Mechanisms of Scientific Life
Notes and Links

Doctoral Researchers of the
DFG Research Training Group "AlgoSyn"

December 2015

These notes are co-authored by the following members of the research training group AlgoSyn:

Wolfgang Thomas (Speaker of AlgoSyn)

Benedikt Brütsch
Marcus Gelderie
Florian Göbe
Oliver Göbel
Friedrich Gretz
Katrin Hölldobler
Sebastian Junges
Benjamin Kaminski
Gereon Kremer
Simon Leßenich
Ulrich Loup
Johanna Nellen
Christian Meirich
Mathias Obster
Deni Raco
Frederic Reinhard
Jan Ringert
Thomas Ströder
Andreas Tönnis
Melanie Winkler
Sarah Winter

Preface

Doctoral researchers have a rich program of activities, reaching well beyond the aim of producing scientific results and composing a valuable dissertation. In many cases, work in teaching is involved, or work in a scientific or industrial project. Furthermore, in many universities the doctoral researcher adopts the role of student and takes part in courses and tutorials on various subjects, among them training "soft skills" (or "transferable skills").

In the work of the DFG research training group "AlgoSyn" at RWTH Aachen ("Algorithmic Synthesis of Reactive and Discrete-Continuous Sytems", 2008-2015) it became clear that some issues are often overlooked in these study programs. Besides "skills" of a general nature, young scientists should have solid knowledge about the structures that govern scientific life. These "mechanisms of scientific life" deal with issues like the following:

- How does the scientific competition work, in conferences, journal articles, books?
- How does the market of scientific publishing work?
- How does one evaluate scientific work, how are scientists and institutions ranked?
- What are the basic principles of "scientific conduct" — e.g., in anonymous reviews, own writing, and in editorial work?

We felt that each scientist should know the basic facts and principles on these questions at the start of his/her career, independent of the actual course of career adopted after the doctorate. There are many aspects in this field where good articles, monographs, or web sites exist, for example on the topics closer to the training of skills, e.g., scientific writing, scientific presentations, ethics and plagiarism, and others. The present work book touches such questions, but – when appropriate – just gives pointers to relevant literature. On the other hand, we hope that the reader finds some interesting material here that is complementary to easily available sources.

The text is written from a point of view as represented by our research training group – so our background (and primary audience) is the German scenery of doctoral studies in computer science. Of course, many sections of the present work do not depend on this context and can be of interest for a wider audience.

Most articles were written by doctoral researchers of AlgoSyn, some parts were contributed by advisors. We offer this collection of articles and tutorials on our website. We have to admit that this work is (and will remain) in an

unfinished shape – the subject is too vast, and the set of doctoral researchers involved in the project is under permanent change. So the result achieved at the end of the funding period of AlgoSyn is not complete.

We learnt a lot by the discussions on how to shape this work and by the efforts of writing. We hope that you, the reader, will gain something while leafing through these pages.

Aachen, January 2016

Wolfgang Thomas

Speaker of AlgoSyn

Contents

Chapter 1

Reviewing Scientific Papers

1.1 A Word on Writing

When the first chapter of these notes starts with remarks on reviewing scientific work, the reader may think "before reviewing there is writing" – and ask for a section on the important aspect of how to write scientific work.

We are not entering this subject here since it is covered by dozens of books so that even to compile a list of references is a non-trivial task. Rather than entering this body of literature we single out two key references that are classics today and address the specific problems in the area of mathematical sciences (to which computer science belongs). The first reference is by a great master of mathematics and of mathematical exposition:

> P.R. Halmos, How to write mathematics, L'Enseignement Mathematique 16 (1970), 123-152

available on the website of the Swiss Consortium of University Libraries under

> `http://www.math.uh.edu/~tomforde/Books/Halmos-How-To-Write.pdf`

This paper, which makes for enjoyable reading, has been expanded into a book with three further co-authors:

> N.E. Steenrod, P.R. Halmos, M.M. Schiffer, J.R. Dieudonné: How to write mathematics, American Mathematical Society 1973

The second reference is by one of the pioneers of computer science, Don Knuth, together with two co-authors:

> D.E. Knuth, T.L. Larrabee, P.M. Roberts: Mathematical Writing, Mathematical Association of America, 1989

In the remainder of this chapter, we address aspects which become relevant when a scientific paper (or book chapter, or book) is written. It is read and checked by other members of the scientific community. This process is called reviewing.

1.2 Structure of the AlgoSyn Review Training

In the research training group AlgoSyn, reviewing was treated in two parts. In a first part, a tutorial with background information was given, with contents as presented in the subsequent subsections. In a second part, reviewing was practised by the doctoral researchers:

- The steps and roles implemented in the conference management system Easychair were explained and tested.
- Test articles were distributed, and reviews prepared by the doctoral researchers.
- The reviews were discussed in a session of all doctoral researchers, simulating a program committee session, and moderated by the speaker of the research training group. Here the emphasis was on the quality of the reviews rather than the quality of the test papers.

1.3 Significance of Reviewing

Scientific reviewing is a cornerstone of scientific life. It serves as the central mechanism of self-control of the scientific community to ensure true scientific progress. Manuscripts are checked by experts on the respective subject. These experts are usually not known to the author(s) – a fact which helps to obtain frank and honest statements on the manuscript under consideration.

The result of reviewing is twofold: It serves as a filter (the yes-no-decision whether a manuscript is accepted for publication in a scientific journal, at a conference, or another venue of publication), and it gives feedback to the author to improve the work (either in a new manuscript or in a revised version prepared for publication).

This rigorous procedure of quality control is not found in other types of literature (belletristic or journalistic works); and in fact it is the key mechanism to create trust in science. This refers to trust in science by the greater public (e.g., in political discussions), but also to trust among scientists who need to be certain about the scientific results they are using and relying on.

To participate in the (often strenuous) task of reviewing papers is a duty of every scientist; the willingness of researchers to take review jobs is a prerequisite of scientific progress.

There are many kinds of danger that might harm the process of quality control by reviewing. Let us list three of them:

1. Lack of time and other reasons for hasty or careless reviews: By the increasing number of papers produced world-wide, the workload of reviewing increases, and researchers may be confronted with two many requests for reviews, sometimes with strict deadlines (of e.g. of just a month, when a conference paper is to be reviewed).
2. Lack of independence, usually in forms of “conflict of interest”, i.e., a too close relation to the author(s), for example if the author has been a recent co-author of the reviewer.

3. Issues of "money". This is a delicate point which may harm the reputation of science in a most disastrous way. Examples are known from the research on climate change and medicine. Economic interests may interfere with the results of scientific work, and reviewing may be misused to either push papers that are a support for such interests, or to downscale or even stop work that has the potential to harm such interests. As another illustration, consider the field of e-education (for schools or universities). While projects in this field need scientific control, most of the possible reviewers are themselves involved in such projects and rely on the respective funds – this makes it hard to obtain objective opinions on the perspectives of such projects.

The last point already indicates that reviewing does not only cover the evaluation of scientific papers to be published in journals or conference proceedings. Reviews are also required for thesis work (e.g., doctoral dissertations) and for books, and also research projects are reviewed. Here the responsibility of a reviewer is especially high since usually the decision on large amounts of funding depends on the reviewer's statements. The highest degree of responsibility is reached when reviews have to be written on the quality of research of individual researchers (e.g. in procedures of appointment for professorship positions).

1.4 How Peer Reviewing Works

1.4.1 What is peer review?

Peer review is used by scientific publishing organs and conferences as part of the editorial decision process that leads to the acceptance or rejection of a scientific manuscript for publication. The peer review process usually consists of the following steps:

1. Scientists submit their manuscript with their research results to a publisher for consideration.
2. The editor asks other scientists who are working in the same research field as the authors (also called "peers") and are therefore expected to have the required expertise to be able to understand and evaluate the manuscript for a review.
3. Each reviewer (also called "referee") sends their review of the manuscript back to the editor. The review should consist of a general evaluation, feedback for the authors, comments and a recommendation for acceptance or rejection for the editor.
4. The editor provides the authors with the reviewer's feedback and might give them the chance to revise their manuscript and to resubmit it.
5. Eventually the manuscript gets accepted or rejected.

1.4.2 How to become a peer reviewer?

Typically an editor will ask people to serve as reviewers who have previously been involved with the journal or conference by successfully having published their own manuscripts.

1.4.3 Why become a peer reviewer?

Although reviewers don't get paid for their work, becoming a reviewer has other benefits. Serving as a reviewer is a valuable contribution to the scientific process in general and commonly something that is expected from researchers in academia. It is also an opportunity for a researcher to get informed about new scientific contributions to their field early on.

1.4.4 When to reject a review request?

There are situations in which it is better to reject a review request. Firstly, a review request should be rejected, if by self-assessment one realizes that the topic of the manuscript falls outside of one's scope of expertise. It is also better to reject a review request if meeting the deadline for the review cannot be guaranteed in advance. In the interest of impartiality and fairness a review request should also be rejected when one has any personal involvement with any of the authors that leads to a bias in favour or against them personally. Besides the aspect of personal connections, there is also the possibility that the paper under review is in direct competition with one's own work; in this case, a conflict of interest exists. The editor who recruits reviewers cannot always know whether potential conflicts of interest exist, so that the reviewers themselves bear most of the responsibility to guarantee a fair peer review process.

1.4.5 What types of peer review are there?

Single blind review The reviewers know the identity of the authors, but not vice versa. This is the most common type of peer review.

Double blind review The reviewers and the authors do not know each other's identity.

