Preliminary Analysis of TKF "Book & Bench" texts using unsupervised semantic mapping of natural language with Leximancer concept mapping

RESEARCH UPDATE

Keith G. Tidball, Erika S. Svendsen, Lindsay K. Campbell, Nancy Falxa-Raymond, and Kathy Wolf

In 2011, our research team successfully competed for grant funds from the TKF foundation to explore "Landscapes of Resilience: Understanding the creation and stewardship of open spaces and sacred places in Joplin, MO and Detroit, MI." Within this work we proposed to conduct a preliminary analysis of TKF Open Space Sacred Place (OSSP) journal entries collected from multiple sites from the years 2000-2010. These journals are blank books or diaries located at the sites in which any and all visitors to or users of the site are free to write. This research update represents a brief introduction to the Leximancer approach and method used to conduct the preliminary analysis, followed by "snapshots" of the analysis itself. This research update is not be construed as final results, but rather as indications of our approach and demonstrations of the research products in development.

Leximancer method

Smith and Humphreys (2006), the originators of the Leximancer approach and software for automated content analysis, argue that there are several reasons why one would want an automated system such as the Leximancer software package for content analysis of text. First, they argue that it is known that human researchers and decision makers are potentially subject to influences that they are unable to report (Nisbett & Wilson, 1977). Furthermore, Smith and Humphreys contend, the mitigation of subjectivity in human analysis requires extensive investment of time and money in the content analysis process; code books or dictionaries must be validated, coders must be trained, and intercoder reliability must be tested (see, e.g., Weber, 1990). Therefore, Smith and Humphreys advocate for increasing the automation of this process on grounds of reduced costs and the allowance of more rapid and frequent analysis and reanalysis of text. They also point to how such a system is useful for extremely large quantities of text where there is little possibility of intense human analysis, such as the TKF journals characterized by multiple authors, voluminous text, and multiple sites. An advantage in using Leximancer is that it helps make the analyst aware of the global context and significance of concepts and helps avoid fixation on particular evidence, which may be atypical or erroneous.

The form of semantic mapping supported by Leximancer software has been published elsewhere (A. E. Smith, 2000a; A. E. Smith, 2000b; A. E. Smith, 2003). The Leximancer system performs a style of automatic content analysis, and goes beyond keyword searching by discovering and extracting thesaurus-based concepts from the text data, with no requirement for a prior dictionary. These concepts are then automatically coded into the text, using the thesaurus as a classifier. The resulting asymmetric concept co-occurrence information is then used to generate a 'concept map.'

Journals from TKF OSSPs were collected and transcribed, including entries from 2000-2010 from 33 sites. The journal entries were combined into one text document. The text document was then loaded into the Leximancer software for analysis.

Preliminary Results: themes, sub-clusters, overlaps, and network maps

Eleven themes emerged from the analysis, as depicted below. In this research update, we will briefly elaborate on a select group of these themes.

Theme	Connectivity	Relevance
day	100%	
love	94%	
God	74%	
life	68%	
time	68%	
heart	49%	
bench	40%	
drawing	38%	
walk	26%	
spot	10%	
someone	06%	

Figure 1. Themes, connectivity, and relevance from analysis of TKF OSSP journals.

