How to Characterise a 1914 Architect using Big Data

Adrian Humphris, Wellington City Archives and Geoff Mew, Independent Scholar

ABSTRACT: "Big data" analytics is a means that can be used for examining large and varied data sets to uncover hidden patterns, unknown correlations, market trends, customer preferences and other potentially useful information. The analysis of large data sets has become possible, and much more relevant, as such data has come increasingly accessible, and computing technology has advanced considerably. In the study reported here we have taken the first steps in using this technique to investigate whether big data can tell us more about the characteristics of New Zealand architects and their work. Part of the research for our book *Raupo to Deco* resulted in a large data set of newspaper tenders. Tender notices enabled us to identify architects, the time frames they worked, the types and numbers of buildings they designed, and the connections and relationships between individuals and practices. In addition, we used genealogy techniques to work out birth and death dates, to study obituaries and to track movement patterns. We have continued to build this data set, adding in architects from all over New Zealand and relevant data about them; currently we have around 20,000 individual building tenders for buildings across the country between 1840 and 1940.

This paper analyses our data to give a view of "an architect" in 1914, the year architects in New Zealand were first required to be registered. How many were there? How old were they and how long had they been practicing? And what else can we find out? We also discuss the advantages and pitfalls of dealing with a large data set, and explore how we can ensure the validity and accuracy of the results.

Introduction

Our recently published book on architecture and architects, *Raupo to Deco*, provided an history of the architects practising in Wellington in the first century of European settlement. As important as the styles of architecture and development of the city were, the people, their background and training as well as their experience and work as architects were also considered of major interest. To research such a broad topic a very diverse range of sources needed to be consulted: local authority building permit and property records, historic building plans, biographies, indexes, newspapers and others.

Since publication we have considered whether this study could be extended to other areas of New Zealand, or used to discuss architects or architecture at a national level. Or, what else can we learn or glean from the information compiled to write the book? We have therefore broadened our research into looking at other questions we can answer, or ways our data can be shared and used. One way is to further research the characteristics of architects were, and how they changed over time.

One resource we have created is a database that captures information about both architects and their buildings. This resource is the focus of this paper, which we have used to both present some initial results found using the data, but also to discuss the questions raised in assembling such data on a national scale. We are attempting to utilise the concept of big data; that is using a large data set to define concepts and detect otherwise less obvious trends.

While analysis of large data sets is usually focused on more commercial aspects, such as consumer behaviour and sales, health care or business process optimisation,[1] our focus applies it to historic data to see what characteristics about architects we can find, as

[1] Marr "How is Big Data Used in Practice?" np.

well as looking for changes and trends over time. The ability to do this has become possible, and much more relevant, as the data needed has become increasingly accessible, and the computing technology needed to interrogate and analyse the data has advanced considerably.

Data Sources

So what sources of data can help us to characterise architects or their output? While numerous sources of information exist, we have focused in particular on two; local authority building records and newspaper tender records.

Our earlier book, *Ring Around the City*, analysed the relationship between the extension of tramways and the built development of suburbs in Wellington in the early 1900s. To do this analysis we needed the dates of construction for new dwellings in the suburbs of Kilbirnie and Kelburn between 1900 and 1930. While building permit records were accessible at the Wellington City Council Archives, many weren't linked to a current street address. Historically addresses weren't recorded in many permit records even until the 1950s, so much research was required to find these. As this involved checking the plans, names of builders and architects were also captured where found. Collating this data allowed us to: summarise the numbers of buildings designed by given architects, determine dates when their work first appeared and when it ended, and the types of buildings that specific architects designed.

Similar data was got from newspaper tenders using *PapersPast*. While the level of detail is often far less (for example, the actual street address for the building is only very infrequently given, whether it was built at all, no details of its size and construction), collating this information again allows a data set to be built up around designs over time by given architects.

Both sources also provide a way to identify the actual architects themselves. We only had a very incomplete list to start with, sourced from secondary sources, such as almanacs, trade directories, and New Zealand Institute of Architects records. So, as well as finding detail about the work architects were doing, analysing our two data sources also helped us assemble a comprehensive list of names for further investigation.