Open review Reviewers and authors know each other's identity.

1.4.6 Structure of the review

A review of a research paper usually consists of three kinds of statements

Rating This section contains the overall rating of the manuscript and the judgement whether the reviewer considers it suitable for publication.

Comments to the editor This section contains comments that will only be read by the editor. It will contain all confidential remarks.

Comments to the author This section contains comments that will be sent to the author of the manuscript. They give the author a summary and assessment as perceived by the reviewer, and it can offer comments that should help the author(s) to improve their manuscript, e.g. by pointing out flaws and possible repairs.

1.4.7 Evaluation Critera

The reviewer should evaluate the manuscript according to certain evaluation criteria, of which the most common are the following:

Correctness Are all the results and claims that are contained in the manuscript correct and do the authors provide sufficient evidence and proof for their claims?

Significance Does the manuscript contain research results that are novel and significant enough to warrant publication?

Readability Is the manuscript readable enough in terms of style, notation and grammar to warrant publication?

1.5 Practical hints for writing reviews

1. Before entering the job, ask yourself whether you are able to judge the paper under consideration, what your level of "confidence" is, and whether there is a conflict of interest.
2. Do not misuse the fact that your identity remains secret – a good principle is to write a review in a way that you would not mind having it made public under your name.
3. Before entering evaluation, describe the contents of the work in your own words, hinting at the novelty, the technical substance, and the relation to existing literature.
4. Give your evaluation in two parts, on the correctness, originality, and technical merit, and on the clarity of exposition as well as grammatical and terminological correctness.
5. Most important in the case of critique: Do not insult the author but stay objective and constructive.

Especially the last point is vital. Insulting words are usually not tolerable even in personal dispute, even more so insults have to be avoided when statements are made from behind a wall (of anonymity).

The last point is also a vital principle for reading reviews. One should always take critique which might be stated in a review as a useful feedback, enabling an improvement of the submitted work.

1.6 References and Links

Peer review guidelines of publishers

Elsevier Reviewer Guidelines:
`http://www.elsevier.com/reviewers/reviewer-guideline`

Springer Peer Review Academy:
`http://www.springer.com/authors/journal+authors/peer-review-academy`

Peer review guidelines of scientific organisations

Deutsche Forschungsgemeinschaft (DFG, in German):
`http://www.dfg.de/foerderung/grundlagen_rahmenbedingungen/quo_vadis_antrag/index.html`

Other peer review online tutorials

A Guide for New Referees in Theoretical Computer Science by Ian Parberry:
`http://jmlr.org/reviewing-papers/p92-parberry.pdf`

Kapitel 2

Das wissenschaftliche Publikationswesen im Umbruch – Beobachtungen eines Informatikers[1]

Wolfgang Thomas

2.1 Prolog

Wir Wissenschaftler glauben zu wissen, wie das Publikations- und Verlagswesen funktioniert. Es ist ja seit Jahrzehnten, seit Jahrhunderten die bewährte Schnittstelle zwischen den einzelnen Wissenschaftlern und der Gemeinschaft der Fachleute (heute sagt man "Research Community") des jeweiligen Forschungsgebiets – zum Teil auch zwischen den Wissenschaftlern und der weiter gefassten Öffentlichkeit. Mit dem Internet hat sich das Szenarium für wissenschaftliches Publizieren, wie jeder weiß, radikal verändert. Doch sind damit andere Veränderungen einhergegangen, die zu paradoxen Resultaten geführt haben, zum Beispiel zu einer Erschwernis des wissenschaftlichen Austauschs. Diese Trends sind es wert, genauer betrachtet zu werden, auch weil wir, die Wissenschaftler, nicht nur Beobachter, sondern Mitspieler in diesem Szenarium sind – und es selbst in der Hand haben, Verbesserungen zu erreichen.

Es sei betont, dass sich die nachfolgenden Überlegungen nur auf die wissenschaftlichen Publikationen in Zeitschriften und Konferenzbänden beziehen – auf dem Hintergrund einiger Jahrzehnte eigener Erfahrung als Autor und Herausgeber. Zudem geht es nur um die Gebiete Informatik und Mathematik; es ist klar, dass sich die Verhältnisse in anderen Fächern anders darstellen.

Danksagung: Viele Entwicklungen im Bibliothekswesen sind mir durch meine Frau Renate Eschenbach-Thomas, die Leiterin der Fachbibliothek Informatik an der RWTH Aachen, nahegebracht und erklärt worden. Ohne ihre zahlreichen Hinweise auf Quellen (u.a. Diskussionen im Internet) wäre die vorliegende Arbeit nicht möglich gewesen. Ihr sei an dieser Stelle ganz herzlich Dank gesagt. Auch

[1] Der vorliegende Text hat seinen Ursprung in Vorträgen des Autors zum Projekt *Mechanisms of Scientific Life* bei Workshops des DFG-Graduiertenkollegs "AlgoSyn" (2010-2015).

möchte ich Marc Herbstritt (Leibniz-Zentrum für Informatik Schloss Dagstuhl) und Prof. Klaus Tochtermann (ZBW - Leibniz-Informationszentrum Wirtschaft, Kiel) herzlich danken für ihre wertvollen Hinweise.

2.2 Ein Rückblick auf die "alte Zeit"

Wir beginnen mit einem Ereignis des Jahres 1923 – eine Zeit, als Göttingen weltweit als Mekka der Mathematik galt und als die deutsche Szene der Mathematik und Physik durch Namen wie Hilbert, Courant, Born, Einstein, Sommerfeld glänzte. Ausweis der führenden Rolle der deutschen Wissenschaftler in diesen Gebieten war die Reputation der Zeitschriften des Springer-Verlags, *Mathematische Annalen* und *Mathematische Zeitschrift*, und auch der berühmten "gelben Reihe", einer Folge von Monographien unter dem Titel *Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen*, die für Jahrzehnte der Spitzenort mathematischer Literatur war.

Die herausragenden Vertreter der Mathematik und Physik in Deutschland dokumentierten ihre hohe Wertschätzung für "ihren Verleger", Dr. Ferdinand Springer, mit einer bemerkenswerten Glückwunschadresse zum Erscheinen der 25. Nummer der Mathematischen Zeitschrift. Wir lesen[2]:

> Sehr verehrter Herr Dr. Springer!
>
> Die unterzeichneten Redaktionsmitglieder der Mathematischen Annalen und Mathematischen Zeitschrift möchten den Tag, an welchem der fünfundzwanzigste in Ihrem Verlage erscheinende Band mathematischer Zeitschriften herausgegeben wird, nicht vorübergehen lassen, ohne Ihnen gegenüber zum Ausdruck zu bringen, wie sehr sie von der Wichtigkeit Ihres Eintretens für unsere Wissenschaft durchdrungen sind und wie hoch sie dieses Eintreten einzuschätzen wissen. Wenn nicht Ihre opferbereite Unternehmungslust und Ihre umsichtige Energie sich der Sache der mathematischen Wissenschaft angenommen hätten, so würde heute die Mathematik in Deutschland nicht mehr lebensfähig sein. Dass sie es noch ist, und dass die deutschen mathematischen Journale nach wie vor zu den angesehensten der Welt gehören, haben die deutschen Mathematiker und damit die Wissenschaft überhaupt wesentlich Ihnen zu verdanken.
>
> Nehmen Sie an diesem Tage mit unseren Glückwünschen den Ausdruck der Hoffnung entgegen, dass der beschrittene Weg unter gegenseitigem Vertrauen und Verständnis unbeirrt weiter gegangen wird.
>
> Berlin und Göttingen, den 27.7.23
>
> *Klein D. Hilbert. C. Caratheodory R. Courant*
> *H. Bohr M. Born Bieberbach Knopp A. Einstein*
> *Erhard Schmidt I. Schur O. Hölder G. Herglotz L. Lichtenstein*
> *C. Neuman F. Schur A. Kneser E. Hecke W. v. Dyck*
> *O. Perron E. Landau W. Blaschke A. Sommerfeld*

[2] zitiert nach V.R. Remmert, U. Schneider: "Ich bin wirklich glücklich zu preisen, einen solchen Verleger-Freund zu besitzen" – Aspekte mathematischen Publizierens im Kaiserreich und in der Weimarer Republik, Mitteilungen der DMV 14-4 (2006), 196-205.

Wenig später, im Jahre 1930, verlieh die Universität Göttingen auf Betreiben der Göttinger Mathematiker (vertreten durch Courant, Hilbert, Landau und Herglotz) Ferdinand Springer sogar die Ehrendoktorwürde.

Selten hat sich so schön wie in diesen Ehrungen die Wertschätzung der Wissenschaft, repräsentiert durch ihre angesehensten Forscher, für die Leistung von Verlegern gezeigt. Die oben zitierten anrührenden Zeilen künden von einer funktionierenden Symbiose, die durch wechselseitiges Verständnis, ja wechselseitige Hochachtung geprägt ist. Was einen Verleger wie Ferdinand Springer kennzeichnete, war – neben einer Beherrschung des operationalen Geschäfts – der Blick für Qualität und die tiefe Verbundenheit mit den Anliegen der Wissenschaftler. Diese Verbundenheit hat dazu geführt, dass der Springer-Verlag die besten Fachvertreter für sich gewinnen konnte; und dies war letztlich die Grundlage für den Erfolg des Verlags, der sich auch in den Nachkriegsjahren fortsetzte.

