

---

## SOFTWARE METAPAPER

# OACoder: Postcode Coding Tool

Muhammad Adnan,<sup>1</sup> Alex Singleton<sup>2</sup> and Paul A. Longley<sup>1</sup>

<sup>1</sup> Department of Geography, University College London, London, United Kingdom

<sup>2</sup> Department of Civic Design, University of Liverpool, Liverpool, United Kingdom

---

Geodemographic classifications are small area classifications of social, economic and demographic characteristics. The Output Area Classification (OAC) is a free geodemographic classification. It is an Office of National Statistics validated measure that summarises neighbourhood conditions at the Output Area Level across the United Kingdom. Linkage of these valuable statistics has been problematic for users more used to address records that are georeferenced using unit postcodes. OACoder resolves this problem by allowing users to link corresponding OAC codes to each of the postcode addresses. OACoder is an open source software, and it is developed and tested to work on different versions of windows operating systems. It is stored in Figshare. The source code of the OACoder is stored in SourceForge. As open source software, OACoder has reuse potential across a range of applications. The functionality of OACoder can be extended to work with the new version of OAC (2011 OAC). It is also possible to reuse the source code and extend the functionality to work on different operating systems other than Windows. Different components of the software can be reused for the purpose of reading/writing CSV files and handling large data sets.

---

**Keywords:** Output Area Classification, OACoder, postcode coding

### Funding statement

This work was completed as part of the EPSRC research Grant “The Uncertainty of Identity: Linking Spatiotemporal Information in the Real and Virtual Worlds” (EP/J005266/1).

---

## (1) Overview

### Introduction

Geodemographic classifications are small area summary indicators of the social, economic and demographic characteristics of neighbourhoods. They are based upon the fundamental premise that place and population are inextricably linked with one other. Knowing about where somebody lives, can reveal a lot of information about that person's identity [1]. Commercial geodemographic classifications remain very popular, despite their creation as ‘black box’ systems [2]. Such classifications use closed methods and provide little documentation of the data inputs, weighting and normalisation procedures or the specific methods of clustering. The 2001 Output Area Classification (2001 OAC: [1]), by contrast, is an open geodemographic classification built using 2001 Census data, and has been very widely used across a range of applications. The 2001 Output Area Classification assigns each Output Area in the United Kingdom into one of the following categories, based on the socio-economic characteristics of the population:

- Blue Collar Communities
- City Living Areas
- Countryside

- Prospering Suburbs
- Constrained by Circumstances
- Typical Traits
- Multicultural Areas

However, the Output Area geography has been devised by ONS for the single and explicit purpose of disseminating results of the UK Census of Population. By contrast, address records typically consist of lists of unit postcodes. It is possible to match unit postcoded records to the OAC by using database software. But it requires knowledge of database programming, and there is no software utility which could be used for a simple match of the postcodes to their corresponding OAC codes.

Recently, Office of National Statistics (ONS) has created an open license version of the National Statistics Postcode Directory (ONSPD: [3]), which includes an OAC Code for each unit postcode in the UK. The OACoder software uses the ONSPD and allows the users to read in a CSV file with list of postcodes. It then appends the corresponding OAC code for each of the unit postcodes. For this paper, we have used the latest version of the ONSPD available (snapshot date: November, 2012).

OACoder is written in Java. It uses the standard Java packages to read in a CSV file containing unit postcodes,

codes each unit postcode to the corresponding OAC category, and then writes each of the resultant records into a separate CSV file. OACoder is an open source software and is stored in Figshare (<http://dx.doi.org/10.6084/m9.figshare.156599>). The source code of the software is stored in SourceForge (<http://sourceforge.net/projects/oacoder>). As an open source software, OACoder has reuse potential across a range of applications.

