Content Analysis of Submissions by *Leximancer*

A content analysis was undertaken to identify the main themes, concepts and ideas identified in the submissions. The major findings of this analysis showed:

1. A large number of diverse issues are discussed in the submissions, revealing a rich and densely populated collection of concerns, and indicating a highly complex system of innovation involving many components.

2. There is a consensus about the high value of innovation and its contribution of economic and social life, and recognition of the central importance in Australia of the connections between research, industry and innovation.

3. There is little common purpose between the various elements of Australia's national innovation system or mutual understanding of how it is configured and how connections within it are enabled.

The findings reveal that although the importance of Australia's national innovation system is widely appreciated and contains a large number of different elements, it is a disconnected system where there are few bridges between its major players. The constituents of Australia's national innovation system do not agree about its most important core elements.

The tool of analysis

The analysis was undertaken using Leximancer, a software tool used to find meaning from text-based documents.[1] Leximancer is a small Australian company based in Brisbane.[2] The software automatically identifies key themes, concepts and ideas by data mining large amounts of text, and visually represents information in 'concept maps' showing the main relationships. These relationships can be examined in more detail by exploring major connections. Analysis was undertaken of all submissions to the Review that were in English, not hand-written or diagrammatical. Confidential submissions were not included. In all, 606 submissions were analysed. The maps display the frequency of the concept by the boldness of its text; the extent of the connectedness of the concepts is shown by size of concept point. The closer to the blue end of the spectrum, the more important the concept overall, reflecting frequency and connectedness.

The benefits of content analysis

This form of analysis is highly inclusive inasmuch as all submissions are important. Every submission contributes to the analysis and to overall understanding. The picture that emerges from this analysis captures 'the wisdom of crowds' and is in essence a user-driven representation of Australia's national innovation system. It is drawn by those sufficiently concerned with the national innovation system to present a submission.

---

[1] The analysis was undertaken by Mark Dodgson, Stuart Middleton, David Rooney and Julia Cretchley

[2] www.leximancer.com

The data mining occurs automatically.[3] In the following analysis there is no 'seeding' of particular concepts; they appear only because of their high incidence in the texts.

Findings

The Review's Terms of Reference and Call for Submissions ensured that broad ranging and diverse submissions were received. Figure 1 shows the large and varied number of themes, concepts and ideas discussed in the submissions. Each of the 63 concepts on this map is raised over 1000 times in the submissions (see the ranked list in Table 1). The core concepts and major connections in the analysis are, and lie between, *Research*, *Innovation* and *Industry*. This reveals widespread recognition of the importance of the relationships between them.