Wir leben heute in anderen Zeiten. Eine Grußadresse wie die von 1923 ist heute kaum vorstellbar; stattdessen gibt es in der Community der Wissenschaft Klagen über Klagen über die Lage des wissenschaftlichen Publikationswesens im Allgemeinen und die Rolle der Verlagshäuser im Besonderen. Man muss eine tiefe Entfremdung zwischen Wissenschaftlern und (Groß-) Verlagen konstatieren. Zwar haben Verlage immer noch ihre traditionellen Funktionen, nämlich

1. die Entscheidung über die Publikation eines Werks (Artikel, Buch) mit Hilfe von hierzu herangezogenen Wissenschaftlern (den Herausgebern einer Zeitschrift, einer Buchreihe, eines Sammelbandes),

2. die Produktion von Druckwerken (Setzen, Drucken von Zeitschriften und Büchern, Binden von Büchern),

3. Werbung und Vertrieb,

4. Vergütung der Autoren gegen Übergabe des Copyright.

Doch haben sich beispielsweise die Funktionen unter 2 und 3 auf dem Hintergrund der Digitalisierung von Literatur dramatisch verschoben. In den folgenden Abschnitten diskutieren wir diese und andere damit verbundene Entwicklungen, welche die heutige Situation prägen, und gehen der Frage nach, welche Perspektiven sich daraus ergeben.

2.3 Die Digitalisierung: TeX, LaTeX und das Internet

Ein besonderes Merkmal der mathematischen (und auch naturwissenschaftlichen) Literatur ist das Auftreten mathematischer Formeln, seien es nun Integrale, Matrizen oder logische Ausdrücke. Diese oft komplexen Textbestandteile mussten mit spezifischer Expertise gesetzt werden. Früher war das eine der zentralen Dienstleistungen der Verlage. In den 1970er Jahren kam der “Fotosatz” auf, bei dem die Autoren den Text in einer Form ablieferten, der nur noch Seite für Seite fotografiert werden musste. Die Autoren, nicht mit der Setz-Technologie ausgestattet, füllten dann Sonderzeichen per Hand ein, und es gab auch spezielle Vorrichtungen für klassische Schreibmaschinen, mathematische Symbole wie

griechische Buchstaben zu tippen: Man setzte eine mit Feder gehaltene Type (etwa "alpha") an der Einschlagstelle der Typenhebel ein, tippte irgendeinen Buchstaben (etwa "e"), und der e-Typenhebel schlug dann den alpha-Typenhebel aufs Farbband und dieses auf das Papier. Später konnte man auch die Kugelkopfmaschinen verwenden, was allerdings das Wechseln von Kugelköpfen erforderte, sobald man verschiedene Buchstabensätze nutzte.

Alle diese Umstände wurden durch die Pioniertat eines einzigen Wissenschaftlers (und seiner Mitarbeiter) überwunden: Donald Knuth, einer der führenden Köpfe in der Herausbildung des Fachs "Informatik". Das Mammutwerk *The Art of Computer Programming* (erschienen bis Band 4) war (und ist immer noch) eine Referenz für Generationen von Informatikern. Knuth entschied sich jedoch, für etwa ein Jahrzehnt – ca. 1977-1986 – ein "praktisches" Projekt zu verfolgen, nämlich die Konzeption und Entwicklung von Software, die den Vorgang des Setzens von Text und Formeln automatisiert. Dieses Vorhaben erforderte Problemlösungen auf zwei Stufen: Erzeugung der Einzelbuchstaben und Symbole, und Anordnung der Buchstaben zwecks Bildung eines ansprechenden Schriftbilds.

Für das erste Problem entwickelte Knuth die Programmiersprache *Metafont*. Mit ihr konnte man die graphische Gestalt aller benötigten Zeichen programmieren, seien es nun Integralzeichen, mathematische Zeichen oder Buchstaben (lateinisch, griechisch, hebräisch, Fraktur, etc.). Dies richtig zu tun, erforderte Expertise aus Mathematik, Buchdruckerkunst und schlicht Ästhetik.

Für das zweite Problem, die richtige Anordnung von Zeichen auf dem Papier, entwickelte Knuth ein weiteres Programmiersystem: TeX (wobei das "X" für das griechische "Chi" steht). Als Programmiersprache steuert TeX geometrische Größen wie die Abstände zwischen Buchstaben, Wörtern, Zeilen, Absätzen, und sie gliedert den Text in inhaltliche Blöcke. So wird vermieden, dass beispielsweise ein Seitenumbruch eine einzelne Zeile eines Absatzes erzeugt (auf der Erst- oder der Zweitseite) – was die Setzer "Schusterjunge" bzw. "Hurenkind" nannten. Ebenso wird vermieden, dass ein Seitenumbruch eine Überschrift vom Folgetext trennt. Hervorgehobene mathematische Formeln werden korrekt zentriert, mit angemessenem Abstand zum Vor- und zum Folgetext.

In TeX schreibt man also Text und Formeln gemeinsam mit "Setzanweisungen", etwa der Information, dass es sich bei einem Wort um eine Überschrift handelt. Das TeX-System ist dann in der Lage, das Druckbild (unter Hilfe von *Metafont*) automatisch zu erzeugen; es rechnet dazu aus, wie eine Seite am besten ausgefüllt wird, zum Beispiel durch nicht merkliche Veränderung der Zeilenabstände, um etwa die Überschriften korrekt platzieren zu können. TeX übernimmt also vollständig die Aufgabe des Setzers, mit dem Unterschied, dass das Ergebnis nicht wie in alter Zeit eine Anordnung von Metalltypen ist, sondern die Konfiguration der Schwarz-Weiß-Bereiche durch Angabe, welche Symbole wo gesetzt sein sollen. TeX wurde durch Leslie Lamport, ebenfalls ein Grundlagenforscher, der wie Knuth den Turing-Preis (den "Nobelpreis für Informatik") erhielt, zum System "LaTeX" erweitert. Hiermit konnte – anders als bei TeX – der "durchschnittliche Wissenschaftler" bequem umgehen. Dazu wurden leicht lernbare Standardbefehle definiert, mit denen die TeX-Programmierung vereinfacht wird. Heute schreibt praktisch jeder Forscher im Bereich Mathematik und Informatik seine Arbeiten in LaTeX.

TeX und LaTeX haben die Erstellung von mathematischer Literatur revolutioniert. Es sind so ausgefeilte Systeme, dass man keinen Qualitätsverlust gegenüber den Ergebnissen des klassischen Buchdrucks feststellen kann (jedenfalls soweit

man wissenschaftliche Literatur betrachtet).

Wenn ein LaTeX Dokument erstellt ist, kann es automatisch in die Form einer "pdf-Datei" übersetzt werden. Das "Portable Document Format" (pdf) der Firma Adobe ist heute ein Weltstandard für die Darstellung von Dokumenten.

TeX, LaTeX und das Dokumentenformat pdf ermöglichen den Verlagen enorme Kostensenkungen; die Verlage verlangen in der Regel die jeweilige Literatur im pdf-Format und können sich mehr als früher auf den Vertrieb konzentrieren.

Doch auch diese Funktion der Verlage hat sich durch die globale Infrastruktur, die das Internet bereitstellt, gewandelt. Zwar werden noch Bücher gedruckt und an Buchhändler und Bibliotheken versandt, doch erfolgt der Wissensaustausch mit Zeitschriftenartikeln heute zumeist auf digitaler Basis. Das "World-Wide Web" bietet dafür die technischen Voraussetzungen: Auf den Servern der ganzen Welt liegen wissenschaftliche Artikel im pdf-Format vor, zugänglich gemacht durch sogenannte "Links", also Verweise. Ein Autor kann zum Beispiel auf seiner persönlichen Webseite seine Publikationsliste veröffentlichen und dafür sorgen, dass beim Anklicken eines dieser Titel die Druckausgabe (gespeichert im pdf-Format) erscheint. Diese Technologie erlaubt es also im Prinzip, auch die Funktion des Vertriebs von Literatur aus den Händen der Verlage zu lösen.

Eine völlig neue Funktion wird den Wissenschaftlern außerdem dadurch eröffnet, dass Suchmaschinen zur Verfügung stehen, die es erlauben, im Internet gezielt nach relevanter Literatur zu suchen.

Diese Entwicklungen, die sich bisher über den kurzen Zeitraum von vielleicht 20 Jahren erstrecken, haben die Arbeit der Wissenschaft in einer Weise verändert, wie es zuvor nur mit dem Aufkommen des Buchdrucks – in einer viel längen Zeitspanne – geschehen ist.

Zugleich stellt man fest, dass die meisten Funktionen, die bisher durch Verlage wahrgenommen wurden, nun auch – im Prinzip – ohne deren klassische Dienstleistungen erfüllt werden können. Viele Wissenschaftler haben daher erwartet, dass die Literaturbeschaffung nicht nur drastisch vereinfacht würde, sondern dass auch viel geringere Kosten anfallen würden. Es ist anders gekommen.

2.4 Die Kommerzialisierung

Parallel zur technologischen Revolution, die sich mit der Digitalisierung der wissenschaftlichen Literatur vollzog, hat sich auch das wissenschaftliche Verlagswesen geändert. Die Chancen, die sich mit der Globalisierung des Wissenschaftsbetriebs ergaben, wurden von einigen Verlagen schnell erkannt und zielstrebig wahrgenommen – schneller als dies die Wissenschaftler selbst taten. Das Publikationswesen wandelte sich zum "Big Business" und wurde in einer Weise kommerzialisiert wie dies der Verleger Friedrich Springer im Jahr 1923 wohl nicht für möglich gehalten hätte. Wir illustrieren diese Kommerzialisierung durch einige Daten der Verlagsgeschichte des Hauses Springer:

1842: Gründung durch Julius Springer (1817-1877), fortgeführt durch Ferdinand Springer sen. (1846-1906) und Ferdinand Springer jun. (1881-1965)

1999: Bertelsmann übernimmt Springer von der Familie Springer (Vorstandsvorsitzender bei Bertelsmann damals Thomas Middelhoff), seitdem massiver Ausbau des digitalen Angebots, Rückbau der Rolle der Printmedien.