To capture as much information as possible, as well as being able to cross-reference some information, we also used a wide range of other sources; for example, publications and theses about architects, local authority heritage inventories, and index sources such as the University of Auckland architecture archive and indexes for other historic and archive collections. While these added to our knowledge, the quantity of data provided was far less than the volume from our two identified sources, and the volume we need to do our research.

Our Data Set

Our data set covers over 21,000 individual buildings across the country between 1840 and 1940, as well as nearly 1,500 architects or architectural firms in practice during this period. These are buildings which were either being tendered for or built, and they are linked to our data set of people who called themselves architects. As our interest has been in the characteristics of the architects as well as their buildings, we have accessed other data sources to find out more about them. These have included street directories, electoral rolls, births deaths and marriage records, probate records, military service records, shipping arrival records, architectural institute records, gazette notices and other

family history sources, not only in New Zealand but worldwide.

Our data is split into two related sets; building information [linked to architect], then information about the architects themselves. The key attributes we have focused on are:

Building information captures:

* Location: Region – City – Suburb and/or Address
* Date: either the permit was issued or the tender appeared
* Description of work: both the actual description, and also a building type category we have assigned for analysis purposes
* Who the building was for: either the owner, or sometimes in permits the applicant for the work
* The architect, or architectural firm: either given in the tender or recorded on the permit plans/specification
* The builder and value of work: mostly from building permit records, although sometimes newspapers record a list of tenderers and their tender price, and which was successful
* Source: newspaper title and date, or building permit reference
* "Notes": any other detail about that building that may help; for example, when the tender closes, if a description of the extent of work was given, any additional initials of draftspeople off the permit plans

Architect information captures:

* Name of the architect or firm
* Date of birth and death
* Date of first and last architectural reference
* Practice locations
* Partnerships or firms they were connected with
* Education, qualifications, and what they 'called themselves': Architect, Building Surveyor, etc
* Other association, memberships, awards or jobs that may be relevant or of interest
* Number of building designs identified, either on their own or as part of a partnership

Note that we can only capture the information that is available for any reference we check. This means we can't always capture all the above attributes for every building or architect. For example, newspaper tenders rarely provide the owner/builder or value of the work, whereas generally this is all provided as part of a building permit application. Similarly, it is sometimes not possible to find precise dates of birth and death. Because of this we have had to use different subsets of our data to answer different questions. For example, while for 1914 we have found just over 300 architects, we only have dates of birth for around 180. Over time this number will increase as our research progresses, but we have assumed our sample size is large enough to be representative of the whole.

Also, we focus on new construction rather than recording every record of work found. We have included significant alterations; for example, an additional wing to a building, or a complete rebuild, but have disregarded minor alterations or additions. We decided to balance the extra time taken it would have taken to capture such data against the assumption that new builds would give a suitable reflection of an architect's work. Therefore, we have assumed there were not significant number of architects who spent their lives doing alterations only.

A First Rough Cut

For this paper we have experimented with our data to try and summarise what we know

about an "architect." We have chosen to look at 1914; the year registration became required under the NZIA Act 1913. We have sought to answer the following:

* How many architects were there, and where were they practising?
* How old were they, and how long had they been in practice?
* Were they in partnerships, or in sole practice?
* What sort of qualifications did they have?

There are several other aspects we are also working on with our research. One is to look at patterns or trends over time; for example, here we are looking at 1914, but how does this compare with 1904, or 1894? Can we compare different periods to see if any trends emerge? For example, does the average age of architects decrease as architectural courses become available locally and increasing numbers of students graduate from them? What happens to the proportion of overseas-qualified architects? And is there a move from small partnerships or individual architects practicing to the establishment of larger firms, with students studying architecture and registering as architects more likely to do draughting work in a firm than set up their own practice?