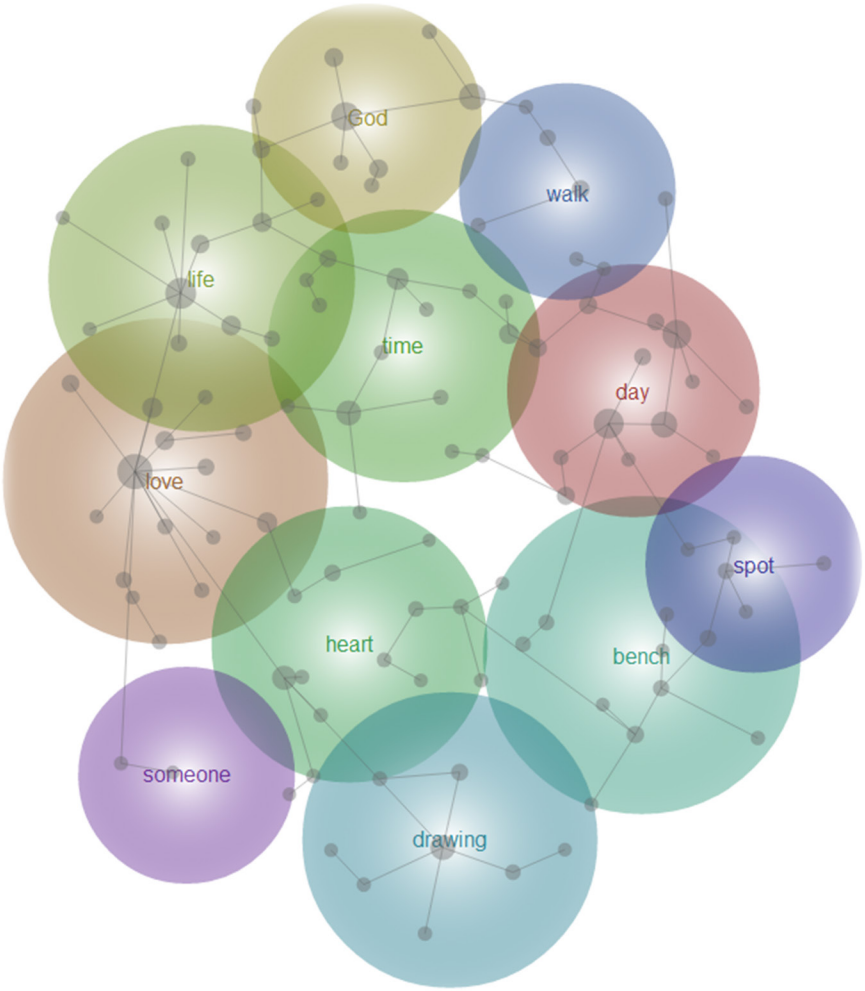


Figure 2. Clusters of themes identified through Leximancer textual analysis. The map is an indicative visualization that presents concept frequency (brightness), total concept connectedness (hierarchical order of appearance), direct interconcept relative co-occurrence frequency (ray intensity), and total (direct and indirect) interconcept co-occurrence (proximity).

It is useful to explore the relationships between the themes to seek further understanding and interpretation of the Leximancer analysis and presentation of the data. It is worth noting that the size of the theme spheres indicates the expansiveness of the theme in terms of the concepts and linkages it contains, rather than its "importance" in a purely hierarchical sense. For instance, though the themes "love" and "life" appear largest, the theme "day" has the largest values for connectivity and relevance.

Sub-clusters of themes

Contemplating Figure 2 one can ascertain two sub-clusters of themes related to the central TKF notion of "sacredness" in the left hemisphere and a more "spatial-temporal" sub-cluster in the right hemisphere. The "sacredness" sub-cluster includes the themes God, life, time, love, heart, and someone, respectively. The "spatial-temporal" sub-cluster includes the themes walk, day, spot, bench, and drawing, respectively. The "spatial-temporal" sub-cluster alludes to key notions and design elements emphasized by TKF in creation of OSSPs.