2003: Die Beteiligungsgesellschaften Candover und Cinven kaufen Springer. Fusion mit Kluwer zu "Springer Science and Business Media".

2009: Die Beteiligungsgesellschaften EQT und GIC kaufen Springer für 2,3 Milliarden Euro.

2013: Die Beteiligungsgesellschaft BC Partners kauft Springer für 3,3 Milliarden Euro.

2015: Holtzbrinck und Springer fusionieren (und Holtzbrinck übernimmt die Mehrheit an Springer von BC Partners), Unternehmenswert jetzt 5 Milliarden, Börsengang möglich. Fusion mit MacMillan Education zur "Springer Nature Group".

Die Entwicklung seit 2003 zeigt zunächst, dass aus dem Dreieck Autor – Verlag – Leserschaft nun ein Viereck geworden ist: Autor – Verlagspersonal – Verlagseigentümer – Leserschaft. Man muss auf der Verlagsseite unterscheiden zwischen dem Personal, welches das Verlagsgeschäft betreibt (und zwar in einer langfristigen Bindung an den Verlag, mit hochkompetenten Mitarbeitern), und den Verlagseigentümern, die keine professionelle Bindung zum Verlagsgeschäft mehr haben. Die Eigentümer sind (anders als seinerzeit der Verleger Friedrich Springer) üblicherweise an zahlreichen Firmen beteiligt und haben keine primär verlegerischen Interessen mehr; die rein wirtschaftlichen Interessen in diesem Kontext der Hedgefonds-Kultur haben zur Folge, dass die Eigentümerschaft in so kurzen Fristen wie Fünfjahreszeiträumen wechseln kann. Die Dimension dieses Geschäfts wird klar an der Wertsteigerung des Unternehmens Springer um eine Milliarde in nur vier Jahren (2009-2013).

Diese Kommerzialisierung geht einher mit einer rapiden Konzentration des Verlagsmarktes auf wenige global wirtschaftende Unternehmen. So gehören die traditionsreichen Verlage Vieweg und Teubner heute zum Springer-Konzern. Die Verlage Academic Press, Morgan Kaufmann Publishers, Pergamon Press und andere gehören heute zu Elsevier, einem Verlag der Aktiengesellschaft Reed Elsevier (seit 2015 "RELX Gruppe").

Es wäre eine interessante Aufgabe für Wirtschaftshistoriker, diese Umwälzung im Verlagswesen im Einzelnen zu studieren und vielleicht zu erklären. Ein wesentlicher Treiber dieser Umwälzung ist zweifellos das Internet, das dazu aufruft, globale Geschäftsmodelle zu entwickeln. Dies haben Großverlage wie Elsevier, Springer, Wiley mit beeindruckender Zielstrebigkeit getan.

Ein entscheidender Schritt war der Aufbau einer digitalen Infrastruktur, nicht nur für den inneren Betrieb (etwa zur Unterstützung der Begutachtung von Artikeln durch die Herausgebergremien) und für die Vernetzung von Literatur (was den Forschern bei ihrer Arbeit hilft), sondern vor allem für den Vertrieb. Die Auslieferung von Printmedien wurde nach und nach zurückgefahren; stattdessen wurde (für zunächst ermäßigte Preise, später immer teurer) der Zugriff zu der auf den Verlagsservern liegenden Literatur verkauft. Die Universitätsbibliotheken wurden auf diese Weise allmählich immer härter in die Zange genommen: Deren Verpflichtung, Zeitschriften zu abonnieren, wurde ausgenutzt, und so kam es seit den 1990'er Jahren zu Preissteigerungen von 10 % und mehr pro Jahr. Zur "Peitsche" kam das "Zuckerbrot" hinzu durch Angebote, erträgliche Preise pro Zeitschrift dadurch zu erreichen, dass man Zeitschriften im Bündel offerierte (was die Bibliotheken nötigte, auch ungewünschte Titel zusätzlich zu abonnieren).

Die Umstellung von Printmedien auf digitale Medien hatte einen desaströsen Effekt auf den Status der Universitätsbibliotheken: Sie haben sich von Eigentümern (der Literatur auf den eigenen Regalen) zu Mietern gewandelt: Sie besitzen nicht mehr die Literatur selbst, sondern erwerben (auf Zeit) nur noch die Zugriffsrechte dazu. Damit nähert sich der Status der Bibliotheken dem der IT-Zentren in den Universitäten an. Und so ist die Frage einer zeitgemäßen Katalogisierung in den Vordergrund gerückt, da hier noch spezifische bibliothekarische Kompetenz gefragt ist. Wie weit das trägt, ist offen. Die klassische Aufgabe der Universitätsbibliotheken, die wissenschaftliche Literatur für zukünftige Generationen zu archivieren und zugänglich zu halten, ist auf ein Minimum geschrumpft. Die Besitzer der Literatur sind andere. Die Dynamik in den Besitzverhältnissen bei den Verlagen vermittelt zudem – anders als es bei den Universitätsbibliotheken der Fall wäre – nicht das beste Gefühl für eine Beständigkeit, wie man sie mit einer wissenschaftlichen Archivierung verbindet.

Ein anderer Druck auf die Wissenschaft wurde parallel erzeugt durch den Aufbau neuer Infrastrukturen der Wissenschaftsevaluation. Das Internet erlaubt es, mit Hilfe der verfügbaren Suchmaschinen und darauf aufbauender Softwaresysteme, die Artikel und Bücher eines Autors (oder der Autoren einer Universität) ziemlich vollständig zu erfassen und – mehr noch – die Zitationsstrukturen offenzulegen. So hat sich etwa der "Hirsch-Index" als Maßstab für die "Qualität" eines Wissenschaftlers etablieren können. Er beruht auf der Integration zweier Parameter: Anzahl der veröffentlichten Arbeiten, Anzahl der Zitate einer Arbeit in anderen Werken. Hat jemand 20 Arbeiten geschrieben, die jeweils in 20 anderen Arbeiten (die eigenen ausgenommen) zitiert wurden, so ist der Hirsch-Index mindestens 20; allgemein ist es die größte Zahl x, so dass x Arbeiten mindestens x Fremdzitate haben. (Es hilft also wenig, nur eine Arbeit zu schreiben, die alle zitieren, oder 100 Arbeiten, die niemand zitiert.)

In diesem feinen Räderwerk der Evaluation arbeiten heute die Autoren (und – in Ebenen darüber – die Departments, Rektorate und Hochschulräte): Sie müssen im globalen Wettbewerb darauf hinarbeiten, dass ihre Kennzahlen zur Spitzengruppe gehören. Die Bereitstellung des umfangreichen statistischen Materials und die Erstellung der vielen "Rankings", die daraus erwachsen, haben ebenfalls Wirtschaftsunternehmen übernommen (allen voran Thomson Reuters, gegen gutes Geld, das die Universitäten zu zahlen haben), ehe die akademische Welt oder deren Geldgeber, d.h. staatliche Institutionen, auf die Idee kamen, das selbst in die Hand zu nehmen. Das "Eichamt" der Wissenschaft ist derzeit in den Händen privatwirtschaftlicher Unternehmen (wie Scopus von Elsevier oder das ISI Web of Science von Thomson Reuters), und diese verfahren nicht immer transparent oder auch nur nachvollziehbar (hierzu unten mehr).

Da entschlossene Unternehmer die Chancen der Digitalisierung und des globalen Marktes viel dynamischer verfolgt und genutzt haben als die akademische Welt, haben die Gesetzmäßigkeiten des Kapitalismus dazu geführt, dass die großen Wissenschaftsverlage heute einen wirtschaftlichen Erfolg haben, der seinesgleichen sucht. Er beruht auf einem wirkmächtigen Transfer öffentlicher Gelder in private Kassen. Der Staat bezahlt für wissenschaftliche Leistungen mehrfach, nämlich für das Gehalt der Autoren in den Universitäten, für deren Tätigkeit als Gutachter und Herausgeber, für den (Rück-) Erwerb der in Artikeln dokumentierten Resultate durch die Universitätsbibliotheken und schließlich für die Wertung der von den Wissenschaftlern erbrachten Leistungen.

2.5 Anfänge einer "Reconquista" durch die Academia

Die Aufwendungen für wissenschaftliche Literatur haben sich allmählich so stark erhöht, dass die Arbeit der Wissenschaftler dadurch merklich eingeengt wird. Seit etwa 10 Jahren ist diese Entwicklung auch ein Thema der breiten Öffentlichkeit. Ein prominentes Beispiel hierfür ist der Artikel von G. Monbiot "The Lairds of Learning" (The Guardian, 30. August 2011), übersetzt erschienen in den *Mitteilungen der Deutschen Mathematiker-Vereinigung* unter dem Titel "Die Lehnsherren des Lernens"[3]. Diese Philippika beginnt so:

> Wer sind die rücksichtslosesten Kapitalisten der westlichen Welt? Wessen monopolistische Praktiken lassen WalMart wie einen Tante-Emma-Laden und Rupert Murdoch wie einen Sozialisten aussehen? Du wirst nie im Leben darauf kommen. Es gibt viele Anwärter, aber ich votiere nicht für die Banken oder die Ölkonzerne und auch nicht für die Krankenkassen, sondern – und jetzt halte dich fest – für die Wissenschaftsverlage [...]

