Aalto University publication series
DOCTORAL DISSERTATIONS /

Analysis of Cumulative and Temporal Patterns in Science

Pietro della Briotta Parolo

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Science, at a public examination held at the lecture hall TU1 of the school on 4 December 2017 at 12.

Aalto University
School of Science
Department of Computer Science

Supervising professor
Kimmo Kaski

Thesis advisors
Mikko Kivelä
Santo Fortunato

Preliminary examiners
Alexander Petersen,
UC Merced
School of Engineering 2
Suite 315 5200 N. Lake Road
Merced, CA 95343

Angel Sánchez, Prof
Departamento de Matemáticas and Institute UC3M-BS for Financial Big Data
Universidad Carlos III de Madrid,
Av. de la Universidad, 30, 28911 Leganés, Madrid, Spain

Aalto University publication series
DOCTORAL DISSERTATIONS /

ISBN (printed)
ISBN (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:

Unigrafia Oy
Helsinki

Finland

NORDIC ECOLABEL
441 697
Printed matter

Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Pietro della Briotta Parolo

Name of the doctoral dissertation
Analysis of Cumulative and Temporal Patterns in Science

Publisher School of Science

Unit Department of Computer Science

Series Aalto University publication series DOCTORAL DISSERTATIONS /

Field of research Science of Science

Language English

☐ **Monograph** ☒ **Article dissertation** ☐ **Essay dissertation**

Abstract

The goal of science has always been to investigate the world and its phenomena, by collecting data from all possible events that take place around us, breaking them down into their most simple elements and trying to come up with models able to explain and predict the outcome of these events. For centuries, the primary focus of science was mainly on natural events, but as the new technologies allowed to gather data from human interactions, it was natural for scientists to use this new information in order to apply the same logic to social systems, including science itself.

Since the late 19th century, when the first modern scientific journals were published, science has seen a constant rise in both its size and productivity, thanks to the standardization of research practices and the building of an international community that actively helps to push forward the limits of human knowledge. As science itself went from being a purely intellectual endeavor to a complex social, economical and political system, it is no surprise that a lot of attention has been dedicate in recent years to the study of the underlying mechanisms of science, aided by the explosion of means of communication that allow collaborations and exchange of information at instant speed across the globe, leaving behind digital traces that provide valuable data to study. The continuous exponential growth of science however, causes also difficulties in analyzing objectively the patterns and statistics that scientific data can reveal: for example a paper from the early 20th century would rarely get more than 100 citations, while now it is not uncommon for publications to pass the 10 thousand citation mark.

This thesis follows these attempts in trying to grasp how science works, by investigating the connections, i.e. citations, that exists between scientific publications and how these connections create structures and patterns. It shows that typical patterns in citation count and diffusion of information between fields is heavily influenced by the rate of growth of science, thus suggesting to use the number of publications as a better measure of time. It shows that there is a lag between breakthrough discoveries and the time when they are recognized, thus suggesting that we might be either running out of discoveries or rather having too much of them, in either case an extreme phenomenon. It shows that the community of publications which builds around an original successful paper has a typical life cycle, with an initial clustering, followed by an inevitable breaking down. Finally, it offers a new way of quantifying the impact of publications across time based on their cumulative impact on the overall corpus of scientific material.

Keywords science of science, scientometrics

ISBN (printed)	**ISBN (pdf)**	
ISSN-L 1799-4934	**ISSN (printed)** 1799-4934	**ISSN (pdf)** 1799-4942
Location of publisher Helsinki	**Location of printing** Helsinki	**Year**
Pages 100	**urn** http://urn.fi/URN:ISBN:	