation for the LOF service
- Code repos.: [Anomaly Detection Service]

SKILLS

Strengths

Data Science Pipeline

Python

Java

Semantics and Linked Data

Languages

Arabic

English

REFEREES

Dr. Adegboyega Ojo

eGov Unit Leader

@ adegboyega.ojo@insight-centre.org

✉ eGov Unit, Insight Centre, NUIG Galway, IE

Prof. Mathieu D'Aquin

Insight Centre Director, and DKE Unit Leader

@ mathieu.daquin@insight-centre.org

✉ DKE Unit, Insight Centre Director, NUIG Galway, IE

Dr. Wassim Derguech

Technical Project Manager, Derilinx

@ wassim@derilinx.com

✉ Derilinx, 11 Baggot Ct Dublin, IE

Langs.	python
Frameworks	flask, ckan, pip, git
Libs.	pylof
Envs.	unix
Tools	pycharm, trello, jupyter notebook
Data types	json
Algorithms	Local Outlier Factor

M.Sc.: A Government Support System based on Citizen Interactions

Proposing a framework for linked public policy management.

Contribution-a: Core Public Policy Vocabulary

- RDF vocabulary to conceptualize public policies
- Documentation: [CPPV, CPPV ckan demo]
- Publications: ["Standardising Public Policy Documentation to Foster Collaboration Across Governmental Agencies" Rezk et al. 2017a, "Towards a Vocabulary for Open Public Policies" Rezk et al. 2017b]

Contribution-b: Public Policy Collaboration Mining

- Mining collaboration opportunities between public policies stakeholders by analysing meta data of published data sets on open data platforms.
- Documentation :[collaboration miner]
- Publications/posters: ["Mining Governmental Collaboration Through Semantic Profiling of Open Data Catalogues and Publishers" Rezk, Adegboyega Ojo, and Hassan 2017, poster panel at the AI and smart enterprise event at DCU]

Contribution-c: Public Policy Acceptance Prediction

- Analysis of the past public policies' related social media interactions, to forecast citizens reactions towards proposed public policies before it is published to the public.
- Documentation : [GOV-DAF]
- Publications: ["A Government Decision Analytics Framework Based on Citizen Opinion" Rezk, Khayat, et al. 2016, "A Predictive Government Decision Based on Citizen Opinions:ToolsResults" Rezk, Adegboyga Ojo, et al. 2018]
- Code repos.: [CPPV, CPPV ckan demo, collaboration miner, GOV-DAF]

Langs.	python, java, uml, sparql, noSQL, javascript
Frameworks	flask, ckan, ckan scheming, pip, git, mongodb
Libs.	twitter api, stanfordNER, dbpedia spotlight, LDA, sigma.js
Envs.	unix
Tools	pycharm, eclipse, jupyter notebook, robomongo, protege
Data types	json, rdf, owl, pdf
Algorithms	Naive Bayes, LDA topic modeling

Project: Your Data Stories

- Galway Traffic data batch processing and restructuring for RDFizing.
- Code repos.: [traffic data transformation]

Lang.	python
Env.	unix
Tools	pycharm, jupyter notebook
Data types	csv

Research Assistant

NewsLab - Insight Centre for Data Analytics

📅 Jan 2015 – Oct 2015

📍 Galway, IE

Project: Social Repository of Ireland

- Built semantic social repository for Ireland to preserve irish social media as part of the irish heritage stored in dri.ie.
- Collected, enriched and linked tweets related to irish events and visualised its stats and aggregations.
- co-authored project deliverables.
- Code repos for this project are internally accessible on gitlab@insight-centre.

Langs.	java, nodejs, javascript, noSQL
Libs.	semantic similarity DISCO, stanfordNER, twitter api
Envs.	unix, maven, npm, git, mongodb
Tools	eclipse, robomongo
Data types	json

Project: Newswire 2.0

- Collecting news articles from rss feeds to build a popularity prediction model based on their NER popularity indexes calculated from social media sentiments.
- Main project deliverables author.
- Code repos for this project are internally accessible on gitlab@insight-centre.