Kann man dies durch ein Beispiel belegen? Ich erwähne hier nur meine erste wissenschaftliche Arbeit, das Kondensat meiner Masterarbeit an der University of Bristol. Unter dem Titel "A note on undecidable extensions of successor arithmetic" sind diese zwei Seiten erschienen im *Archiv für Mathematische Logik und Grundlagenforschung*, Band 17 (1975), S. 43-44. Die Zeitschrift gehörte damals zum Verlag W. Kohlhammer, heute ist sie im Springer-Verlag. Ist die Zeitschrift nicht in einem Universitätsabonnement erfasst, betragen die Kosten für eine digitale Kopie meines Artikels 39,95 USD oder 34,95 Euro. Das ist schlicht unbegreiflich. Ich muss natürlich der Meinung sein, dass bereits meine erste Arbeit sehr gut war, doch hege ich ernste Zweifel, ob sie (zumal nach 40 Jahren) *so* wertvoll ist.

Die Preispolitik der Großverlage, die ganz offenbar durch die Interessen der Verlagseigentümer bestimmt ist, hat in den vergangenen 20 Jahren zu einer so extremen Verteuerung der Zeitschriften und der Konferenzbände geführt, dass Universitätsbibliotheken Zeitschriften im großen Maßstab abbestellen mussten. Damit verkehrte sich der Sinn des Publizierens ins Gegenteil: Ein Autor, der möchte, dass seine Arbeiten gelesen werden, erreicht durch die Übergabe seines Artikels an einen Verlag nicht mehr unbedingt, dass die weltweite Verbreitung gesichert ist, sondern dass große (und eventuell auch wachsende) Personenkreise als Leser ausgeschlossen werden.

Eine prominente Rolle spielt in diesen Auseinandersetzungen seit Jahrzehnten das Verlagshaus Elsevier, dessen Preispolitik den Universitätsbibliotheken immer wieder Kopfschmerzen bereitet.

In den U.S.A. regte sich Widerstand bereits vor gut 10 Jahren. Vorreiter war die Cornell University, deren Senat 2004 beschloss, die Abonnements von Elsevier-Zeitschriften zu kündigen. Über die Jahre vollzogen auch weitere Universitäten diesen Schritt, zuletzt im Februar 2015 die Universität Leipzig. In einem Interview, das die *Frankfurter Allgemeine Zeitung* daraufhin mit dem Leiter der UB Leipzig führte (11.2.2015), findet sich auf die Frage *"Es heißt, eine einzelne Zeitschrift kostet bei Elsevier bis zu 21000 Euro. Wie kommt es zu*

[3] Mitteilungen der DMV 19 (2011), 148-152.

diesen hohen Summen?" die Antwort: *"Gier. Es gibt keine andere Erklärung. Die Gewinnmargen von Elsevier sind bekannt. Sie lagen in den vergangenen Jahren bei dreißig Prozent. Das ist Geldschneiderei."* Gegenwärtig (Dezember 2015) verfolgt die akademische Öffentlichkeit die Auseinandersetzung in den Niederlanden (wo Elsevier in Amsterdam auch residiert). Der Verbund der niederländischen Universitäten erwägt, alle Abonnements bei Elsevier zu kündigen.

Diese eher defensiven Aktionen werden – nach bescheidenen Anfängen – mehr und mehr ergänzt durch Initiativen aus dem Wissenschaftsbereich, alternative Publikationsmodelle zu schaffen.

Als erstes Beispiel sei eine weitere Pioniertat der Cornell University erwähnt, nämlich die Beheimatung des Dokumentenservers arXiv (seit 1999). Hier können Wissenschaftler ihre Resultate kostenlos publizieren; die Artikel werden damit sofort (aber ohne Prüfung durch einen Begutachtungsprozess) weltweit zugänglich. In den Bereichen Physik, Astronomie, Mathematik, Informatik wird der wissenschaftliche Austausch durch arXiv sehr beschleunigt. Wichtige Resultate finden heute ihren Weg in die akademische Öffentlichkeit durch Platzierung auf dem arXiv Server. Die Kosten des Betriebs von arXiv sind erheblich (es sind ca. 900 000 USD jährlich), wenn auch minimal verglichen mit den Budgets der kommerziellen Verlage; sie werden getragen von der Cornell University Library, der Simons Foundation und unterstützenden Institutionen (wie der Max-Planck-Gesellschaft). Im Januar 2014 halfen 174 Institutionen mit Jahresbeiträgen von 300 bis 4000 USD bei der Finanzierung.

Ein zweites Beispiel betrifft die prominente Elsevier-Zeitschrift *Journal of Algorithms*. Nach Jahren immer schnellerer Preissteigerung für das Abonnement reagierte einer der Herausgeber, der schon erwähnte Donald Knuth, mit einem denkwürdigen Brief an seine Kollegen des *Editorial Board*. Dieses 14-seitige Papier vom 27. Oktober 2003 war bald auch in der Wissenschafts-Community über das Internet bekannt. Es mündete in der Aufforderung an das *Editorial Board*, geschlossen zurückzutreten und eine neue Zeitschrift in alternativer institutioneller Anbindung zu schaffen. Das geschah, und so wurde bei der amerikanischen Vereinigung für Informatik, der *Association of Computing Machinery (ACM)* die Zeitschrift *ACM Transactions on Algorithms* ins Leben gerufen. Elsevier musste das *Journal of Algorithms* einstellen. (Die Berufsvereinigung der Informatiker in den U.S.A., die zugleich als Verlag agiert, war hier der Rettungsanker. In Europa sind die Verhältnisse – noch – anders. Zum Beispiel hat sich die deutsche *Gesellschaft für Informatik (GI)* mit ihrer Hauszeitschrift und weiteren Projekten an den Springer-Verlag gebunden und blendet damit einen Teil ihres Mandats zur Interessenwahrnehmung der Wissenschaft aus. Das *EMS Publishing House* der *European Mathematical Society* ist ein Alternativmodell.)

Die jüngeren Forscher haben nicht die Muße und das Gewicht, in solche Auseinandersetzungen durch eigene Initiativen einzugreifen; sie können eigentlich nur durch Entscheidung, wo sie ihre Artikel publizieren, Einfluss nehmen. Umso mehr Bedeutung kommt der Haltung von Spitzenforschern wie Knuth zu. Im Bereich der Mathematik hat Tim Gowers eine solche Initiative gestartet. Er ist Träger der Fields-Medaille (was dem Nobelpreis in Mathematik entspricht) und wurde 2012 durch königlichen Erlass zum "Sir" geadelt. Im gleichen Jahr verbreitete er auf der von ihm initiierten Webseite *thecostofknowledge.com* den Aufruf an die Wissenschaftler, zukünftig weder als Autor noch als Gutachter noch als Herausgeber für Elsevier tätig zu werden. Viele sind diesem Aufruf durch öffentlich sichtbare Zeichnung gefolgt (Ende 2015 waren es über 15000).

Wenn es auch manche Kritik an diesem Boykott gibt (etwa dass es eigentlich nicht nur um Elsevier geht und dass auch kommerzielle Verlage eine – wenn auch teure – Leistung erbringen), so hat "The Cost of Knowledge" doch einiges bewegt für den notwendigen Wandel der Publikationslandschaft.

Viel hängt davon ab, dass die Wissenschaftler sich bemerkbar machen.[4] Eine solche Stimme aus jüngster Zeit ist ein erfrischender Artikel von L. Hilty über seine Erfahrungen zu den Leistungen und Fehlleistungen eines kommerziellen Verlages bei der Herausgabe eines Sammelbandes[5].

Allmählich hat sich auch die Politik in einigen Staaten (etwa in Großbritannien oder in den Niederlanden) des Themas angenommen, vor allem um zu erreichen, dass die Ergebnisse staatlich finanzierter Projekte offen (d.h. ohne Bezahlschranken für die Leserschaft) publiziert werden.

Es gibt inzwischen auch zahlreiche Initiativen für den Aufbau einer nicht-kommerziellen Publikationskultur. Das Schlagwort ist hier "open access" (also die freie Zugänglichkeit der wissenschaftlichen Arbeiten, häufig verbunden damit, dass das Copyright beim Autor verbleibt). Anfänglich wurde hierunter (insbesondere in der "Netz-Community") die völlig kostenfreie Verbreitung von Literatur verstanden – was natürlich eine Illusion ist. Es muss ja eine digitale Infrastruktur geschaffen und betrieben werden, mit Personal- und Sachkosten. Inzwischen haben sich mehrere Modelle der open-access-Publikation entwickelt. Auf der einen Seite stehen Projekte der Wissenschaftler (etwa arXiv), die – als Faustregel – mit Kosten in der Größenordnung von 100 Euro pro Artikel kalkulieren, auf der anderen Seite die inzwischen verfügbaren Angebote der kommerziellen Verlage, welche Beträge in der Größenordnung des Zehnfachen verlangen. Der Autor kann durch Zahlen eines solchen Betrags die Freischaltung seines Artikels sicherstellen.

