# 1 INTRODUCTION

TBD

## 1.1 MOTIVATION

In the 19th century, in Ireland, there was no standard of the spelling of names, handwriting could be difficult to read, contractions or abbreviations were used. Many people was not lliterate, so they asked or dictated literate people to write their names. This way even names with the same pronounciation could be written in many different ways, depending on who wrote them.

Apart from various ways of spelling one's name, people from the time use Irish names which equivalent to modern names, for example, Irish version of *Smith* could be *Gowan*. There are also some Irish prefixes like *O'*, *M'*, *Mac*, etc. When combined together this would result in *O'Gowan* or *M'Gowen*, and so on.

An example list of possible equivalent Irish names of *Smith* could be as follow

> *Smith*, *Smyth*, *Smythe*, *Smeeth*, *Going*, *Gowing*, *Maizurn*, *McGhoon*, *MaGough*, *M'Ghoon*, *MacGivney*, *MacGivena*, *M'Givena*, *MacGhoon*, *M'Evinie*, *McGivney*, *MacEvinie*, *McGivena*, *M'Givney*, *McEvinie*, *MacAvinue*, *M'Avinue*, *McAvinue*, *McCona*, *MaGowen*, *MaGowan*, *MaGovern*, *MaGeown*, *McGowan*, *McGoween*, *McGown*, *M'Cona*, *MeCowan*, *MeGowan*, *MacGown*, *MacGoween*, *MacGowan*, *MacCona*, *M'Gowan*, *M'Gowen*, *M'Gown*, *Ogowan*, *O'Gowan*, *Gowen*, *Gowan*, *Gow*, *Goan*.

At present time, when historical researchers try to trace people back using historical records, they would encounter this problem of name variations.

Various solutions have been created to find matching different names that refer to the same person. However, for our extent knowledge, there is yet no public system which encodes those solutions together and provides service of name matching. We then decided to create one platform to archieve this.

## 1.2 RESEARCH QUESTION

From the motivation, we address our research questions as follow.

1. Can we provide a web service to match names, where matching can be a complicated process because of the way people record their names.
2. Can the web service act as a platform for general names or words matching system so that it can be extended to other languages as well.

The first question derives directly from the motivation. The second question is an enhancement for the system. It can be designed as a more general purpose matching system rather than just specified only for Irish names. Therefore it should be extensible for any further matching methods to be developed in the future.

## 1.3 OBJECTIVE AND AIMS

The objective of this project is to provide a web service that encodes several of matching methods and produces matching results between two list of names.

*Matching methods implemented at the initial state of this project are Levenshtein Distance, Soundex, Irish Soundex, and Lookup Table.*

The project aims to be a part of a bigger system, such as genealogy research. These client systems, at some point, they might need a service of a name matching on demand, so then they can use this web service, providing their lists of name, methods be be used, and threshold as inputs, and get matching results for their further usage.

We would start by focusing on Irish *surname* first. For any further kind of names we would leave it for future works.

## 1.4 REPORT STRUCTURE

Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

# 2

# RELATED WORK

Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo.

# Part I

# THE SHOWCASE

You can put some informational part preamble text here. Illo principalmente su nos. Non message *occidental* anglo-romanic da. Debitas effortio simplificate sia se, auxiliar summarios da que, se avantiate publicationes via. Pan in terra summarios, capital interlingua se que. Al via multo esser specimen, campo responder que da. Le usate medical addresses pro, europa origine sanctificate nos se.

# 3 MATH TEST CHAPTER

Aliquam lectus. Vivamus leo. Quisque ornare tellus ullamcorper nulla. Mauris porttitor pharetra tortor. Sed fringilla justo sed mauris. Mauris tellus. Sed non leo. Nullam elementum, magna in cursus sodales, augue est scelerisque sapien, venenatis congue nulla arcu et pede. Ut suscipit enim vel sapien. Donec congue. Maecenas urna mi, suscipit in, placerat ut, vestibulum ut, massa. Fusce ultrices nulla et nisl.

## 3.1 SOME FORMULAS

Due to the statistical nature of ionisation energy loss, large fluctuations can occur in the amount of energy deposited by a particle traversing an absorber element[1]. Continuous processes such as multiple scattering and energy loss play a relevant role in the longitudinal and lateral development of electromagnetic and hadronic showers, and in the case of sampling calorimeters the measured resolution can be significantly affected by such fluctuations in their active layers. The description of ionisation fluctuations is characterised by the significance parameter $\kappa$, which is proportional to the ratio of mean energy loss to the maximum allowed energy transfer in a single collision with an atomic electron:

*You might get unexpected results using math in chapter or section heads. Consider the `pdfspacing` option.*

$$\kappa = \frac{\xi}{E_{max}} \tag{1}$$

$E_{max}$ is the maximum transferable energy in a single collision with an atomic electron.

$$E_{max} = \frac{2m_e\beta^2\gamma^2}{1 + 2\gamma m_e/m_x + (m_e/m_x)^2},$$

where $\gamma = E/m_x$, $E$ is energy and $m_x$ the mass of the incident particle, $\beta^2 = 1 - 1/\gamma^2$ and $m_e$ is the electron mass. $\xi$ comes from the Rutherford scattering cross section and is defined as:

$$\xi = \frac{2\pi z^2 e^4 N_{Av} Z\rho\delta x}{m_e\beta^2 c^2 A} = 153.4\frac{z^2}{\beta^2}\frac{Z}{A}\rho\delta x \quad \text{keV},$$

where

1 Examples taken from Walter Schmidt's great gallery: http://home.vrweb.de/~was/mathfonts.html

| | |
|---|---|
| $z$ | charge of the incident particle |
| $N_{Av}$ | Avogadro's number |
| $Z$ | atomic number of the material |
| $A$ | atomic weight of the material |
| $\rho$ | density |
| $\delta x$ | thickness of the material |

$\kappa$ measures the contribution of the collisions with energy transfer close to $E_{max}$. For a given absorber, $\kappa$ tends towards large values if $\delta x$ is large and/or if $\beta$ is small. Likewise, $\kappa$ tends towards zero if $\delta x$ is small and/or if $\beta$ approaches 1.