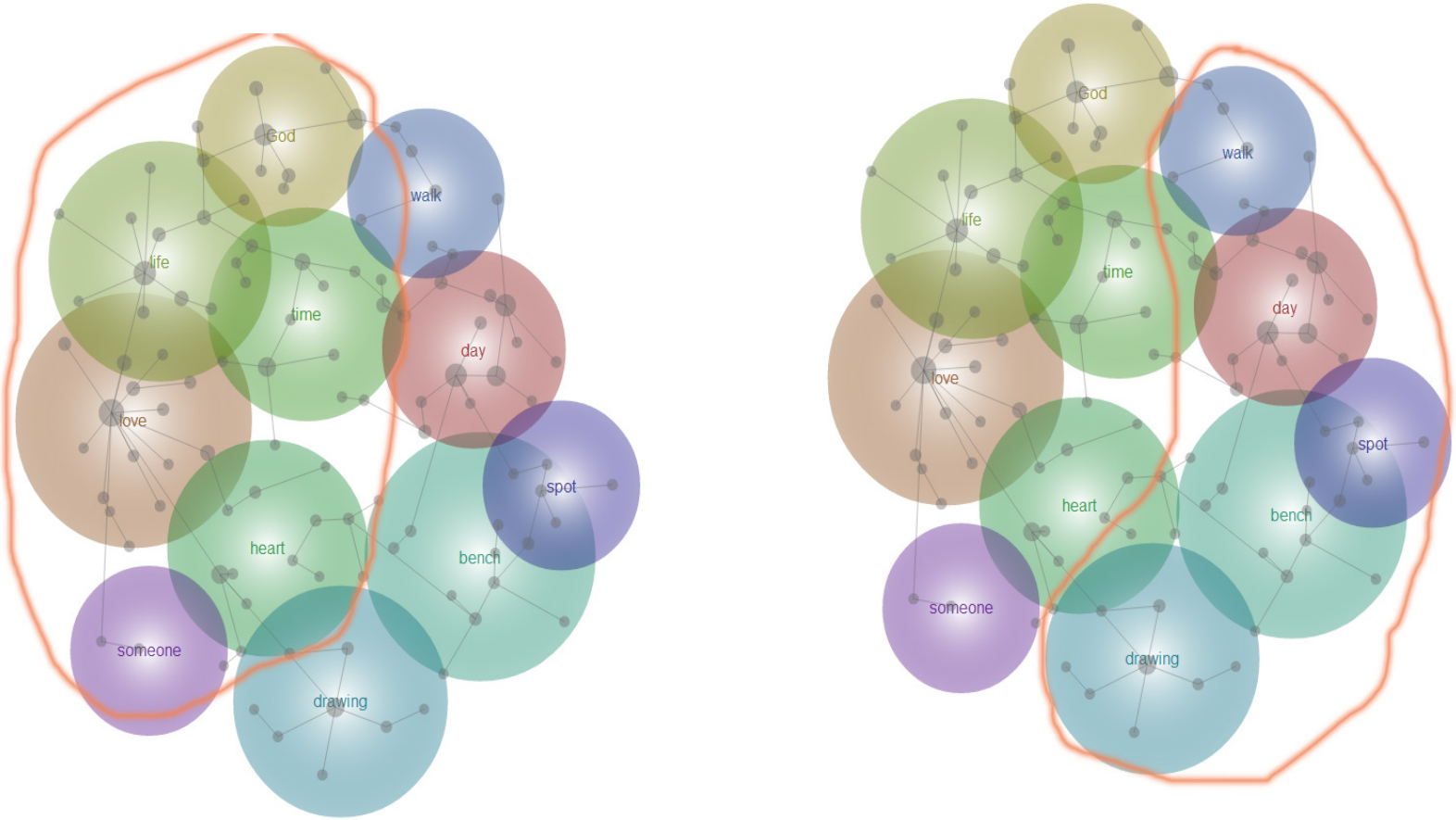


Figure 3. Sub-clusters within the map of themes. On the left is a "sacredness" sub-cluster, and on the right a "spatial-temporal" sub-cluster.

Triple overlaps

There are three occurrences of "triple overlap" of themes. The first occurs within the "sacredness" sub-cluster at the intersection of the three themes "life," "love," and "time." The concept node that is located in the triple overlap is "remember" indicating the importance of personal reflection, remembrance, "memory work," and potential importance of "memorialization mechanisms" (Doss, 2008, 2010; Foote, 1997; Santino, 2006; Tidball, Krasny, Svendsen, Campbell, & Helphand, 2010; Young, 1994) within OSSPs and their host social-ecological systems.