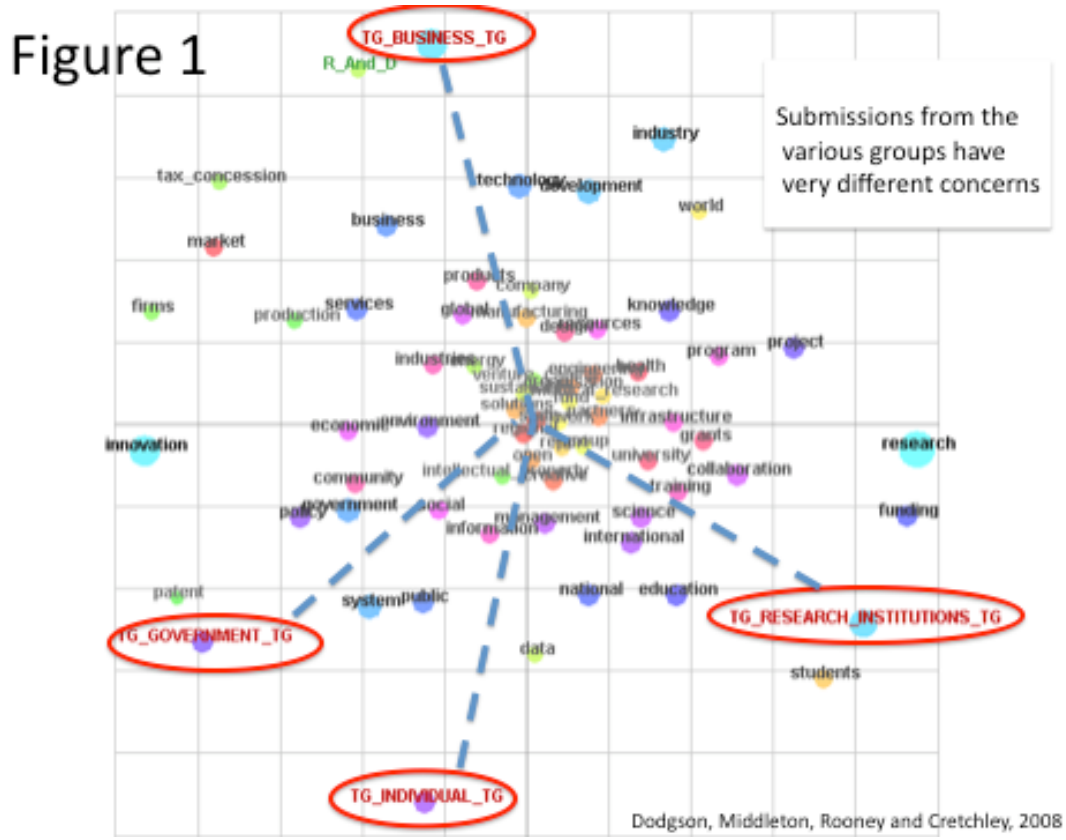
Table 1

Top concepts (mentioned 1000 + times)

research	economic	university
innovation	collaboration	creative
industry	program	company
funding	community	production
development	science	manufacturing
system	management	regional
government	firms	engineering
technology	social	energy
business	global	organisation
public	information	open
services	resources	report
project	infrastructure	solutions
national	world	state
education	training	partners
policy	tax concession	group
R&D	industries	patent
knowledge	data	fund
market	products	medical research
environment	grants	sustainable
students	design	network
international	health	intellectual property

However, the analysis also shows little agreement on what are the core elements connecting Australia's national innovation system. Figure 1 reveals a wide distance between the four groups of submissions: Business, Government, Research Institutes and Individuals. If all four groups agreed on the concepts and features of Australia's national innovation system – and how research, innovation and industry are brought together - the four widely dispersed red ovals in Figure 1 would overlap in a central point.

[3] Although the data mining process is automated, interpretation is required and it relies on the 'craft skills' of its users, fully engaged with the texts and the strengths and limitations of the tool.



Although each group would be expected to represent its understanding of, and position on, issues, the Call for Submissions asked for fresh, forward-looking views and ideas with the aim of improving Australia's national innovation system. It was requesting opinions on a common concern. Whilst diversity is an important feature of healthy and evolving systems, the degree of specialization in concerns shown by the different groups is most valuable when there is evidence of connections, shared concerns and common interests within the national innovation system. Figure 1 shows, to the contrary, that the most important concepts lie at the periphery of the map, and where concepts are shared - moving towards the centre of the map - these are progressively less frequent, connected and important.

This is shown more clearly in Figure 2 which displays the top 50 per cent of the major concepts identified, in practice all those concepts mentioned over 3,000 times in the submissions. The doughnut shape clearly reveals a gap or hole in core of the system.