We are also looking at other ways we can use our data, for example analysing the actual descriptions of buildings recorded. Residential tenders use a wide range of descriptions; dwellings, villas, houses, bungalows, residences. Even residences are subdivided into villa residences through to family and gentlemen's residences. What were the differences? Was there a relationship between the buildings being constructed and the names given, or was it more the fancy of the architect placing the advertisement? And when did different terms come in to use or stop being used?

Further to this, what can building data for each architect tell us about their output; how busy they were, what type of buildings they designed, and so on? Were some architects predominantly designing houses, where others focused on commercial work, or produced a wider range of buildings?

The Architects of 1914

Using our data set we have 307 architects practising in New Zealand in 1914. Around 10% of these were employed by local or central government, the other 90% in private practice. It was, as to be expected, a male-dominated profession. Our only female architect is Lucy Adelaide Greenish, who had registered in 1914 and was working in the office of Atkins & Bacon.

The "average" architect was 42 years old, and had been in practice, or "involved" in architecture for 13 years (remembering our start data for an architect often includes the period they were a pupil or articled to an architect, not necessarily just when they started designing buildings).

The average length of practice was 13 years, although 7% or our architects were in their "first year," and 38% have been in practice for under five years. 75% of architects had been in practice for less than 20 years, while only 25% had been practicing 20 years or longer. A number of suppositions could be made for this pattern of experience, but before any conclusions can be reached we need to make some comparisons with previous periods.

The age range of architects appears much more consistent; although youngest to eldest spans some 70 years (we have an 18-year old William John McKeon working as a draftsman for Hoggard and Prouse in Wellington through to an 88-year old Fitzgibbon Louch nearing the other end of his career in

Auckland), there is no one age group significantly more represented than another. 30% of architects were aged between 31 and 40.

In terms of qualifications, 62% our architects had registered under the NZIA by 1914. A further 4% recorded overseas qualifications, predominantly registration with Australian institutes or the Royal Institute of British architects. This leaves 34% without a record of architectural "qualification."

Looking at the timing of registration under the NZIA, two-thirds of architects had actually joined the NZIA prior to 1913, when registration became compulsory; nearly 20% had actually belonged from 1905. By 1920 the number of architects we have found without registration was just under 20%; some architects had retired by then, while others had gone through the registration process. A number of the architects not registered were teaching or working for central government or local councils; potentially registration was less relevant for them. Others we have found would go on to seek professional standing as members of the Incorporated Association of Architects & Surveyors, which was formed in 1925 and offered another path for training and education.

Our data analysis also indicates how architects worked. Of those in private practice, we found 39 partnerships, accounting for 88 architects, or around one third. A further 12 architects had others working for them, so we could say there were 51 "firms" that accounted for 40% of our architects. The remaining architects appeared to work independently.

Finally, architects appear to have been spread across the country, with individuals found in 39 towns or cities. We have grouped these together by the regions first used by the NZIA; Auckland, Wellington, Canterbury, Otago and Southland; partly to simplify our data, but also to compare how close our numbers are to those in the NZIA 1915 list of registered architects.

Auckland and Wellington accounted for three-quarters of our architects, the final quarter working in the South Island. Canterbury and Otago had similar numbers in each major city. Interestingly half of all architects worked in the Wellington region, compared to a quarter in Auckland. In 1914 Wellington still had most of the government departments and company head offices as well as a relatively large manufacturing base and port.

Our results are similar to analysing the NZIA list from 1915, although we have under-represented Auckland and over-represented Wellington. This may be due to how we have linked our architects to each region, although it has to be remembered we also we have included a large number of "architects" who were not registered. Our numbers are similar however for the split between North and South islands.

Understanding and refining our architects

In the process of writing *Raupo to Deco* we encountered several issues relevant to handling large data sets such as our architect's data.

One of the most basic is the definition of an architect, particularly prior to the Registration Act of 1913. The simplest solution is to assume that because a person put "architect" after his or her name on a tender notice or on the plans submitted for a permit then he or she was one; but we know there was a considerable range in skills and qualifications behind that, even including some duplicity. Objective grading

of qualifications for the nineteenth- and early twentieth-century architects is not easy but probably needs to be attempted.

