Conventionally the PhD has been positioned as training at the beginning of an academic career, but this original purpose is now under question in many countries. This discomfort about the appropriateness of the PhD, even for its stated purpose of training academics, has a long history. The earliest paper we located on the topic was from 1930 (Dale, 1930) two decades before Australia started offering research training degrees. For at least two decades the changing workforce structures have been noted as a problem for PhD graduates in finding academic employment (Pazeley, 1996). In Australia the number of graduates who end up working outside the academy is likely to be as high as 60% (McGagh, 2016), making it more normal to engage in a non academic career path on graduation than to stay on and become an academic. Governments are starting to recognise the importance of highly trained graduates to innovation, which is increasingly positioned as a future route to increased national prosperity (Access Economics, 2010). In Australia this has translated to policy makers putting pressure on universities to re-think PhD curriculums to make them more ‘industry relevant’. However, there is a paucity of data to guide this transition to an innovation economy with a high number of research workers. Data from employers is mostly anecdotal, consisting of retrospective self report from individual employers, chambers of commerce and professional accreditation bodies. This study aims to fill this gap by examining what non academic employers want by looking at job advertisements; aspirational statements about the ‘ideal candidate’ the company is wishing to hire.

This research builds on an exploratory study which analysed job adverts for roles specifying a PhD as a required or desired criteria in academic roles (Pitt and Mewburn, 2016). This first paper outlined a content analysis approach to understanding what academic employers are looking for and compared them to the ‘researcher development framework’ designed by the Vitae organisation in the UK. This approach had value, but there was one problem to solve first. While academic jobs could be sourced from higher education providers directly, most jobs are advertised through large scale job search sites and newspapers. Our first challenge was to identify a set of ‘PhD shaped’ jobs to form a set which could be subjected to a deeper analysis. One of the problems of analysing the non academic job market is that many of the highly knowledge intensive jobs are hidden. Most employers do not have good exposure to PhD programs or knowledge of what PhD graduates can do and do not routinely add ‘PhD’ as a qualification on their job ad postings. There is a common assumption that PhD graduates are narrow specialists (‘geeks’) who do not have a wide range of transferable skills that would be valuable in contemporary workplaces.

Our challenge was to develop a machine learning algorithm which could sort through a large data set to find these hidden ‘PhD shaped’ jobs. Accordingly, a large data set of job advertisements was sought and provided by the Australian careers search site.
Seek.com.au. The Seek data set contained over 25,000 job advertisements from a wide range of industry sectors, including academia. Obviously this does not represent all jobs on offer in Australia in 2015; no listing of job advertisements will ever be comprehensive as many are not formally advertised. However seek.com.au is the largest job advertisement site in the country, so we would argue this is an excellent sample of the jobs that are potentially available to PhD graduates outside of academia.

The natural language learning process began with the development of an ontology - a set of high level statements describing what a PhD shaped job looks like. The Vitae framework was too large and complex to suit this task, so a new ontology needed to be developed from scratch. Through an iterative process, which began with an expert workshop, nine high level statements were developed. which were used as coding categories. The statements were as follows:

1. Creative Approach to Problems
2. Ability to Perform High Level Analysis & Synthesis
3. Comfortable with Uncertainty
4. Deep Learning & Self Reflective
5. An ‘Expertly’ Trusted Advisor
6. Ability to Work with Minimal Supervision
7. Ability to Act with Integrity in relation to External Governance Requirements
8. Emotional Orientation to the Work
9. Attention to Detail, Quality, & Accuracy

These statements were used to code job advertisements and then a judgment was made about the whole advertisement and it was categorised as either high knowledge intensity, medium knowledge intensity and low knowledge intensity. This data was used to build a concept machine that could ‘read’ and ‘analyse’ the job ads and produce a set of ads matching each of the knowledge intensity criteria. These jobs could be subjected to discourse analysis techniques that enable us to unpack and examine the needs and desires expressed by employers to inform curriculum development and transferable skills training. Further, the large data set enables us to do an industry sector analysis to see where high knowledge intensive jobs are required and in what quantities. At time of writing experiments are still being performed to see if building this machine is possible.

In this paper we report on an initial analysis of a smaller data set of 40 jobs identified in the process of designing and testing the ontology. A discourse analysis of these jobs reveals which of the already well identified skills and attributes of research degree graduates were desired by non academic employers and - most importantly - what sort of skills and attributes may be required that presently fall outside of our ontology. By making this analysis visible to research students, research managers, businesses and government we can find ways to better connect Australian businesses with Australia’s highly skilled research workforce and provide new directions for those engaged in supporting PhD students in their careers post PhD.

Access Economics (2010). Australia’s future research workforce: supply, demand and influence factors