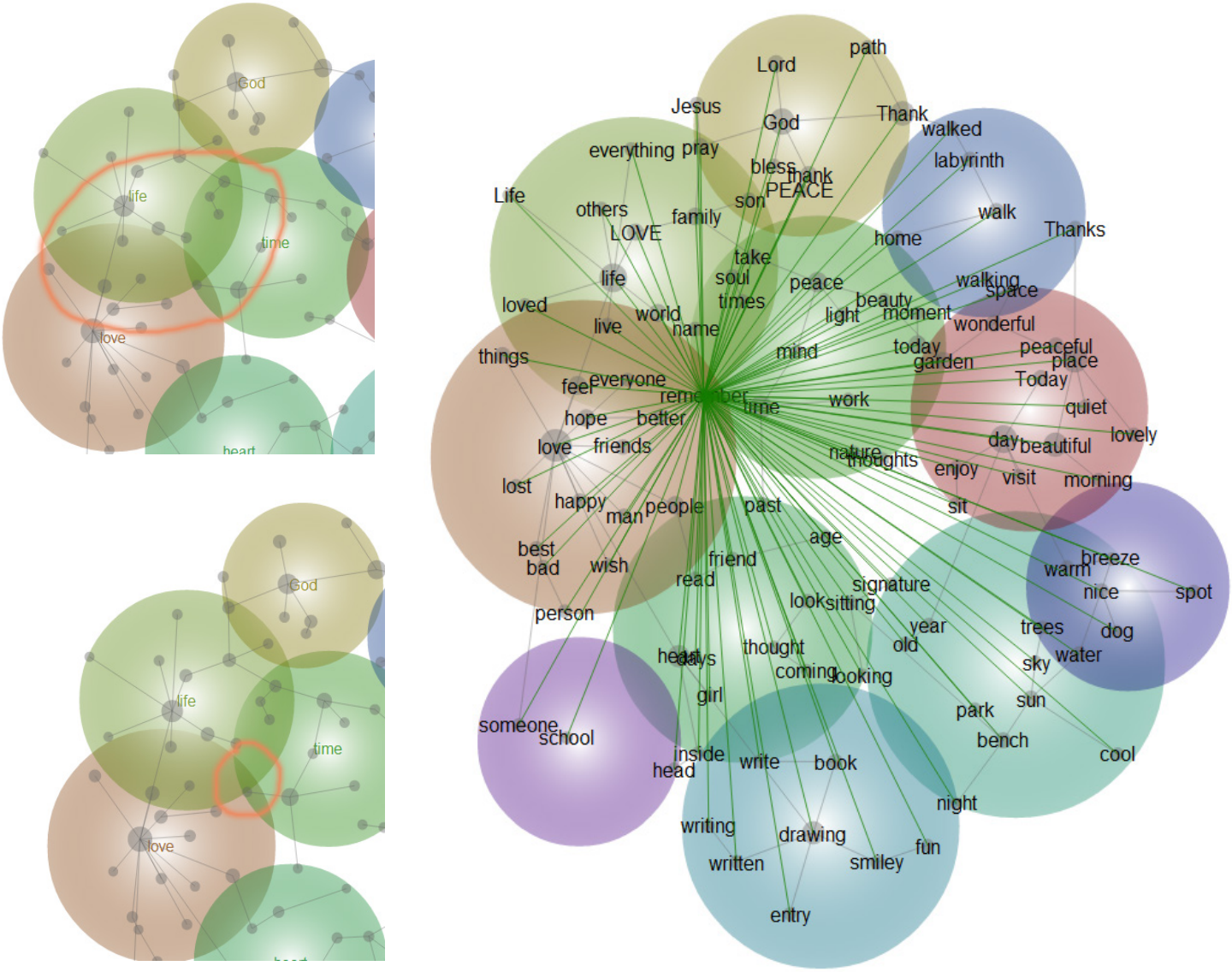


Figure 4. Triple overlap 1, with themes of "love," "life," and "time," overlapping. Note the concept "remember" located in triple overlap region.

The second triple overlap also occurs in the "sacredness" sub-cluster amongst the themes "life," "God," and "time." There is no concept node at the triple overlap point, though nearby concepts include "soul," "son," "times" and "take," possibly relating to broad themes of spirituality and sacredness.

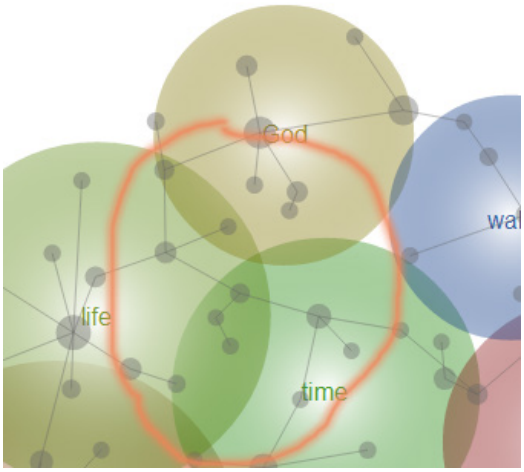


Figure 5. Triple over lap 2, with themes "life," "time," and "God" overlapping.

The third triple overlap occurs in the "spatial-temporal" sub-cluster among the themes "day," "spot," and "bench." Once again there is no concept node at the point of triple overlap. This overlap references OSSP users' experiences interacting with a particular space, place, and time.