There can also then be challenges with the genealogical research to find basic data about architect's lives. While this is often relatively straightforward, there are instances when names are misspelt, duplicated, extremely common, or totally or partially absent from the available records. Usually these problems can be overcome by additional, but time-consuming, research.

There are also a number of assumptions that we have to make to allow us to analyse some of our data attributes. For example, we use the first and last recorded activity to help determine an architect's length of practice. While our aim is to validate this period, with the numbers of architects we are finding, it is time consuming to check all of them. There may be cases where architects took long breaks over this period, or worked multiple jobs so architecture was not their main focus. However, to determine things such as the average length of practice at a particular date we have to rely on these two dates at the present time.

Completeness of our data

Another issue relates to the extent and completeness of our data, particularly in assessing total quantity and types of buildings designed by a particular architect during his or her entire career. Although *PapersPast* provides a very large amount of information for us to work through, it is a work in progress and not complete. While many newspapers are online, many provincial newspapers have quite limited date ranges - which could present problems if architects moved around the country at frequent intervals or if they only practised in small centres such as Ekatahuna. We previously overcame this for our major architects by checking tender notices using the comprehensive microfilmed newspaper collection held by the National Library, but this approach is impractical for the volume of data we are now dealing with - so some data subsets are currently incomplete.

Our approach to data collection to date may also affect how complete a record we have. Our research focus was initially specific individuals, mainly Wellington, architects. Since broadening our scope we have still often been searching by specific architect, as opposed to systematically recording tender information by date, starting from the 1840s and working forward. This has meant sometimes we find a tender that suddenly extends a practice by a decade, or on closer inspection find many more tenders than we initially recorded. Could this affect our accuracy? Or given the sheer number of architects and volume of tenders, can we assume on average our data is representative?

For example, we had a single recorded tender of a dwelling designed by a JJ Morley in 1911. Was Mr Morley an architect, and do we include him in our study? Some more specific research means we now know James John Morley was practising as an architect in Hastings from at least 1910-19, and was responsible for 25 buildings over that period. This extra detail suggests he should be included. This also highlights another challenge we face. While looking into Morley we then found a single tender for Grant & Ball, neither of whom were then on our list. Some further researched indicated John Ball as an architect, but Mr Grant remains a mystery – or a builder? While Hawkes Bay now has the *Hastings Standard* online up until 1922, no Napier paper has been scanned after around 1902.

We are also making an assumption that the tenders recorded in papers represent an architect's output reasonably accurately. The reality is that although we are capturing data from a wide range of sources, newspaper tenders are the most practical way to build up the volume of information we need. It would be impractical to visit and research all the building permits from every local authority, even if those records were still complete and extant.

Conclusion

Our preliminary work has shown that previously unrecognised characteristics of a "typical architect" in a particular year can be identified using a subset of big data assembled for New Zealand architects, 1840-1940.

We have determined that it is insufficient to reply on a single source such as tender notices to build a comprehensive database, although such a source forms the foundation for building the big dataset on architects and their buildings we need for our research.

To answer as wide range of questions as possible, it is essential to compile from, and cross-reference, other databases, such as those available from the New Zealand Institute of Architects and Births, Deaths and Marriages. Many other sources also need to be consulted.

It is important to get agreement on definitions in order to make the data manageable in terms of defining who we class as architects. This will help with the validation of data quality and establishing consistency.

The work to date suggests that, with suitable refinement and ongoing addition of data, analysis of big data will provide a powerful tool for finding out more about New Zealand architects and their buildings.

REFERENCES

Humphris, Adrian and Geoff Mew *Ring Around The City, Wellington's New Suburbs 1900-1930* Wellington: Steele Roberts Publishers 2009.

Marr, Bernard "How is Big Data Used in Practice? 10 Use Cases Everyone Must Read" *Bernard Marr & Co* (2017) https://www.bernardmarr.com/default.asp?contentID=1076

Mew, Geoff and Adrian Humphris *Raupo to Deco, Wellington Styles and Architects 1840 - 1940* Wellington: Steele Roberts Publishers 2014.

.