Zu diesen Fragen seien einige persönliche Eindrücke und Einschätzungen angefügt. Dem Aufruf von "The Cost of Knowledge" habe ich mich vor drei Jahren angeschlossen. Aber schon länger konzentriere ich meine Publikationsaktivitäten praktisch vollständig auf nicht-kommerzielle Projekte (im Bereich der theoretischen Informatik). So konnte ich bei der Gründung der Zeitschrift *ACM Transactions on Computational Logic* (im Jahre 2000) mitwirken und war für acht Jahre im Editorial Board. Von 2004 bis 2012 war ich dann Mitglied im Gründungs-Editorial Board der Zeitschrift *Logical Methods in Computer Science*, deren Organisation (als "overlay" auf arXiv) Jiri Adamek an der TU Braunschweig in selbstloser Weise übernahm. Die Verstetigung dieses Projekts über die Emeritierung von Jiri Adamek hinaus ist ein komplexer Prozess angesichts der Frage, wie sich eine Zeitschrift des "open access" finanzieren soll.

Die Finanzierung von "open access" hat sich noch nicht im Sinne einer für die Wissenschaft tragfähigen "Kultur" stabilisiert. Die Devise "author pays", die den Vorteil hat, verführerisch einfach zu sein und die derzeit von vielen Seiten propagiert wird, erscheint mir völlig verfehlt. Sie verabschiedet sich von dem Prinzip, dass an guter wissenschaftlicher Arbeit durchaus auch allgemeines Interesse (und nicht nur das individuelle Interesse des Autors) besteht, und sie benachteiligt Forscher, die in finanzschwachen Institutionen arbeiten oder in

[4] vgl. etwa den Blogeintrag "Wake Up!" des Fieldsmedaillienträgers David Mumford: http://www.dam.brown.edu/people/mumford/blog/2015/WakeUp.html, u.a. mit einer Wiedergabe der in Abschnitt 2 erwähnten Urkunde.

[5] L. M. Hilty: Was leisten Wissenschaftsverlage heute eigentlich noch?, netzpolitik.org, 26. August 2015, siehe `https://netzpolitik.org/2015/was-leisten-wissenschaftsverlage-heute-eigentlich-noch/`, oder *Informatik-Spektrum* 38 (2015), 302-305.

Entwicklungs- oder Schwellenländern leben. Es ist nicht absehbar, auf welche Weise die Autoren flächendeckend mit Mitteln ausgestattet werden können, um die Freischaltung gerade in kommerziellen Medien zu garantieren (und es ist sehr fraglich, ob diese Art der Finanzierung überhaupt im Interesse der Wissenschaft wäre). Es ist andererseits bemerkenswert, wie still sich die Universitäten und ihre Bibliotheken von ihrer Pflicht fernhalten, die durch Wissenschaftler in mühsamer Arbeit aufgebauten nicht-kommerziellen Publikationsstrukturen zu unterstützen. So wie die Cornell University für das Betreiben des arXiv-Servers von deutschen Wissenschaftsorganisationen Mittel erhält, so sollte dies auch für Zeitschriften wie *Logical Methods in Computer Science* geschehen. Zweierlei wäre dazu nötig: die Erstellung von *Download*-Statistiken, und die Übereinkunft der Bibliotheken, nach ihrem damit bestimmten Anteil die (sehr geringen) Kosten zu übernehmen. Dieses Modell erfordert natürlich einige Investitionsmittel für die Installation, doch wäre es eine System-Transformation, die endlich für eine im allgemeinen Interesse liegende Steuerung der Geldflüsse sorgen würde. Es ist merkwürdig, dass diese Möglichkeit, die Universitätsbibliotheken als Sachwalter der Wissenschaft wieder mehr ins Spiel zu bringen, in einschlägigen Papieren wie den Empfehlungen der Max-Planck-Gesellschaft[6] oder der Berlin-Brandenburgischen Akademie der Wissenschaften[7] keine Beachtung findet. Wenn die Universitäten und ihre Bibliotheken die Dinge auf diesem Felde weiterhin treiben lassen und nicht endlich selbst in die Hand nehmen, indem sie open-access-Plattformen der eigenen Wissenschaftler wirksam und *dauerhaft* unterstützen, dann marginalisieren sie sich selbst, und es wird neben den kommerziellen Verlagen und jenseits der eigenen Kontrolle eine Parallelwelt des scheinbaren "open access" entstehen; Musterbeispiel hierfür ist die Infrastruktur Researchgate, die zur Zeit noch das Image der individuellen Selbsthilfe der wissenschaftlichen Autoren pflegt, aber nach Sammlung eines hinreichenden Datenbestandes diesen Vorteil in Geld umwandeln dürfte.

Bei Konferenzbänden besteht eine Finanzierungsoption, die Zeitschriften nicht haben: Man kann die Teilnehmerbeiträge der Konferenz für die Publikationskosten heranziehen. Dies entlastet den einzelnen Autor (der ja in der Regel einer der Konferenzteilnehmer ist) und wird auch bei den meisten Konferenzen so gehandhabt.

Im Zusammenhang mit einer Konferenzserie habe ich den Schritt vom kommerziellen zum nicht-kommerziellen Publikationsmedium selbst miterlebt, mitgestaltet und dabei interessante Erfahrungen gemacht.

Es ging um das jährlich stattfindende *Symposium on Theoretical Aspects of Computer Science (STACS)*, in diesem Bereich eine der beiden Spitzentagungen in Europa. Seit der Gründung Jahre 1984 waren die Konferenzbände bei den *Springer Lecture Notes of Computer Science (LNCS)* publiziert worden. Das erste Editorial Board von LNCS (ab 1973 G. Goos und J. Hartmanis, später kam J. van Leeuwen hinzu) trat in 2004 zurück, worauf die Qualität der Bände abnahm, die Anzahl enorm erhöht wurde (auf etwa zwei Bände *pro Tag*) und gleichzeitig die Kosten für das Abonnement stark anstiegen. In Spitzenkonferenzen wie STACS wuchs daher die Unzufriedenheit; man fühlte sich in dieser Serie nicht mehr wohl, und immer weniger Institutionen konnten sich LNCS leisten.

[6] https://www.mpg.de/9201460/flaechendeckende-umstellung-auf-open-access-moeglich
[7] www.bbaw.de/publikationen/stellungnahmen-empfehlungen/wisspublikation

Als die Tagung STACS 2007 bei uns in Aachen stattfand und von Springer keine Zusicherung kam, nach einer Karenzzeit (von ca. drei Jahren) die Konferenzartikel freizuschalten, entschloss sich das Steering Committee von STACS, gemeinsam mit der indischen Konferenzserie *Foundations of Software Technology and Theoretical Computer Science (FSTTCS)* die Publikation vollständig auf "open access" umzustellen. Das Informatikzentrum Schloss Dagstuhl (ein Leibniz-Institut im Saarland) unter seinem Leiter Reinhard Wilhelm war bereit, eine solche Plattform zu etablieren, die *Leibniz International Proceedings in Informatics (LIPIcs)*. Nur Konferenzen mit hoher Qualität wurden in diese Serie aufgenommen; heute, nach sieben Jahren, ist LIPIcs ein großer Erfolg, und ich freue mich, hier als Chair des Editorial Board mithelfen zu können. Aber auch hier fehlt die Unterstützung durch die Bezieher der publizierten Literatur, nämlich durch die Universitäten und ihre Bibliotheken. Was oben zu *Logical Methods in Computer Science* gesagt wurde, gilt auch in diesem Fall. Man gewinnt den Eindruck, dass sich eine Schere auftut zwischen den vielen jungen Wissenschaftlern, die in Vollversammlungen auf Konferenzen oder in elektronischen Abstimmungen für LIPIcs als Publikationsort stimmen, um sicherzustellen, dass sie ihr Copyright behalten und dass ihre Arbeiten auch *gelesen* werden, einerseits, und dem Überbau der staatlichen Institutionen (Universitäten, Wissenschaftsorganisationen, Ministerien) andererseits, die LIPIcs keinerlei Finanzierung gewähren ("author pays", "conference pays") und damit ihre Klientel im Regen stehen lassen.

Eine Merkwürdigkeit im Zusammenhang mit der Gründung von LIPIcs war die Behandlung der bei LIPIcs erscheinenden Konferenzen durch das ISI Web of Science (Thomson Reuters). Der Wechsel des Publikationsorts führte bei den betroffenen Konferenzen sofort zu einer Elimination in der Indizierung, obwohl sich an der inhaltlichen Exzellenz gar nichts geändert hatte. Eine befriedigende Klärung dieser Sache gelang nicht, da Thomson Reuters auf die meisten Anfragen überhaupt nicht reagierte; nur in einem einzigen Fall, nämlich zu FSTTCS, gab es eine Reaktion, allerdings nicht mit einem Sachargument[8]. Dieses privatwirtschaftliche Unternehmen ist eben nicht in die Regelkreise integriert, welche die Wissenschaftler selbst geschaffen haben. Es ist sehr fraglich, ob ein Unternehmen wie Thomson Reuters als "Eichamt der Wissenschaft" (dies ist sein gegenwärtiger De-facto-Status) geeignet ist. Wenn sich die akademischen Institutionen in zentralen Fragen der Wertung akademischer Arbeit privatwirtschaftlichen Unternehmen ausliefern (indem sie z.B. akzeptieren, dass "open access" Journale keinen Impact-Faktor haben und folglich "open access" Publikationen aus solchen Wertungen herausfallen), dann zeugt das nicht nur von mangelndem akademischen Selbstbewusstsein, es missachtet auch (vielleicht ungewollt) die ureigenen Interessen der Wissenschaft.

