Chaomei Chen

Turning Points

The Nature of Creativity

Chaomei Chen

Turning Points

The Nature of Creativity

With 82 figures, 18 of them in color

高等教育出版社
HIGHER EDUCATION PRESS

Springer

Author
Dr. Chaomei Chen
College of Information Science and Technology
Drexel University
3141 Chestnut Street, Philadelphia
PA 19104-2875, USA
E-mail: chaomei.chen@drexel.edu

ISBN 978-7-04-031703-9
Higher Education Press, Beijing

ISBN 978-3-642-19159-6 e-ISBN 978-3-642-19160-2
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011920985

© Higher Education Press, Beijing and Springer-Verlag Berlin Heidelberg 2011
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.
The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

Among the uniquely human capabilities is the capacity to create and discover. Understanding how humans create innovative art, music, poetry, or novels and discover scientific principles patterns, or relationships requires a recursive form of creativity and discovery.

The foundations for human creativity and discovery depend on passion for solving problems and fluency with social contexts that promote solutions. The passion produces persistence over time and enables devotion to solving important problems, filling troubling gaps, stretching annoying boundaries, or opening doors to fresh opportunities.

The fluency with social contexts helps researchers to see problems more clearly, bridge disciplines, and apply methods from one knowledge domain to another. The social context also provides powerful motivations that encourage varied forms of competition and collaboration. Sometimes competition is fierce, other times it can be friendly. Sometimes collaboration is narrow and limited to dialogs between trusted partners, other times it can be broad and long-term, producing lively conversations among thousands of contributors who are united by the passion to solve a problem. Innovators who protect their nascent ideas too closely will miss the opportunity to get feedback about their progress or learn about related ideas.

Researchers are increasingly attracted to study the dynamics of creativity and discovery. For the first time in history the databases of human scientific activity are sufficiently large and widely available. For the first time in history the tools for analyzing this data are capable of performing appropriate analyses and becoming widely available.

Retrospective citation analysis of scientific papers remains the major approach, sometimes complemented by informed ethnographic observations and interviews by researchers with sufficient knowledge-domain understanding to recognize important steps, controversies, or mistakes. However, analysis of patents, patent citations, trade journal articles, blogs, emails, twitter posts, and other social media will provide a finer-grained, more diverse, and more immediate record of how scientific breakthroughs emerge.

Citation analysis goes far beyond simple counts of who cited whom, but expands to author co-citation and document co-citation networks, while adding potent metrics such as betweenness centrality to find boundary-spanning papers that bridge knowledge domains. An important tool for these analyses is network visualization, which sometimes surprises researchers by showing important clusters, revealing bridging papers, or spotting important papers that may be tragically ignored for many years or become very hot quickly.

This latest book from Chaomei Chen makes important contributions to research on creativity because he brings a remarkably broad perspective to this topic, weaving together several strands of research. Chen clarifies existing theories, applies interesting metrics, and shows compelling visualizations. He lets readers know exactly what his point of view is: "transformative discoveries are likely to emerge from the twilight zones where multiple fields meet." This strong conviction is validated by retrospective analyses and case studies from impressively diverse branches of science.

The importance of this book, *Turning Points The Nature of Creativity*, is that Chen has a greater ambition than to look back, he wants to be in the moment by offering researchers the capacity to see what is currently happening in their knowledge domains, so as to spot important contributions early. The capacity to predict which papers will eventually be highly cited would be a wonderful gift to researchers, government policy planners, and industry managers. This goal is not easy to attain, but Chen suggest some promising possibilities.

The even more ambitious challenge that Chen takes on is to spot opportunities for interesting research by identifying "structural holes" or missing intersections of related knowledge domains. This is not easy since there are many unproductive intersections, so it takes informed expertise to make the right judgments or spot early signs of progress. This is a seductive idea, but Chen warns of many forms of "biases, pitfalls, and cognitive traps." Still he boldly offers a powerful claim: "a paper with a high betweenness centrality is potentially a transformative discovery. In addition, it would be possible to use this metric to identify potential future discoveries by calculating the would-be betweenness centrality of a hypothetical connection between two disparate areas of existing knowledge networks....Thus, betweenness centrality can be translated into interestingness, which can be in turn translated into actionability."

Readers should take time to reflect on the goals Chen lays out and appreciate the diverse sources he draws from. They should also carefully consider the metrics he proposes and study the visualizations from his CiteSpace system. Chen admirably lays out his emerging ideas, seeking constructive dialogs and engaging in fruitful conversations. This makes for provocative reading and stimulates fresh thinking. Readers can respond with even better theories, data, metrics, and visualization.

Ben Shneiderman
University of Maryland
July 2011

Preface

Research assessment has become a central issue for more and more government agencies and private organizations in making decisions and policies. New indicators of research excellence or predictors of impact are popping out one after another. However, if we look behind the available methods and beyond the horizon decorated by the various types of indicators, then we will encounter a few questions again and again: What is the nature of creativity in science? Is there a way that we can tell great ideas early on? Are there ways that can help us to choose the right paths? Can we make ourselves more creative?

