BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Tanja Bekhuis

POSITION TITLE: CEO and Founder, TCB Research and Indexing LLC; Principal Scientist, Evidence in Documents, Discovery and Analytics (EDDA) Group™; Adjunct Faculty, Biomedical Informatics, University of Pittsburgh School of Medicine

eRA COMMONS USER NAME (credential, e.g., agency login): TBEKHUISPI

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Florida International University, Miami, FL, USA	BA	05/1973	Liberal Studies, Natural Sciences, Psychology
University of Miami, Coral Gables, FL, USA	MS	05/1982	Educational Research
University of North Carolina, Chapel Hill, NC, USA	PhD	05/1991	Quantitative Psychology
University of Pittsburgh, Pittsburgh, PA, USA	MLIS	12/2005	Library and Information Sciences
University of Pittsburgh, Pittsburgh, PA, USA	Postdoctoral Scholar	09/2010	Biomedical Informatics

A. Personal Statement

I am the CEO and Founder of TCB Research & Indexing LLC, a company that supports authors and investigators in the health and social sciences. I am also the Principal Scientist of the Evidence in Documents, Discovery and Analytics (EDDA) Group™ (www.tcbinfosci.com/edda). At the University of Pittsburgh, I am a member of the Adjunct Faculty, Biomedical Informatics, School of Medicine. I am a distinguished member of the Academy of Health Information Professionals (AHIP).

My informatics research is about discovery of knowledge in texts to support evidence synthesis. In 2016, I completed a project funded by the NIH US National Library of Medicine. It had to do with reducing the labor associated with screening thousands of titles and abstracts for inclusion of studies in systematic reviews. The classification task pertains primarily to nonrandomized and observational studies. We used text-mining methods, natural language processing, and knowledge representation approaches.

<…>

I am particularly well suited to serve as Principal Investigator for this <…> proposal because of my multidisciplinary experience. While I raised my family and before I began my informatics career, I served as a statistical consultant and programmer, measurement specialist, and freelance science writer and technical editor. I have formal training in research methods and design, library and information sciences, and language technologies. I also have experience as an evaluator of systematic reviews for the Centre for Reviews & Dissemination (University of York, UK) and the *Journal of Evidence-based Dental Practice*. I am a member of the Cochrane Oral Health Group and the Cochrane Information Retrieval Methods Group. Since 2008, I have collaborated with Dr. Heiko Spallek (Dean of the Dental Faculty, University of Sydney, Australia) on several projects, including one involving analysis of more than 14,000 text messages posted to an online community of dentists; for that study, I developed mixed methods combining qualitative analysis with automated, natural language processing. At the university, I led a software development team <…>.

B. Positions and Honors

Positions (related to information sciences and biomedical informatics)

2006 - 2007	Adjunct Faculty, Department of Biology, Bioinformatics Lab, Juniata College, Huntingdon, PA, USA
2007 - 2008	Adjunct Faculty, College of Information Science & Technology, Drexel University, Philadelphia, PA, USA
8/2008 – 9/2010	NLM/NIDCR Postdoctoral Scholar, Dept. of Biomedical Informatics/Center for Dental Informatics, University of Pittsburgh, Pittsburgh, PA, USA
6/2009 - 7/2009	Postdoctoral Scholar, NLM Rotation for Medical Informatics, National Center for Biomedical Communications, US National Library of Medicine, Bethesda, MD, USA
10/2010 - 6/2011	Postdoctoral Associate, Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA, USA
7/2012 - 2016	Assistant Professor, Dept. of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA, USA (primary appointment)
7/2013 - 2016	Assistant Professor, Dept. of Dental Public Health, University of Pittsburgh, Pittsburgh, PA, USA (secondary appointment)
3/2014 - 2016	Director of Translational Research Methods, Center for Informatics in Oral Health Translational Research (CIOHTR), School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, USA
2015 - 2016	Assistant Professor, School of Information Sciences, University of Pittsburgh, Pittsburgh, PA, USA (secondary appointment)
2016 - present	Adjunct Faculty, Dept. of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA, USA
2016 - present	CEO and Founder, TCB Research & Indexing LLC, Pittsburgh, PA, USA
2016 - present	Principal Scientist, Evidence in Documents, Discovery and Analytics (EDDA) Group™, a division of TCB Research & Indexing LLC, Pittsburgh, PA, USA