Nicht nur bei Thomson Reuters fehlt die Transparenz, die im wissenschaftlichen Betrieb eigentlich unverzichtbar sein sollte. Auch die Kosten für die aus Steuergeldern bestrittenen Zeitschriftenabonnements werden geheim gehalten, obwohl es ein klares öffentliches Interesse an diesen Daten gibt. Die Vertraulichkeit ist in den Abonnementsverträgen festgeschrieben; die Bibliotheken akzeptieren dies bisher und schweigen somit über ihre Abschlüsse. Doch gibt es hier leise Anzeichen für eine Wende. In der Schweiz ist es vor kurzem und nach jahrelangen

[8] Der Kern bestand in der Auskunft "there is a limit on the number of conferences we can index annually and it is dependent on our product needs."

Bemühungen (des unermüdlichen Christian Gutknecht) gelungen, Universitäten zur Offenlegung ihrer Aufwendungen an die Verlage zu verpflichten[9]. Wann wird es in Deutschland so weit sein?

2.6 Schlussbemerkung

Die Emanzipierung der Wissenschaft im Kontext des kommerziellen Verlagswesens ist ein langwieriger und schwieriger Prozess. Es ist offen, ob und in welcher Weise der Hauptzweck des Publizierens, nämlich die weltweite Verbreitung wissenschaftlicher Erkenntnisse, in Zukunft zu nachvollziehbaren Kosten erreicht wird. Viel hängt vom Verhalten der Wissenschaftler selbst ab, von ihrem Selbstverständnis – zum Beispiel ob sie sich als besoldete Dienstleister oder als verantwortliche Mitglieder der Gemeinschaft der Wissenschaft sehen. Und vielleicht noch mehr hängt von den wissenschaftlichen Institutionen (Universitäten, Deutsche Forschungsgemeinschaft, Max-Planck-Gesellschaft usw.) ab, deren Aufgabe es ist, die Interessen der Wissenschaft zu vertreten.

[9] `http://wisspub.net/2016/01/03/zahlungen-der-universitaet-zuerich/`

Chapter 3

Further Material on Publishing

Sharing scientific findings with others is a very important process for the exchange and progress within a research field, but also for the scientist's reputation (see Chapter 4). *Publish or Perish* is a common mantra in this context. This chapter covers some general information about the publishing process and characteristics of different forms of publications. Scientific writing itself will not be addressed here since lots of material already exists.

3.1 Categories

Publications may take place in different formats. Conferences are the most common ones with only a few months between submitting a first draft and the final publication.

3.1.1 Conferences

Publishing your work in CS often happens in Conferences:
Many conferences take place (bi-)annually. (I.e. CAV).
They publish a Call For Papers (CFP, usually on their websites and via newsgroups).
Call for Papers can also be found on sites like wikicfp.com.
In a CFP one can expect to find:

- General information about the conference (and its history)
- Its aims, a rather general statement of the goals of the conference.
- The scope, a list with the "topics of interest include but are not limited to".
- When the conference will take place (notice that for conferences, one of the authors is expected to be there and give a presentation)

- What to submit.
  - What kind of work (new research, case studies, tool papers, ongoing work).
  - How to present it (The length and the formatting).
    - poster, short paper, regular paper.
    - LNCS, ACM
- How to submit
  - I.e. easychair.
- Deadlines
  - Abstract deadline, usually a week before the paper deadline. One is expected to submit a preliminary abstract to help the program chairs select reviewers.
  - Paper submission. The final deadline for the paper to be reviewed. Here the timezone is often stated explicitly. Anywhere on Earth means that there must be a region on earth where it is still the stated date.
  - Rebuttal period, about 6-12 weeks later. Not all conferences feature such a period. If they do, then the authors get feedback from the reviewers and the opportunity to improve.
  - Notification, about 8-12 weeks after the paper submission. The authors are informed whether their paper got accepted together with the feedback from the reviewers.
  - Camera-ready copy, about 4-6 weeks after the notification. The authors of accepted papers are requested to submit their final version, to be published.
- The publisher of the work.
- The program chairs, the organizers
- The program committee, which will take part in the reviewing.
- Invited talks at the conference.
- Whether some papers might be invited to send their work to a journal.

From: CAV 2014 CFP, STACS 2014 CFP, DSN 2014 CFP, TACAS 2014 CFP.

3.1.2 Journals

Topics covered here:

- Journals in general feature a much longer reviewing process ...
- Journals address different scientific stage From: Benson & Silver: *What Editors Want*

3.1.3 Book Chapters

Writing Passages or chapters for Monographs or even whole books is also possible.

3.2 Some References and Links

On commercial publishing vs. Open Access:

- Berlin Declaration on Open Access:
  `http://openaccess.mpg.de/286432/Berlin-Declaration`
- Budapest Open Access Initiative:
  `http://www.budapestopenaccessinitiative.org/`
- Donald Knuth's letter regarding the Journal of Algorithms:
  `http://www-cs-faculty.stanford.edu/~knuth/joalet.pdf`
- Boycott of Elsevier:
  `http://www.theguardian.com/science/2012/feb/02/academics-boycott-publisher-elsevier`
  `http://www.economist.com/node/21545974`
- Open Access databases (e.g., `http://www.doaj.org`)
- Springer Open and Elsevier open access journals:
  `http://www.springeropen.com`
  `http://www.elsevier.com/about/open-access/open-access-journals`
- On cash flows between government, researchers, libraries and publishers – "Serial pricing crisis"):
  `http://www.arl.org/storage/documents/publications/in-oldenburgs-long-shadow.pdf`

Chapter 4

Ratings and Rankings

4.1 Motivation

Many decisions in academia, including those about promotions, hiring and funding, are based on an evaluation of the scientific output of individual scientists or institutions.

A very simplistic way to measure the productivity of scientists is to count their publications. However, this may at best give a rough quantitative impression of their scientific work, but fails to measure the quality of that work. There are various metrics that are intended to alleviate that problem by taking into account how often the publications have been cited. We give an overview of some commonly used metrics in Section 4.2.

In Section 4.3, we discuss the rating of journals and conferences. It is not unusual that the quality and impact of the research conducted by individual scientists or institutions is judged based on some rating of the journals and conferences where the research results have been published.

Finally, in Section 4.4, we consider the evaluation of institutions.

It should be noted that the way in which scientific work is currently evaluated has aroused some debate, with critics arguing that it provides undesirable incentives. In particular, the reliance on simplistic quantitative measures and the use of journal impact factors or other types of journal (or conference) ratings as indicators for the scientific value of individual publications has drawn criticism [15, 18, 13].

4.2 Rating Metrics

As described above, counting the number of papers published by a scientist is a simple way to rate his or her scientific output, but does not reflect the quality of the papers. Furthermore, it is easy to manipulate this measure by publishing bad or unimportant work at lowly rated conferences, possibly without any review process.

Other equally simple metrics are the total number of citations or the number of citations on the most cited work. The number of citations on the most referenced work heavily favors one-hit wonders that have a single high-impact publication. Given the fact that the scientists could easily be a lucky co-author

on single influential paper this measure should be used with caution. The total number of citations has similar flaws though less pronounced. Often a small number of papers or articles overshadows the vast majority of the scientists work.

In the following, we describe some more elaborate metrics.

4.2.1 h-index

The *h-index* or *Hirsch-index* was suggested by Jorge E. Hirsch in 2005 [10]. It couples the number of publications with the number of citations in the following way:

A scientist has index h if h of his/her N_p papers have at least h citations each, and the other $(N_p - h)$ papers have no more than h citations each.
– *Hirsch (2005)*

Intuitively the h-index improves over the simple metrics above since it combines quantitative and qualitative criteria in a meaningful way. It is impossible to game the index with a large quantity of low quality publications and the index is robust against singular large outliers in the citation count.

A disadvantage of the h-index is that it is not expressive for young scientists. Often it takes several years for a paper to get somewhat close to the finals number of citations. Therefore the h-index of a doctoral student will usually remain a small single-digit figure throughout his dissertation. Furthermore the robustness towards large outliers can also be a disadvantage. High impact papers are desirable but completely ignored in this metric. Leo Egghe has suggested the g-index as a more accurate tool in rating a scientists work, see Section 4.2.2.

4.2.2 g-index

The *g-index* was proposed by Leo Egghe in 2006 [9] as an improvement to the h-index. In contrast to the h-index it incorporates all citations even for very highly cited articles with diminishing returns.

Given a set of articles ranked in decreasing order of the number of citations that they received, the g-index is the (unique) largest number such that the top g articles received (together) at least g^2 citations.
– *Egghe (2006)*

Every citations on an articles that is within the top g most cited articles of an author contributes towards the g-index. Furthermore the g-index is designed to be always at least as large as the h-index. It rises earlier as the h-index since it does not matter which articles exactly got a citation as long as the the article is within the top $g + 1$ articles of an author. Therefore it reaches meaningful values faster than the h-index and it should have slightly more discriminatory power through its wider range.

4.2.3 e-index

The *e-index* was suggested by Zhang in 2009 [19] for similar reasons as the g-index. It is the square root of the surplus of citations on all papers that are counted in the h-index.

To achieve an h-index of 10 a scientists needs 10 different papers with at least 10 citations each. This sets the bar at 100 citations on his 10 most cited

publications. The surplus is the number of citations he has more than this minimal required number. e.g. a scientist with an h-index of 10 and 200 citations in total on the 10 most cited publications has a surplus of 100 and therefore a e-index of 10.

The aim of the e-index is to differentiate scientists with similar h-index but wildly different citation patterns.

4.2.4 i10-index

The *i10-index* was introduced in 2011 by Google in their academic search engine GoogleScholar. It is simply the number of articles with at least ten citations.