There are only two types of theories no matter what their subjects are: the ones that are instructional and the ones that are not. An instructional theory will explain the underlying mechanisms of a phenomenon in such a way that we can see what we need to do to make a difference. The quest for us in this book is to look for a better understanding of mechanisms behind creativity, especially in the context of making and assessing scientific discoveries. In this book, my goal is to identify principles that appear to be necessary for creative thinking from a diverse range of sources and clarify where we may struggle with biases and pitfalls created by our own perceptual and cognitive systems. Then I will introduce an explanatory and computational theory of discovery and demonstrate its instructional nature through a series of increasingly refined quantitative approaches to the study of knowledge domains in science. Finally, the potential of transformative research is measured by metrics derived from the theoretical underpinning and validated with retrospective indicators of impact. The theory, for example, leads to a much simplified explanation of why some of the good predictors of citation counts of an article found by previous research are due to the same underlying mechanisms.

The conception of the theory of discovery was inspired by a series of intellectual landmarks across a diverse range of perspectives, notably, Vannevar Bush's *As We May Think* and his vision for trailblazing a space of knowledge in his Memex (memory and index), Thomas Kuhn's paradigm shift theory of scientific revolutions, Henry Small's methods for analyzing co-citation networks, Ronald Burt's structural-hole theory, and Peter Pirolli's optimal information foraging theory. The development and use of the CiteSpace system have played an instrumental role in experimenting and synthesizing these great ideas. I have been developing and maintaining CiteSpace since 2003. I have made it freely available for researchers and students to analyze emerging trends and turning points in the literature. The provision of CiteSpace has probably also promoted the awareness of scientometrics, the field that is concerned with quantitative approaches to the study of science. Feedback, questions, and requests for new features from a diverse and growing population of users have also propelled the search for theories to explain various patterns that we see in the literature.

The central thesis of the book is that there are generic mechanisms for creative thinking and problem solving. If we can better understand these mechanisms, then we will be able to incorporate them and further enhance them with computational techniques. Another important insight gained from reviewing the literature across different fields is that creativity is about the ability and willingness to find a new perspective so that we can see something that we take for granted.

The notion of an intellectual turning point has naturally emerged. Kuhn's gestalt switch between competing paradigms and Hegel's syntheses of theses and antitheses are exemplars of view-changing intellectual turning points. We may feel lucky or unlucky, depending on the particular perspective we take. We may miss the obvious if we are looking for something else. I hope that this book can provide the reader with some useful perspectives to study science and its role in society as well as insights into the nature of creativity so that we will be better able to recognize creative ideas and create opportunities for more creative ideas.

I have a few types of readers in mind when I was preparing for this book:

1) anyone who is curious about the nature of creativity and wondering if there is anything beyond the serendipitous view of creativity
2) analysts, evaluators, and policy makers in a situation where tough decisions have to be made that will influence the fate of creative work
3) researchers and students who need to not only keep abreast of their own fields of study but also position themselves strategically with a competitive edge
4) historians and philosophers of science

The first four chapters of the book should be accessible to college students and more advanced levels. The next four chapters may require a higher level of background information in areas such as network analysis and citation analysis. The book may be used for graduate-level courses or seminars in information science, research evaluation, and business management.

Chaomei Chen
Philadelphia, Pennsylvania
April 2011

Acknowledgements

Many people have played an information role in the ideas presented in this book.

My long-term collaborators in interdisciplinary research projects include Michael S. Vogeley, an astrophysicist at the Department of Physics, Drexel University, on a project funded by the National Science Foundation (NSF) (IIS-0612129) to study the interconnections between astronomical literature and the usage of the astronomical data obtained by the Sloan Digital Sky Survey (SDSS), Alan M. MacEachren, at the Department of Geography, Penn State University, on the Northeast Visual Analytic Center (NEVAC) project funded by the Department of Homeland Security, my graduate research assistants and doctoral students Jian Zhang and Don Pellegrino, and international visitors Fidelia Ibekwe-SanJuan (France) and Roberto Pinho (Brazil).

Eugene Garfield and Henry Small, visionary pioneers of citation analysis and co-citation analysis at Thomson Reuters, have been generous with their time and insights. Thomson Reuters' younger generation, David Liu (China), Weiping Yue (China), and Berenika Webster (Australia), are enthusiastic, energetic, and supportive. In particular, Thomson Reuters made generously arrangements for me to have an extensive period of access to the Web of Science while I was on sabbatical leave. I was a recipient of the 2002 Citation Research Award from the ISI and the American Society for Information Science and Technology.

I would like to thank Julia I. Lane and Mary L. Maher, Program Directors at the National Science Foundation (NSF), for their masterminded efforts in organizing the research portfolio evaluation project to explore technical feasibilities of evaluating NSF proposals (NSFDACS-10P1303), Jared Milbank and Bruce A. Lefker at Pfizer Global Research and Development at Groton Labs for collaborating on a Pfizer-funded drug discovery project.

I am also grateful to Zeyuan Liu at the WISELab, Dalian University of Technology, for his enthusiasm, vision, and insights in the use of CiteSpace in mapping knowledge domains in China, Hung Tseng, a biologist-turned NIH program director, for sharing his enthusiasm and insights in issues concerning the evaluation of research and tracing timelines of discoveries from a funding agency's point of view, Rod Miller, Drexel University, for numerous in-depth conversations on my current research and on articulating and communicating complex ideas effectively, and Ying Liu, the editor at the Higher Education Press, China, for her initiative and efforts in getting the book writing project underway.

To Baohuan, Calvin, and Steven, my caring, loving, and cheerful buddies in my sweet family, thank you for everything.

Contents