Honors

1984 - 1985	National Institute of Mental Health, Predoctoral Fellow
1988	Frank Porter Graham Innovative Research Award, University of North Carolina at Chapel Hill
1989	Smith Graduate Research Award, University of North Carolina at Chapel Hill
1990	L. L. Thurstone Psychometric Laboratory Award, University of North Carolina at Chapel Hill
1992	AERA grant, Institute on Statistical Analysis for Education Policy, Atlanta, Georgia
8/2008 - 10/2010	National Library of Medicine / National Institute of Dental and Craniofacial Research, Pittsburgh Biomedical Informatics Training Program 5T15LM007059, Postdoctoral Scholar
6/2009 - 7/2009	NLM Rotation for Medical Informatics, Communications Engineering Branch, Lister Hill National Center for Biomedical Communications, US National Library of Medicine, Bethesda, MD, Postdoctoral Scholar
7/2011 - 12/2016	NIH/NLM Pathway to Independence Award (K99/R00)
8/2013 - Present	Academy of Health Information Professionals (AHIP), Distinguished Member

Other Experience and Professional Memberships

1981-present	Phi Kappa Phi Honor Society
2001-present	National Association of Science Writers

2005-present	Medical Library Association
2008-present	Beta Phi Mu, International Library and Information Studies Honor Society
2008-present	American Medical Informatics Association
2008-present	Cochrane Oral Health Group, Cochrane Collaboration
2012-present	Association for Computational Linguistics
2013-present	Academy of Health Information Professionals (AHIP), Distinguished Member
2013-present	Association for Information Science and Technology
2013-present	Cochrane Information Retrieval Methods Group, Cochrane Collaboration
2014-present	American Medical Writers Association
2016-present	Special Libraries Association

C. Contribution to Science

Link to My Bibliography: www.ncbi.nlm.nih.gov/myncbi/collections/mybibliography

C.I. Machine learning studies towards a decision-support system for systematic reviewers

i. **Bekhuis T**, Demner-Fushman D. Towards automating the initial screening phase of a systematic review. Stud Health Technol Inform. 2010, 160(Pt 1):146-50. PMID: 20841667.

ii. **Bekhuis T**, Demner-Fushman D. Screening nonrandomized studies for medical systematic reviews: A comparative study of classifiers. Artif Intell Med. 2012 Jul;55(3):197-207. Epub 2012 Jun 5. PMCID: PMC3393813. [**Featured in the AHRQ Scientific Resource Center Methods Library**, November 2012].

iii. **Bekhuis T**, Tseytlin E, Mitchell K, Demner-Fushman D. Feature engineering and a proposed decision-support system for systematic reviewers of medical evidence. PLoS One. Epub 2014 Jan 27; 9(1):e86277. doi:10.1371/journal.pone.0086277. PMCID: PMC3903545. [**Featured in the AHRQ Scientific Resource Center Methods Library**, March 2014]. As of December 20, 2016, number of article views = 2,698.

iv. **Bekhuis T**, Tseytlin E, Mitchell KJ. A Prototype for a Hybrid System to Support Systematic Review Teams: A Case Study of Organ Transplantation. IEEE International Conference on Bioinformatics and Biomedicine (BIBM) Proceedings. Nov 2015; 940-47. doi: 10.1109/BIBM.2015.7359810. PMCID: PMC4742277. [Featured in the AHRQ Scientific Resource Center Methods Library, February 2016].