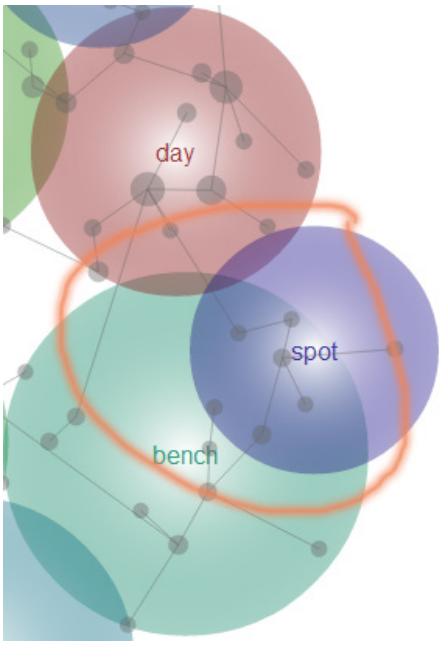
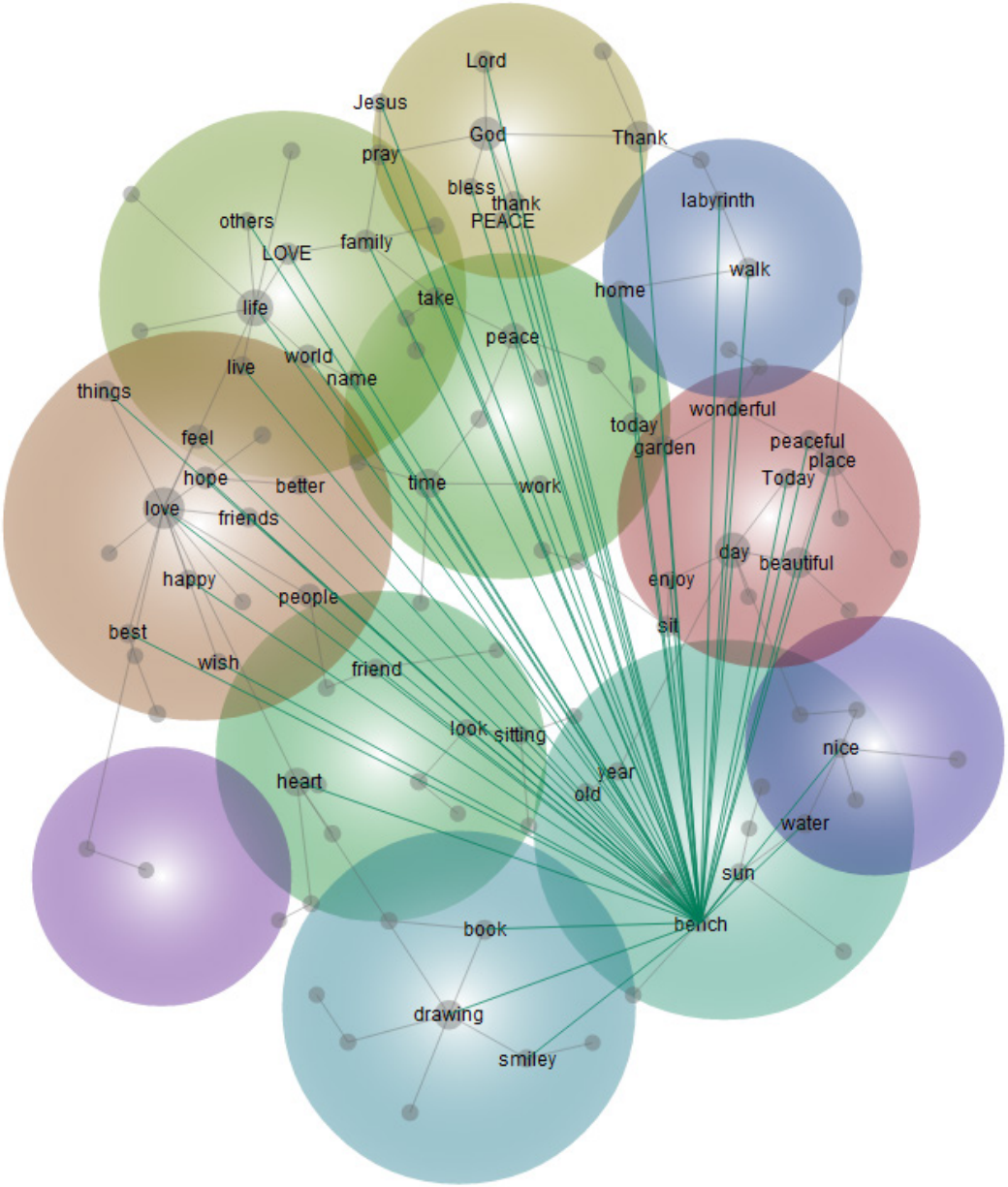