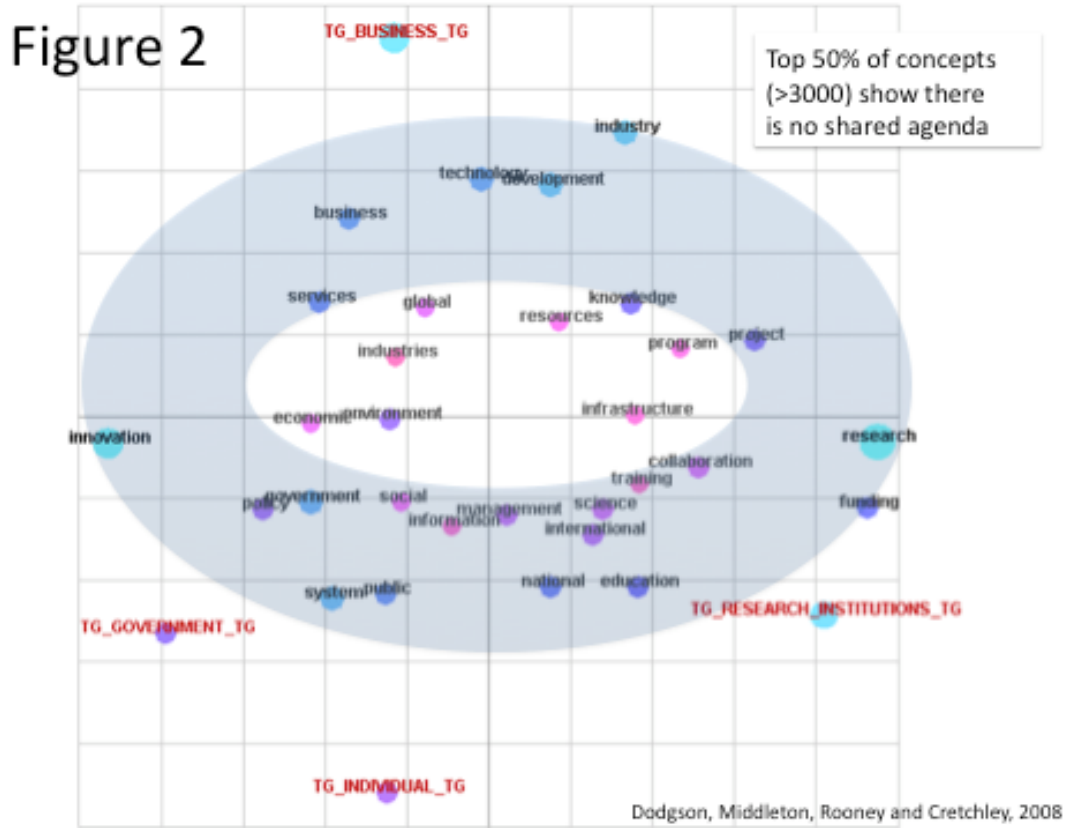


Figure 1 shows *R&D* is very closely allied to the Business group, *Research* is very closely allied to Research Institutions, and *Innovation* is more closely linked to Government. To a major extent this reflects different terminologies used, and the high preponderance of the terms relative to overall submissions within groups. It also shows the absence of a shared language to build common understanding, and may reveal some worrying disconnections. Research Institutions, for example, were 7 times more likely to refer to research, rather than R&D, with its more business-like connotations.

The Business category of submissions was primarily focussed on *industry*, *business*, *technology* and *development*. The Government category identified *system*, *government*, *public* and *policy*. Both these two groups were concerned with *services*. The Research Institutions grouping was primarily concerned with *funding* and *education.* Business and Research Institutions shared a concern with *projects* and *knowledge.*

The most important concepts to be shared by the different groups reveal: recognition of the *economic* importance of innovation, its *global* and *international* characteristics and the need for *resources* and *management*. *Programs* for support and *collaboration* are identified as common interests, although it should be noted that the proximity of *collaboration* to the Research group reveals its comparatively greater concern with this concept than in the other groupings. More detailed analysis showed Research Institutions were twice as likely to refer to collaboration in the submissions, than Business. There is a shared concern for *social* and *environmental* issues. *Science, information, training* and *infrastructure* are commonly identified.

Although all these concepts in the centre of the diagram (within the doughnut) are shared concerns of all groups they are relatively infrequent and unconnected compared to those constructing the doughnut ring.

Leximancer enables the seeding of concepts, to allow for their search and integration into maps. The significance of the most frequently cited institutions in submissions was explored. As might be expected, given its centrality in the Review's Terms of Reference, the *CRC* Program was extensively cited (3113 times), although this was very strongly associated with Research Institutions. The *CSIRO* and *ARC* were seen as more central, however the incidence of their identification was lower (972 and 875 times respectively).

Leximancer offers a tool of analysis rather than prescription. The analysis provides clues to the sorts of public policies and programs needed to build Australia's national innovation system, but its evidence should only be used to support prescriptions derived from other means, such as public discussions, expert opinions and published reports. The analysis supports policies contributing to filling the gaps in the centre of the system. It points to building on those elements already identified by all groups of submission, but not sufficiently important at this stage to enable the scale and connectivity required. The submissions support policies improving the human capital aspects of the system, such education, training, management and knowledge. The submissions identify other resources that play a central role, including technology, infrastructure, services, information and science, with funding being a particular concern to Research Institutes. The role of tax concessions was seen to be important to Business, but ranked only 15th in its list of concepts (it ranked 36th overall in all submissions). The global and international aspect of the national innovation system should be recognized. Collaboration is seen as a key concept, yet it remains strongly within the purview of Research Institutes and better engagement with Business would appear necessary. The submissions reveal very few important institutions connecting the various elements of the system, with only the CSIRO and ARC registering significance. This analysis would complement policy prescriptions derived from other sources that encourage building the institutional framework to better connect Australia's national innovation system.

Mark Dodgson 8th June 2008

The University of Queensland.