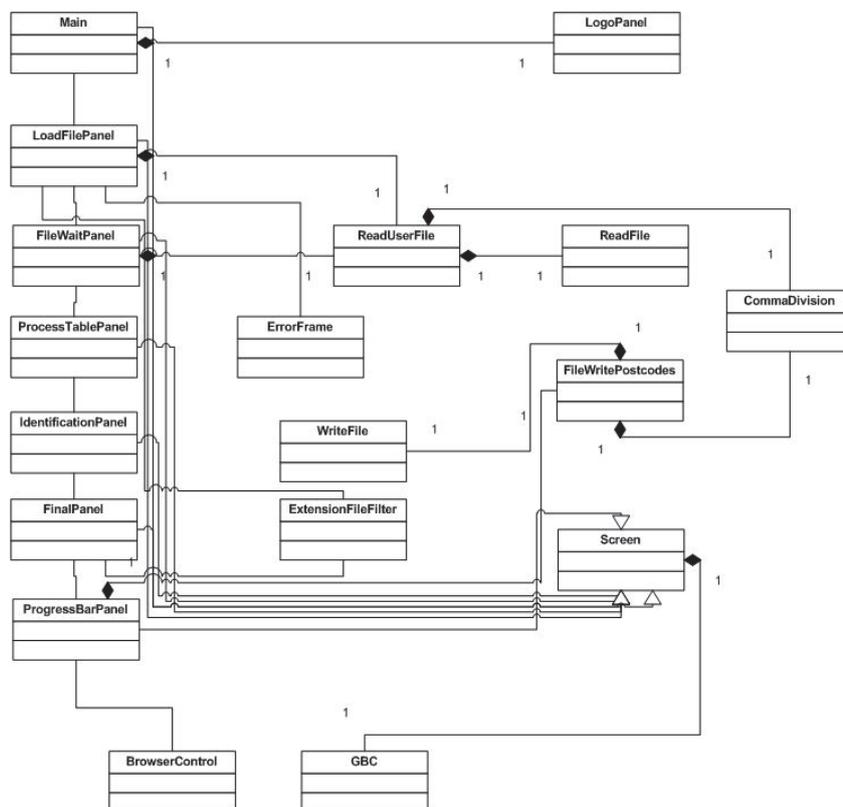
This paper discusses the development and reuse potential of the OACoder software. It outlines the architecture of the software, the testing procedures applied in its development, and the suitable operating environments for running the software. Finally, the paper explains the reuse potential of the different components of the OACoder.

### Implementation/architecture

OACoder was implemented by using the object oriented paradigm, which in practice begins by dividing the software into small components (called Classes). In the later stage, each component of the software is developed. In

an object oriented paradigm, a class diagram is a good way to show the relationship between different components of the software. A class diagram shows the classes that make up a system and the static relationships between them. Classes are defined in terms of their name, attributes (or data), and behaviours (or methods) [4]. The static relationships are association, aggregation, and inheritance.

Classes can be considered as components or individual entities of the system and a class diagram shows each of the components and the relationships between them. A class diagram presents the overall picture of the software or web application, and shows all of the entities that constitute a system. The structure of the classes should be finalized before actual coding of the software or web application begins. This helps the software development team or individual to identify all the objects of the system even before the coding starts. It also shows a clear picture of the software or web application under development. The following **Figure 1** shows the class diagram of the OACoder software:



**Fig. 1:** Class diagram of OACoders.

In a software engineering process, after the architecture of the software has been finalised the next step is to choose an appropriate development technology. A range of development technologies are available for different operating platforms. These include Java, .NET (VB.NET & C#), Visual Basic, and C++. Applications developed in Visual Basic, VB.NET, C++, and C# run only on windows operating systems. However, Java could be used for developing cross platform applications. It means that applications developed in Java can run on different operating systems

e.g. Linux, Unit, Mac OS, and Microsoft Windows. So, Java was chosen for the development of OACoder.

### (2) Availability

#### Operating system

OACoder was developed in Java so that it can run on multiple operating platforms. However, the first iteration of the software was developed to work on Windows operating systems. OACoder was tested on the following operating systems:

- Windows NT, 2000, XP, Vista, 7
- Windows Server 2008

In the future, the functionality of the OACoder can be extended to work on other operating systems e.g. Linux, Unix, Mac OS.

