# kyc-logo.jpg

# Classification of politically exposed persons (PEPs)

written by Jon Camilleri B.Sc (Hons) IS&M

https://mt.linkedin.com/in/jonathancamilleri

*FFIEC* refers to politically exposed persons as including a current or former senior political figure, their immediate family and their close associates.

Corporate governance information systems analysts are particularly interested in politically exposed persons due to the conflict of interest that may arise as a political officer may be different roles, and, this matter is governed and managed through the organizational structure, commonly through the use of a committee that oversees the conflict of interest or the wrongdoing that was otherwise reported. Normally evidence is expected to be provided to the review board and they process this matter through a form of mediation process with an attempt to understand all aspects of the incident. When the committee has reviewed all the necessary documentation, a decision is taken on the matter that gave rise to the reported conflict of interest or wrongdoing.

# Project description

The proposed software implementation serves to propose the automatic classification of politically exposed persons updated within a database where the process shall attempt to classify politically exposed persons into the following classifications:

* closely affiliated i.e. high risk with a score of 1.
* loosely affiliated i.e. moderate risk with a score that is a fraction of 1.
* unpredictable

The training set shall consist of the largest possible representative sample of data that can be used as a training set, and, independent testing shall be carried out to verify the working of the algorithm, as there is a high degree of subjectivity involved.

The research is based on qualitative and quantitative research using online resources available as this is considered to be a low-cost manner of organizing a desktop research. Interviews with politicians within the European Parliament were also requested by the author, however the response was weak, unfortunately. Attempts to find ready data sets have proven to be expensive and therefore the author shall have to resort to sample data created randomly, which reduces the contextual validity of the testing. Testing shall be carried out on a validation and verification basis independently using a separate method for an independent review of the classification algorithm, based on the experience of the author, and, under the supervision of [Dr. George Azzopardi PhD](https://www.um.edu.mt/profile/georgeazzopardi), Professor at the University of Malta.



Figure 1: Logical flow of classification process

# Document control

|  |  |  |
| --- | --- | --- |
| **Document revision** | **Date last revised** | **Author/reviewer** |
| First draft of document. | 10th October 2015 | JC |

# System features

The assignment in question is a research project and is not a complete open or closed system, therefore coding and documentation shall outline the project within the scope defined within the *project description* section of this document.

# Software engineering process

The source code will be written using a fourth generation programming language such as Java Progamming Language, or Python, with the intention to write a classification algorithm that will allow a training data set to serve as a supervised test for the algorithm as part of a research project. The algorithm shall be reviewed to improve the learning ability of the application itself following feedback through the quality assurance process.

Information that is read from reliable sources of information is extracted and simplified to the following format, and, the classifier shall assign a risk based on the rating scored against a model that assesses the risks of politically exposed persons:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Data type | Rating | Weight |
| Name and Surname of politically exposed person | TEXT |  | 0 |
| Political Ideology and Curriculum Vitae | TEXT |  | 0.9 |
| Jurisdiction  | TEXT |  | 1.1 |
| Genetic influence | TEXT |  | 0.9 |
| Affiliations and associations | TEXT |  | 0.9 |
| Residence address | TEXT |  | 0.9 |
| Office address | TEXT |  | 0.9 |
| Risks relevant to the jurisdiction where the politically exposed person resides or holds office | TEXT |  | 1.1 |

In view of the provisions of the [Data Protection Directives](http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML), where information is not publicly available one cannot use information gathered in a private discussion for the scope of academic research unless permission is explicitly granted. Weight is a fraction where the datum is not relevant to the risk. The above serves to provide statistical correlations of how different variables may influence on the risk of conduct.

## Quality assurance

Tests shall be carried out based on ISQETB standards and the following test methodologies have been considered as appropriate for this research assignment. Test objectives are being documented herewith, based on test-driven development, and, using semi-automated methods of verifying calculations, using spreadsheets although there is sometimes a limitation on the use of the software itself.

- Test objectives

## Implementation

??

## Further reading

1. Federal Financial Institutions Examination Council (FFIEC), Politically Exposed Persons Overview, available [online] at https://www.ffiec.gov/bsa\_aml\_infobase/pages\_manual/OLM\_087.htm.

2. Malta Financial Services Authority, website available [online] at http://mfsa.com.mt.

3. Financial Services Authority website, available [online] at http://www.fsa.gov.uk.

4. Accuity PEP Database available [online] at http://www.accuity.com/compliance/pep-due-diligence-database/ , and, article available online titled *Thomson Reuters to deploy IBM Watson technology* available [online] at http://www.automatedtrader.net/news/at/154708/thomson-reuters-to-deploy-ibm-watson-technology, last updated on 9th October 2015.

5. Media article available [online] at http://blog.transparency.org/2015/09/18/football-needs-tough-anti-money-laundering-rules/, written by Arjun Medhi and published on 18th September 2015, and, an article available [online] at https://www.transparency.org/news/feature/sport\_integrity, posted on 21st September 2015.

6. Online software source Github available online at www.github.com.

7. Java ML available online at http://java-ml.sourceforge.net/.

8. io9, *The 12 Cognitive Biases that prevent you from being rational*, available [online] at <http://io9.com/5974468/the-most-common-cognitive-biases-that-prevent-you-from-being-rational>, published on 1st September 2013.



# Appendix A - Code listing

The following is the source code listing provided:

import java.sql.Date;

import java.util.\*;

public class PEP {

 private void Learn()

 {

 }

}

Written using [Eclipse IDE](https://www.eclipse.org/ide/) Mars Milestone 3.