We developed **open source code for topic models and regular expressions** to extract features for machine learning studies:

v. **Bekhuis T,** Mitchell KM, Tseytlin E. Evidence in Documents, Discovery, and Analysis (EDDA) Group. EDDA Extensions for Binominal Text Classification (Topic Models and Regex), released 16 May 2013. Freely available at http://marketplace.rapid-i.com/EDDA . As of December 20, 2016, number of downloads = 5,656.

C.II. Taxonomy of study designs

Early work for the EDDA Study Designs Taxonomy grew out of content analyses of classification tools developed by organizations such as AHRQ and the Cochrane Collaboration. The taxonomy is useful for machine learning studies described in Section C.I. (above) and for information professionals. Foundational work was reported in:

i. **Bekhuis T**, Demner-Fushman D, Crowley RS. Comparative effectiveness research designs in MeSH and Emtree: an evaluation of coverage. JMLA: J Med Libr Assoc. 2013;101(2):92-100. PMCID: PMC3634392. [**Featured in the AHRQ Scientific Resource Center Methods Library**, May 2013].

ii. A **crosswalk between MeSH and Emtree for study designs** is available for information professionals via PMC3634392 as an additional file to Bekhuis et al, JMLA, 2013;101(2):92-100.

An **improved version of the taxonomy** is available via the NCBO BioPortal:

iii. **Bekhuis T,** Tseytlin E. EDDA Study Designs Taxonomy, version 2.0, released 16 July 2016. Creative Commons Attribution–NonCommercial–ShareAlike 4.0 International License (CC BY-NC-SA 4.0). http://bioportal.bioontology.org/ontologies/EDDA.

C.III. Evidence-based dentistry and oral health

i**.** Spallek H, Song M, Polk DE, **Bekhuis T**, Frantsve-Hawley J, Aravamudhan K. Barriers to implementing evidence-based clinical guidelines: A survey of early adopters. J Evid Based Dent Pract. 2010, (10)4, 195-206. PMCID: PMC3011934.

Papers ii and iii are related in that the paper by Bekhuis et al (2011) devised the methods to carry out the paper by Song et al (2013). The BMC paper is about the connection between oral health and systemic disease. It has **"one of the highest ever scores"** computed by Altmetrics for the BMC Oral Health Journal.

ii. **Bekhuis T**, Kreinacke M, Spallek H, Song M, O'Donnell AJ. Using natural language processing to enable in-depth analysis of clinical messages posted to an Internet mailing list: A feasibility study. J Med Internet Res. 2011, 13(4):e98. PMCID: PMC3236668.

iii. Song M, O'Donnell JA, **Bekhuis T**, Spallek H. Are dentists interested in the oral-systemic disease connection? A qualitative study of an online community of 450 practitioners. BMC Oral Health. 2013 Nov 21;13(1):65. [Epub ahead of print] PMCID: PMC3924341. [**Highly accessed article**; also featured in British Dental Journal. 2013; 215(11): 545, News section, online: 6 December 2013. doi:10.1038/sj.bdj.2013.1152.

iv. Straub-Morarend C, Wankiiri-Hale C, Blanchette D, Lanning, SK, **Bekhuis T,** Smith B, Brodie, AJ, Oliveira D, Handysides RA, Dawson D, Spallek H. U.S. Dental students' evidence-based practice knowledge, perceptions and behavior: a cross-sectional study. Journal of Dental Education. 2016; 80(4):430-38.