### Programming language

OACoder requires Java 1.5 or higher installed on the computer.

### Additional system requirements

OACoder requires a minimum of 4GB memory installed on the computer. The use of a Dual Core or higher processor is recommended.

### Dependencies

OACoder does not have any dependencies on other libraries. OACoder uses the standard Java packages, so does not need any additional libraries or frameworks..

### List of contributors

- Muhammad Adnan
- Alex Singleton

### Archive

#### Name

Figshare

#### Persistent identifier

<http://dx.doi.org/10.6084/m9.figshare.156599>

#### License

GPL-3.0

#### Publisher

Muhammad Adnan

#### Date published

13/02/2013

### Code Repository

#### Name

SourceForge

#### Identifier

<http://sourceforge.net/projects/oacoder/>

#### License

GPL-3.0

#### Date published

27/06/2012

#### Language

English

## (3) Reuse potential

Software reuse is an important aspect of any open source software. OACoder has been developed in a way that facilitates its use by other researchers and developers working in different fields. By using the object oriented development paradigm, OACoder was divided into different components called classes. Each of the components works as a separate entity and provides an independent

functional component. Different functional components of the OACoder can be reused by researchers in their own fields. Some of the source code reuse suggestions are given below:

- Extension of OACoder to work on other operating systems: A possible enhancement in the functionality of OACoder is to reuse all of its source code to work on different operating systems. OACoder is developed and tested on the Windows operating systems. However, applications developed in Java can run on multiple operating systems. Mac OS and Linux are used quite extensively in research these days, and the ability for OACoder to run on multiple platforms will increase its user base.
- Use of OAC 2011: OAC was developed by using 2001 census data. A new version of OAC based on 2011 census data will soon come out, keeping in mind that ONS has already started the consultation. The new ONSPD data based on OAC 2011 will replace the existing ONSPD data. Hence, a possible reuse potential of OACoder is to reuse its source code and extend the functionality to work with the new version of ONSPD dataset.
- Reading/Writing CSV files: CSV file format is quite popular in research and thus researchers and developers need source code to read and write CSV files. OACoder has separate components for reading and writing large CSV files. These components can be reused to read/write CSV files for different purposes.
- Handling large data files: Source code in OACoder could also be reused to read and write large data files. OACoder reads a user file containing a list of postcodes, and matches it to a list of 1.8 million postcodes in the UK. OACoder does this very efficiently and quickly. This source code could be reused for potential applications where an efficient read/write is required for large data files.

### Acknowledgements

This work was completed as part of the EPSRC research Grant "The Uncertainty of Identity: Linking Spatiotemporal Information in the Real and Virtual Worlds" (EP/J005266/1).

### References

1. **Vickers, D W** and **Rees P H** 2007 "Creating the National Statistics 2001 Output Area Classification". *Journal of Royal Statistical Society, Series A* 170(2): 379-403, DOI: <http://dx.doi.org/10.1111/j.1467-985X.2007.00466.x>
2. **Longley, P A** and **Singleton, A D** 2009 "Classification through consultation: public views of the geography of the E-Society". *International Journal of Geographic Information Science* 23(6): 737-763, DOI: <http://dx.doi.org/10.1080/13658810701704652>
3. **Office for National Statistics** 2013 "National Statistics Postcode Directory". Retrieved 15th February, 2013. From <http://www.ons.gov.uk/ons/guide-method/geography/products/postcode-directories/-nspp/-index.html>

4. **Copeland L** 2008 "Testing UML Models". Retrieved 15th January, 2009. From <http://www.stickyminds.com/sitewide.asp?Function=edetail&ObjectType=ART&ObjectId=6232>

**How to cite this article:** Adnan, M, Singleton, A and Longley P A 2013 OACoder: Postcode Coding Tool. *Journal of Open Research Software* 1:e5, DOI: <http://dx.doi.org/10.5334/511ba2c94d661>

**Published:** 8 October 2013

**Copyright:** © 2013 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License (CC-BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/3.0/>.