Langs.	java, noSQL
Libs.	twitter api, stanford NER
Envs.	nix, maven, git, mongodb
Tools	eclipse, robomongo
Data types	json, xml

Teaching Assistant

Computers and Information Systems Department, Alexandria University

📅 May 2011 – Jan 2015

📍 Alexandria, Egypt

- Teaching Labs for undergraduate students of Information Systems major [Systems Analysis and Design, OOP, eCommerce, DB Design and Artificial Intelligence].
- Supervising undergraduate students' graduation projects.

Oracle Applications Support Engineer

ICTP project, Faculty of Engineering, Alexandria University

📅 Nov 2010 – May 2011

📍 Alexandria, Egypt

- Maintaining, Extending and Adminstrating Alexandria University MIS applications and servers.

EVENTS & CERTIFICATIONS

- **Open Gov. Intelligence Final Plenary Meeting**, at CErTH, Thessaloniki,GR
[demo - presentation - 29-30 Nov.18]
 - **Blockchain Workshop**, at Insight NUIG, Gawlay,IE
[workshop - training - 21 Nov.18]
 - **Tableau Workshop**, at Insight NUIG, Gawlay, IE
[workshop - training - 9 Nov.18]
 - **Green Screen Workshop**, at Insight Centre, Galway, IE
[training - workshop - project - 1-2 Nov.18]
 - **Introduction to LaTeX**, at NUIG, Galway, IE
[workshop - training - 15 Oct.18]
 - **LOSD Publishing Pipeline Testing Hackathone**, at Insee, Paris, FR
[demo - workshop - testing - 10-11 Sept.18]
 - **AI & the Smart Enterprise event**, at DCU, Dublin, IE
[poster - panel - 7 June.18]
 - **ICE-GOV 18 conference**, at NUIG, Galway,IE
[poster - paper - panel- 4-6 Apr.18]
 - **LOSD Publishing Pipeline Kick-off Hackathone**, at CSO, Dublin,IE
[presentation - brainstorming - 28-29 Feb.18]
-
- **Open Gov. Intelligence 7th Plenary Meeting**, at Enterprise Lithuania, Vilnius,LT
[demo - presentation - 27-28 Nov.17]
 - **PROVE 17 confernece**, at University of Vicenza, Vicenza, IT
[presntations - confernece - paper - 19-20 Sep.17]
 - **Open Gov. Intelligence 6th Plenary Meeting**, at Marirne Institute, Dublin, IE
[demo - presentation - 28-29 Aug.17]
 - **Deep learning Workshop**, at Insight Centre, DCU, Dublin, IE
[workshop - traininig - 27-28 Apr.17]
-
- **Reading Group**, at Insight- NUIG, Gawlay, IE
[training - bi weekly - 1 Nov.15]
-
- **Advanced Java Programming by Oracle**, at ITI, Alexandria, EG
[certificate - training - Feb.13]

PUBLICATIONS

Thesis

- Rezk, Mohamed Adel (2017). *A Government Support System Based on Citizens' Interactions: A Comparative Study of Different Arabic Language Analysis Methodologies*. URL: github.com/mohadelrezk/resume/raw/master/publications/MSc_Dissertation_MohamedAdel_Final_August2017.pdf.

Conference Proceedings

- Rezk, Mohamed Adel, Adegboyga Ojo, et al. (2018). "A Predictive Government Decision Based on Citizen Opinions: Tools & Results". In: *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance*. ACM, pp. 712–714.
- Stasiewicz, Arkadiusz et al. (2018). "Using Linked Statistical Data to Improve Marine Search and Rescue Operations in Ireland". In: *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance*. ACM, pp. 412–418.
- Rezk, Mohamed Adel, Adegboyega Ojo, and Islam A Hassan (2017). "Mining Governmental Collaboration Through Semantic Profiling of Open Data Catalogues and Publishers". In: *Working Conference on Virtual Enterprises*. Springer, pp. 253–264.
- Rezk, Mohamed Adel et al. (2017a). "Standardising Public Policy Documentation to Foster Collaboration Across Government Agencies". In: *Working Conference on Virtual Enterprises*. Springer, pp. 466–477.
- – (2017b). "Towards a Vocabulary for Open Public Policies". In: *Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance*. ACM, pp. 584–585.
- Rezk, Mohamed Adel, Ghada A El Khayat, et al. (2016). "A Government Decision Analytics Framework Based on Citizen Opinion". In: *Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance*. ACM, pp. 27–30.