C.IV. Methods papers

Staff scientists at Lister Hill, US National Library of Medicine, share the narrative review (i) with investigators who are new to text mining:

i. **Bekhuis T**. Conceptual biology, hypothesis discovery, and text mining: Swanson's legacy [**in top ten of most accessed articles**]. Biomedical Digital Libraries, 2006, 3:2. PMCID: PMC1459187. As of January 27, 2014, number of article views = 17,020.

ii. **Bekhuis T,** Kim K, Valappil B, Spallek H. Adapting the Scottish PRIME Fissure Sealant Questionnaire for implementation research in the United States: lessons learned. APA PsycEXTRA [subscription database], American Psychological Association (APA). 2014. doi: 10.1037/e502572014-001.

iii. Frazier JJ, Stein CD, Tseytlin E, **Bekhuis T.** Building a gold standard to construct search filters: a case study with biomarkers for oral cancer. JMLA: J Med Libr Assoc. 2015 Jan;103(1):346-354. doi:10.3163/1536-5050.103.1.005.

iv. Bramer WM, Giustini D, de Jonge GB, Holland L, **Bekhuis T.** Deduplication of database search results for systematic reviews in EndNote. JMLA: Journal of the Medical Library Association. 2016 Jul;104(3):240-3. doi: 10.3163/1536-5050.104.3.014. PMID: 27366130.

v. McIvor, WR, Banergee A, Boulet JR, Weinger MB, **Bekhuis T,** Tseytlin E, Torsher L, Simulation-Based Assessment Research Group. A taxonomy of scenario delivery and documentation deviations during high-fidelity simulations. Simulation in Healthcare. 2016 Dec (published online ahead of print).

C.V. Cochrane protocol and evaluations of systematic reviews

i. **Bekhuis T**, Thyvalikakath T, Oliver R. Interventions for treating ameloblastomas of the jaws [protocol]. Cochrane Database of Systematic Reviews, 2009, Issue 4. Art. No.: CD003975. doi: 1002/14651858.CD003975.pub2.

Note: Evaluations I wrote for the Centre for Reviews & Dissemination (University of York, UK) were published anonymously in the Database of Abstracts of Reviews of Effects (DARE). The following invited evaluations were published in a dental journal:

ii. **Bekhuis T**. Music therapy may reduce pain and anxiety in children undergoing medical and dental procedures [invited evaluation of a systematic review]. Journal of Evidence-based Dental Practice, 2009, 9(4), 213-14. doi:10.1016/j.jebdp.2009.03.002. PMCID: PMC2778574.

iii. **Bekhuis T**. Chlorhexidine varnish may prevent dental caries in children and adolescents [invited evaluation of a systematic review]. Journal of Evidence-based Dental Practice, 2011, 11(2):84-6. PMID: 21605831.

D. Research Support

Active Research Support

NIH/NIDCR U01 (Clinical Trial; Brad Rindal and Heiko Spallek) 8/2/2016 - 7/31/2020
Decision Support System for Tobacco Cessation
Role: Consultant

NSF Innovation Corps (I-Corps) at Carnegie Mellon University Fall 2015 - present
EDDA Analytics Group™
Role: Lead and Principal Scientist

Completed Research Support

4R00LM010943-01A1 (Tanja Bekhuis) 07/01/12 - 12/31/16
NIH/NLM Pathway to Independence Award (R00)
Screening Nonrandomized Studies for Inclusion in Systematic Reviews of Evidence
It is broadly hypothesized that (a) methods based on natural language processing and machine learning can be used to automatically identify topically relevant studies with a mix of nonrandomized designs eligible for inclusion in systematic reviews; and (b) machine performance can consistently reach current human standards with respect to identifying eligible studies.
Role: PI (sole)

1R21DE021494-01 (Heiko Spallek) 02/24/11 - 01/31/13
NIH/NIDCR
Implementing Research Findings and Evidence-Based Interventions into Real-World Dental Practice Settings
Dentistry lags behind medicine with respect to promotion of research findings and evidence-based (EB) knowledge in clinical settings. The goal of this research was to develop informatics solutions for efficient delivery of EB information about better treatments to general dentists at the point of care.
Role: Co-Investigator

1K99LM010943-02 (Tanja Bekhuis) 07/01/11 - 06/30/12
NIH/NLM
Screening Nonrandomized Studies for Inclusion in Systematic Reviews of Evidence
See description above for 4R00LM010943-01A1.
Role: PI (sole)