The value of $\kappa$ distinguishes two regimes which occur in the description of ionisation fluctuations:

1. A large number of collisions involving the loss of all or most of the incident particle energy during the traversal of an absorber.

   As the total energy transfer is composed of a multitude of small energy losses, we can apply the central limit theorem and describe the fluctuations by a Gaussian distribution. This case is applicable to non-relativistic particles and is described by the inequality $\kappa > 10$ (i. e., when the mean energy loss in the absorber is greater than the maximum energy transfer in a single collision).

2. Particles traversing thin counters and incident electrons under any conditions.

   The relevant inequalities and distributions are $0.01 < \kappa < 10$, Vavilov distribution, and $\kappa < 0.01$, Landau distribution.

## 3.2 VARIOUS MATHEMATICAL EXAMPLES

If $n > 2$, the identity

$$t[u_1, \ldots, u_n] = t\big[t[u_1, \ldots, u_{n_1}], t[u_2, \ldots, u_n]\big]$$

defines $t[u_1, \ldots, u_n]$ recursively, and it can be shown that the alternative definition

$$t[u_1, \ldots, u_n] = t\big[t[u_1, u_2], \ldots, t[u_{n-1}, u_n]\big]$$

gives the same result.

Part II

# APPENDIX

# A

# APPENDIX TEST

Aliquam lectus. Vivamus leo. Quisque ornare tellus ullamcorper nulla. Mauris porttitor pharetra tortor. Sed fringilla justo sed mauris. Mauris tellus. Sed non leo. Nullam elementum, magna in cursus sodales, augue est scelerisque sapien, venenatis congue nulla arcu et pede. Ut suscipit enim vel sapien. Donec congue. Maecenas urna mi, suscipit in, placerat ut, vestibulum ut, massa. Fusce ultrices nulla et nisl.

Etiam ac leo a risus tristique nonummy. Donec dignissim tincidunt nulla. Vestibulum rhoncus molestie odio. Sed lobortis, justo et pretium lobortis, mauris turpis condimentum augue, nec ultricies nibh arcu pretium enim. Nunc purus neque, placerat id, imperdiet sed, pellentesque nec, nisl. Vestibulum imperdiet neque non sem accumsan laoreet. In hac habitasse platea dictumst. Etiam condimentum facilisis libero. Suspendisse in elit quis nisl aliquam dapibus. Pellentesque auctor sapien. Sed egestas sapien nec lectus. Pellentesque vel dui vel neque bibendum viverra. Aliquam porttitor nisl nec pede. Proin mattis libero vel turpis. Donec rutrum mauris et libero. Proin euismod porta felis. Nam lobortis, metus quis elementum commodo, nunc lectus elementum mauris, eget vulputate ligula tellus eu neque. Vivamus eu dolor.

## A.1 APPENDIX SECTION TEST

Nulla in ipsum. Praesent eros nulla, congue vitae, euismod ut, commodo a, wisi. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Aenean nonummy magna non leo. Sed felis erat, ullamcorper in, dictum non, ultricies ut, lectus. Proin vel arcu a odio lobortis euismod. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Proin ut est. Aliquam odio. Pellentesque massa turpis, cursus eu, euismod nec, tempor congue, nulla. Duis viverra gravida mauris. Cras tincidunt. Curabitur eros ligula, varius ut, pulvinar in, cursus faucibus, augue.

*More dummy text*

| LABITUR BONORUM PRI NO | QUE VISTA | HUMAN |
|---|---|---|
| fastidii ea ius | germano | demonstratea |
| suscipit instructior | titulo | personas |
| quaestio philosophia | facto | demonstrated |

Table 1: Autem usu id.

Listing 1: A floating example

```
for i:=maxint to 0 do
begin
{ do nothing }
end;
```

## A.2 ANOTHER APPENDIX SECTION TEST

Your own text.

### A.2.1 *Sub*

Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

# BIBLIOGRAPHY


[1] Jon Bentley. *Programming Pearls*. Addison–Wesley, Boston, MA, USA, 2nd edition, 1999.

[2] Robert Bringhurst. *The Elements of Typographic Style*. Version 3.2. Hartley & Marks Publishers, Point Roberts, WA, USA, 2008.

[3] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*. The MIT Press, Cambridge, MA, USA, 2nd edition, 2001.

[4] Donald E. Knuth. Computer Programming as an Art. *Communications of the ACM*, 17(12):667–673, December 1974.

[5] Donald E. Knuth. Big Omicron and Big Omega and Big Theta. *SIGACT News*, 8(2):18–24, April/June 1976.

[6] Ian Sommerville. *Software Engineering*. Addison-Wesley, Boston, MA, USA, 4th edition, 1992.

## COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

http://code.google.com/p/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/

# DECLARATION

Put your declaration here.

*Co. Kildare, Ireland, April 2015*

Phattara Wangrungarun,
May 2, 2015