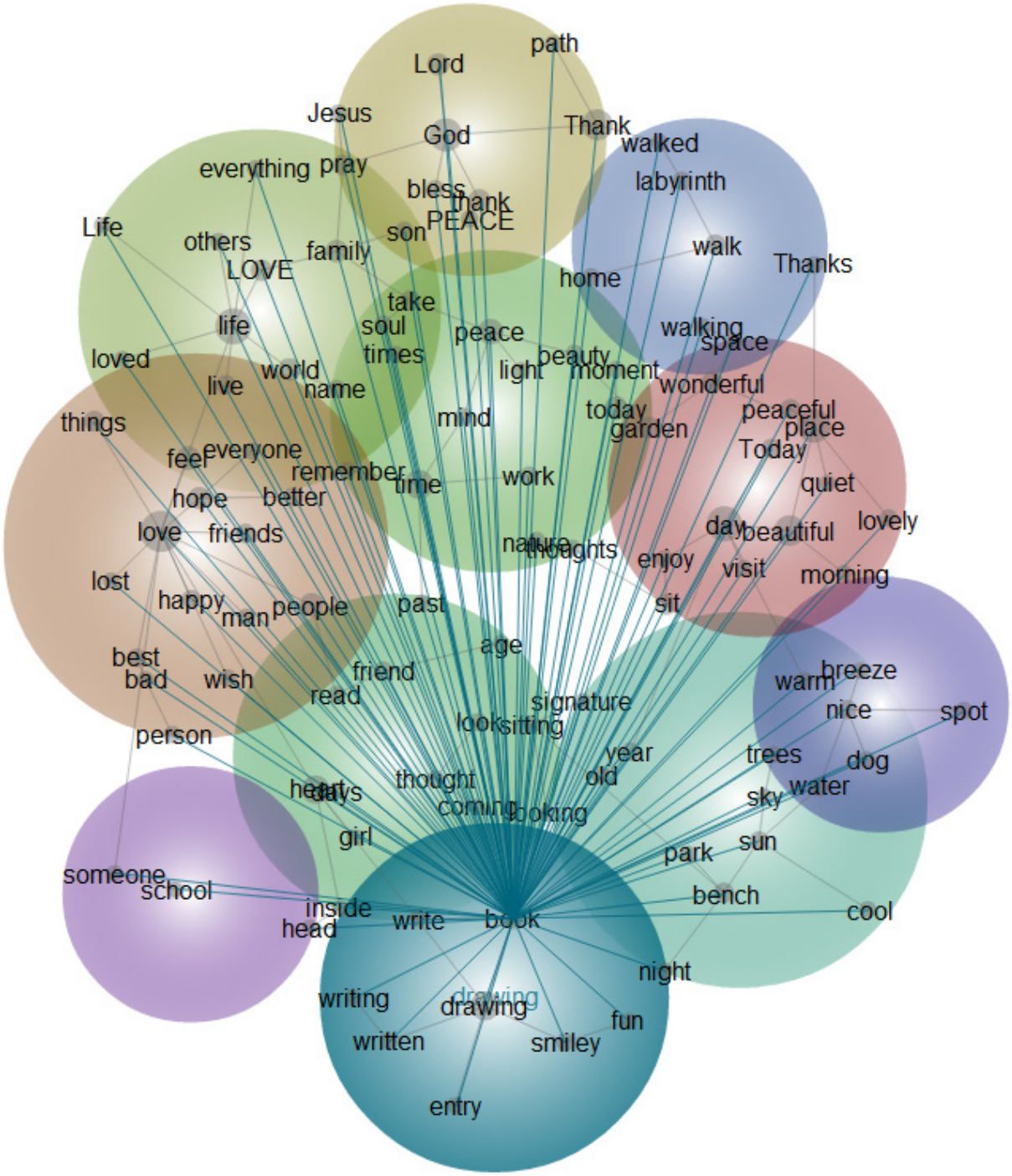
Figure 6. Triple overlap 3, with themes "day," "spot," and "bench."

The eleven themes that the Leximancer analysis identified present multiple and varied opportunities for additional research. In this research update we have explicitly chosen to focus on provocative examples from the textual analysis and have selected a small sample that best illustrate linkages to TKF goals and key philosophical and aesthetic assumptions. Though serving as examples from the multiple entries under particular themes, we acknowledge that these examples are purposively chosen rather than randomly sampled. This does not reduce the utility or accuracy of the automated Leximancer approach, but is intended to supplement it by relating directly to the subjects of greatest interest. The following are concept maps generated by focusing on specific TKF themes of known interest.

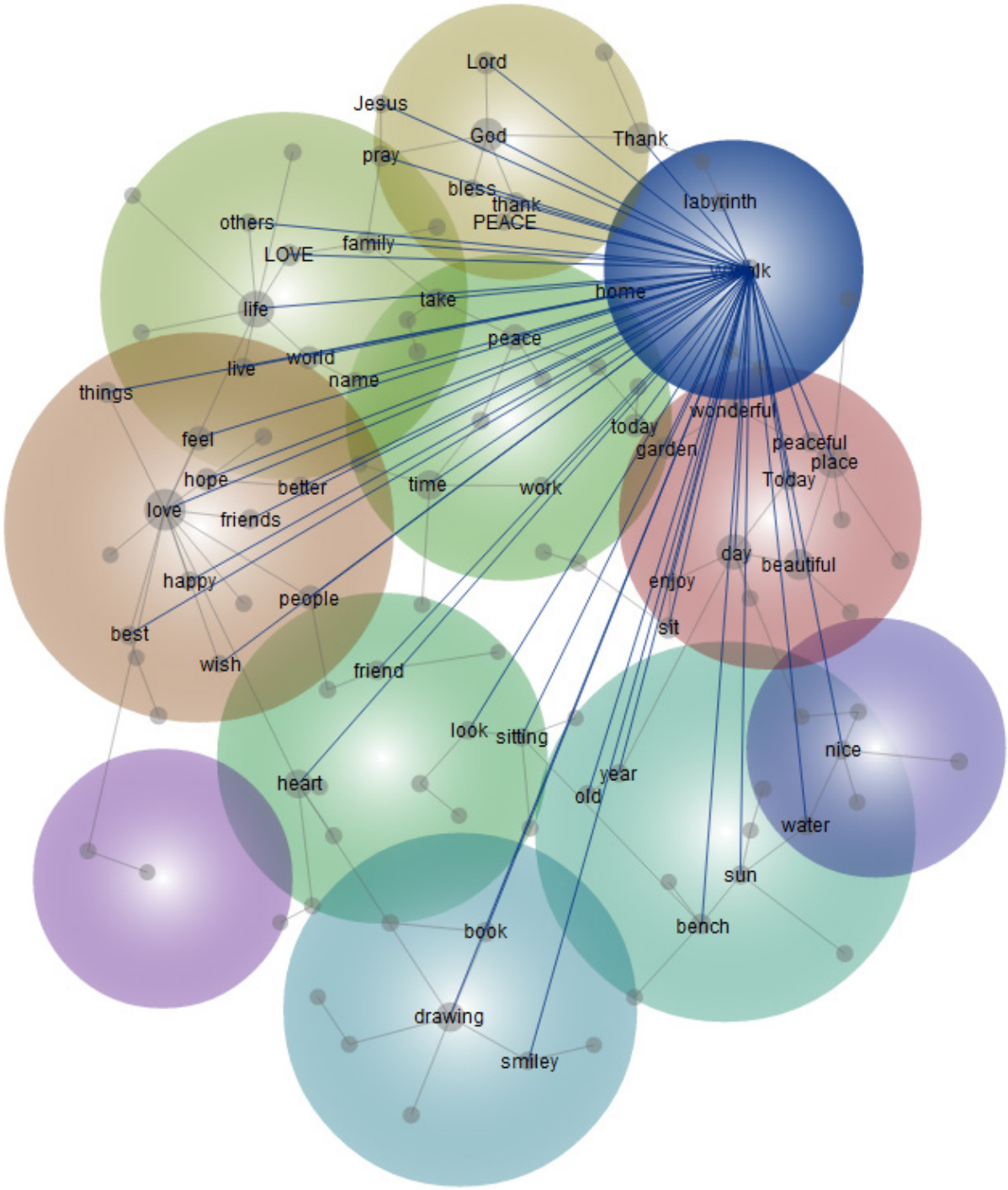
Bench



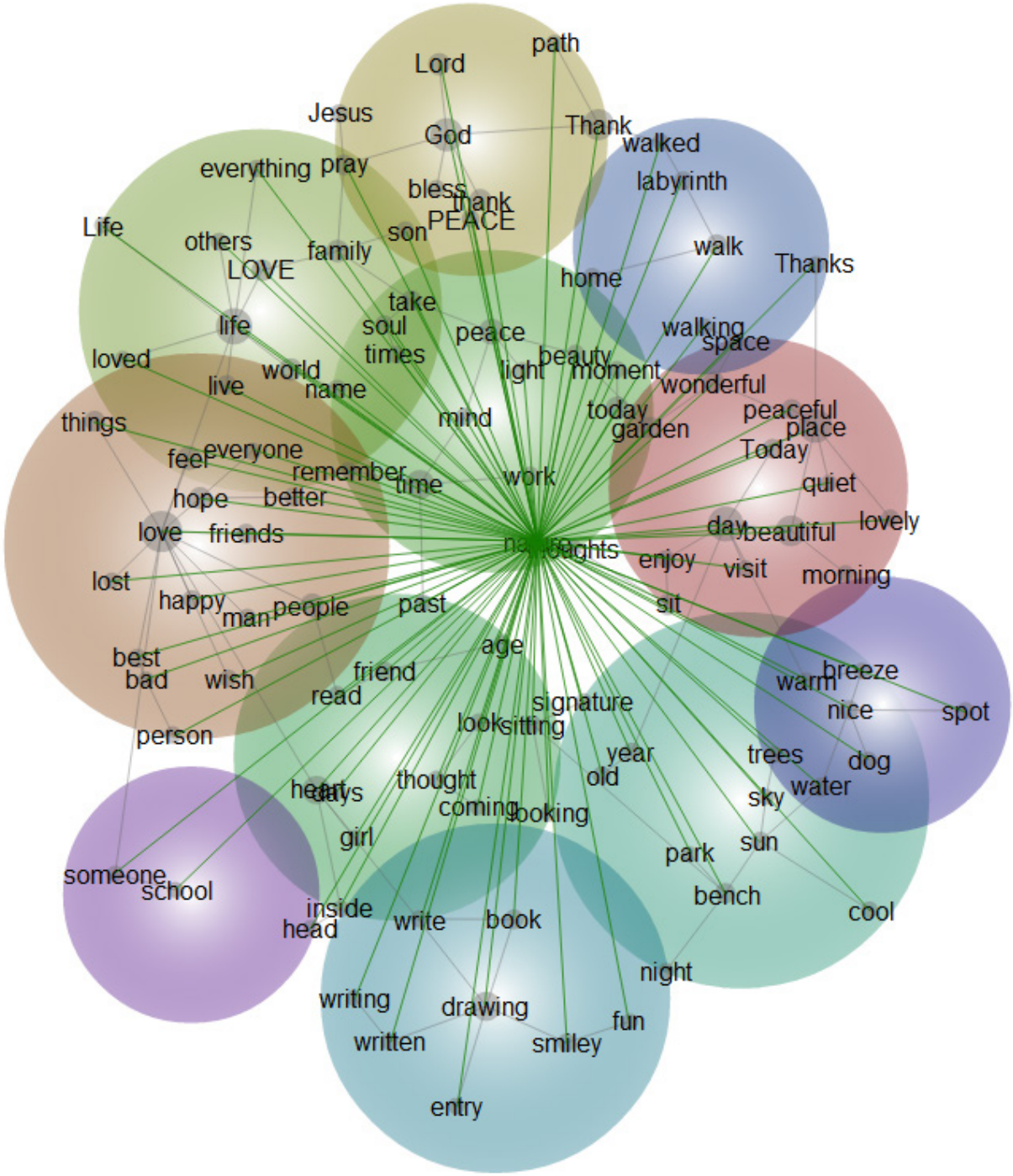
Book


path
Lord
Jesus
God
Thank
walked
labyrinth
everything
pray
bless
thank
PEACE
son
Life
others
family
LOVE
walk
Thanks
home
take
life
soul
peace
walking
space
loved
times
beauty
light
moment
wonderful
world
live
name
things
mind
today
garden
peaceful
place
Today
feel
everyone
remember
time
work
quiet
hope
better
love
friends
day
beautiful
lovely
nature
thoughts
enjoy
visit
morning
lost
happy
man
people
past
sit
age
best
bad
wish
friend
read
breeze
warm
signature
nice
spot
looksitting
person
year
trees
dog
thought
old
water
sky
coming
looking
girl
park
sun
someone
school
bench
inside
head
write
book
cool
night
writing
drawing
fun
written
smiley
entry

Walk (path)


Lord
Jesus
God
Thank
pray
bless
thank
PEACE
labyrinth
others
family
LOVE
life
take
home
peace
world
live
name
things
feel
hope
better
time
work
love
friends
today
garden
wonderful
peaceful
place
Today
day
beautiful
enjoy
sit
happy
people
best
wish
friend
look
sitting
heart
year
old
nice
water
sun
bench
book
drawing
smiley

Nature (note the centrality of the concept…for this type of analysis, this is perhaps the most significant finding)



Works Cited

Doss, E. (2008). *The Emotional Life of Contemporary Public Memorials: Towards a Theory of Temporary Memorials*. Amsterdam: Amsterdam University Press.

Doss, E. (2010). *Memorial Mania*. Chicago: University of Chicago Press.

Foote, K. E. (1997). *Shadowed Ground: America's Landscapes of Violence and Tragedy*. Austin: University of Texas Press.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

Santino, J. (2006). *Spontaneous Shrines and the Public Memorialization of Death*. New York: Palgrave MacMillan.

Smith, A. E. (2000a). Machine learning of well-defined thesaurus concepts. In A. H. Tan & P. S. Yu (Eds.), *Proceedings of the International Workshop on Text and Web Mining* (pp. 72-79). Melbourne: PRICAI.

Smith, A. E. (2000b). Machine mapping of document collections: The leximancer system *Proceedings of the Fifth Australasian Document Computing Symposium.* Sunshine Coast, Australia: DSTC.

Smith, A. E. (2003). Automatic extraction of semantic networks from text using Leximancer *HLT-NAACL 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics: Companion volume (pp. Demo23- Demo24)*. Edmonton: ACL.

Smith, A. E., & Humphreys, M. S. (2006). Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods, 38*(2), 262-279.

Tidball, K. G., Krasny, M., Svendsen, E., Campbell, L., & Helphand, K. (2010). Stewardship, Learning, and Memory in Disaster Resilience. *Environmental Education Research (Special Issue, Resilience in social-ecological systems: The role of learning and education), 16*(5), 341-357.

Weber, R. (1990). *Basic content analysis*. Newbury Park, CA: Sage.

Young, J. E. (1994). *The texture of memory: Holocaust memorials and meaning.* New Haven: Yale University Press.