Presumably Google expects a paper with a moderate academic input to reach at least ten citations. In this case the i10-index attempts to quantify the amount of impactful scientific work a scientists has done.

The i10-index targets a more narrow band in a scientists career. Young scientists in their post-doctoral studies are likely to have a very small i10-index, while senior scientists are likely to achieve 10 citations on most average results. Therefore the index mainly counts the total number of somewhat decent publications, but does not showcase outstanding results.

4.2.5 h-index variants

There are numerous variations to the h-index. As examples we mention the *contemporary h-index* [16] and the *individual h-index* [1]. Both are meant to improve the accuracy of the index in case scientists are hard to compare.

Old scientists for example retain the h-index for their entire life, even though they have not published in several years. If you want to compare the recent scientific output of scientists of different age you skew the citation rates depending on the age of an article. Publish or Perish does this with $\frac{\gamma}{\delta}$ for $\gamma = 4$ and $\delta = 1$. This way recent papers can contribute the the h-index much faster and older papers fall of if they were no major results.

The individual h-index is an attempt to compare different fields. The average number of citations on a paper differs wildly between fields as well as the average number of co-authors. To combat this individual h-indice either divide the citations on a paper through the number of authors, or add count partial papers for the h-index.

4.2.6 aw-citation rate and aw-index

The *aw-citation rate* is attributed to Bihui Jin (2007) [12]. It measures the age-weighted number of citations to a whole body of scientific work. The number of citations on every publication is divided by its age in years.

The *aw-index* is the square root of the age-weighted citation rate to allow for comparison with the h-index. It has approximately the same number as the h-index if the number of citations per year remains constant.

4.3 Rating Journals and Conferences

Jounals and conferences are often compared based on some informal notion of reputation or prestige. However, there are also more formal metrics and rankings that are intended to reflect the quality and impact of these publication outlets.

4.3.1 Metrics for Journal and Conference Ratings

Citation-Based Ratings

Impact Factor [17]: The impact factor of a journal represents the average number of incoming citations of the articles published in the journal (within a certain time window).

Eigenfactor[2]: The Eigenfactor is intended to measure the "total influence of a journal on the scholarly literature". It is determined by the number of incoming citations to articles in the journal, weighted by the importance of the journal containing the citation.

SCImago Journal Rank [14]: Similar to the Eigenfactor, this measure rates journals based on incoming citations, weighted by the importance of the journal containing the citation.

h-index: The h-index and similar measures, originally designed to evaluate the scientific record of individual scientists (see Section 4.2), can also be applied to journals and conferences.

Publication Power Approach

The Publication Power Approach rates each journal according to the frequency at which leading researchers of the respective field publish their results in that journal [11].

Expert Surveys

Another type of rankings is based on surveys where active field researchers rate the quality and impact of journals or conferences.

4.3.2 CORE Rankings

The Computing Research and Education Association of Australasia (CORE), an association of university departments of computer science in Australia and New Zealand, rankings of computer science conferences and journals [7, 8]. In these rankings, journals and conferences are assigned a rank A*, A, B, or C, based on a combination of various criteria.

4.4 Rating Institutions

There exist several rankings of higher order education institutions by various organizations based on various criteria, among them e.g., research performance, student and international opinion, internationality, and eventual career success. These rankings are often produced by magazines, newspapers, websites, governments, or academics.

4.4.1 Global Rankings

Various national and international organizations compile worldwide university rankings, see [6] for an overview.

4.4.2 Regional and National Rankings

A number of regional and national rankings exist, an overview can be found online at [6]. Some rankings evaluate higher education institutions within a single country, while others assess higher education institutions within a continent. Here we present rankings of the Centre for Higher Education Development (CHE), a German association.

CHE University Ranking

The CHE yearly publishes the CHE University Ranking of German higher education institutions. Also Dutch, Belgian, Austrian, and Swiss higher education institutions with German study programmes are ranked. The ranking is primarily targeted at future students, in order to support their search for study programmes and locations. It is based on facts about study programmes, teaching, equipment, research, and student assessments on study conditions. Criteria with respect to research are expenditure of third-party funds as well as number of publications, citations, and doctoral dissertations. The results of the CHE University Ranking are published by the German newspaper Die Zeit [4] and are available both in German and English.

Until 2009, a CHE Research Ranking was compiled using data from the CHE University Ranking focusing on research. It it superseded by the ranking "Vielfältige Exzellenz" with emphasis not only on research, but also on application-orientation, internationality, and programme orientation. It is available in German at [5].

CHE Excellence Ranking

The CHE Excellence Ranking [3] is a university ranking for a selected group of European universities. The universities were selected with focus on research and internationalization indicators. It is designed to aid the search for a master's or doctoral degree programme.

4.5 References and Links

[1] Pablo D. Batista et al. „Is it possible to compare researchers with different scientific interest?“ In: *Scientometrics* 68.1 (2006), pp. 179–189.

[2] Carl Bergstrom. „Eigenfactor: Measuring the value and prestige of scholarly journals“. en. In: *College & Research Libraries News* 68.5 (May 2007), pp. 314–316. ISSN: 0099-0086, 2150-6698. URL: `http://crln.acrl.org/content/68/5/314` (visited on 03/16/2015).

[3] *CHE Excellence Ranking*. URL: `http://www.che-ranking.de/cms/?getObject=632&getLang=en` (visited on 04/23/2015).

[4] *CHE Hochschulranking*. URL: `http://ranking.zeit.de/` (visited on 04/23/2015).

[5] *CHE Vielfältige Exzellenz 2012*. URL: `http://www.che-ranking.de/downloads/CHE_AP164_Vielfaeltige_Exzellenz_2012.pdf` (visited on 04/23/2015).

[6] *College and university rankings*. URL: http://en.wikipedia.org/wiki/College_and_university_rankings (visited on 04/23/2015).

[7] *CORE Conference Rankings*. URL: http://www.core.edu.au/index.php/conference-rankings (visited on 03/17/2015).

[8] *CORE Journal Rankings*. URL: http://www.core.edu.au/index.php/journal (visited on 03/17/2015).

[9] Leo Egghe. „Theory and practise of the g-index". In: *Scientometrics* 69.1 (2006), pp. 131–152.

[10] Jorge E. Hirsch. „An Index to quantify an individual's scientific research output". In: *Proceedings of the National Academy of Sciences of the United States of America* 102.46 (2005), 16569—16572.

[11] Clyde W. Holsapple. „A publication power approach for identifying premier information systems journals". en. In: *Journal of the American Society for Information Science and Technology* 59.2 (Jan. 2008), pp. 166–185. ISSN: 1532-2890. DOI: 10.1002/asi.20679. URL: http://onlinelibrary.wiley.com/doi/10.1002/asi.20679/abstract (visited on 03/17/2015).

[12] BiHui Jin et al. „The R- and AR-indices: Complementing the h-index". In: *Chinese Science Bulletin* 52.6 (2007), pp. 855–863.

[13] David Lorge Parnas. „Stop the Numbers Game". In: *Commun. ACM* 50.11 (Nov. 2007), pp. 19–21. ISSN: 0001-0782. DOI: 10.1145/1297797.1297815. URL: http://doi.acm.org/10.1145/1297797.1297815 (visited on 12/08/2014).

[14] *SCImago Journal Rank*. URL: http://www.scimagojr.com/journalrank.php (visited on 03/17/2015).

[15] Per O. Seglen. „Why the impact factor of journals should not be used for evaluating research". In: *BMJ* 314.7079 (1997), p. 497. URL: http://www.bmj.com/content/314/7079/497.1?eaf= (visited on 12/08/2014).

[16] A. Sidiropoulos, D. Katsaros, and Y. Manolopoulos. „Generalized h-index for disclosing latent facts in citation networks". In: *Scientometrics* 72.2 (2007), pp. 253–280.

[17] *The Thomson Reuter Impact Factor*. URL: http://wokinfo.com/essays/impact-factor/ (visited on 03/17/2015).

[18] Ronald D. Vale. „Evaluating how we evaluate". In: *Molecular Biology of the Cell* 23.17 (Sept. 2012), pp. 3285–3289. ISSN: 1059-1524. DOI: 10.1091/mbc.E12-06-0490. URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3431935/ (visited on 12/10/2014).

[19] Chun-Ting Zhang. „The e-Index, Complementing the h-Index for Excess Citations". In: *PLoS ONE* 4.5 (2009), e5429.

Chapter 5

Plagiarism

5.1 Introduction

The word “plagiarism” has its origin in the Latin language and literally meant “kidnapper”. A contemporary definition, taken from the Oxford Dictionary, defines plagiarism as “the practice of taking someone else’s work or ideas and passing them off as one’s own”.

If you take a look on the wording of the human rights, there is a definition for the property in its section 17. Here it is written that everybody could own property and it is not allowed to bereave somebody of it. No distinction is made between material and immaterial property [3].

Plagiarism can be distinguished by the two categories of work or ideas. The larceny of work, for example a release of an article, is comparatively easily regulated by means of the copyright of the author. An article is (ideally) the property of the writer and nobody is allowed to copy a text without a proper reference. On the other hand, the usage of a published idea can be set against the larceny. Such a usage is possible and legal because everybody can use such an idea and combine it with own contributions. It is the same approach which everybody should do every day to train or acquire the own or new skills [4].

The topic of plagiarism has gained much attention within the last years by some prominent cases, some of them involving leading politicians.

But this is only a symptom. The deeper aspects of the subject are raised by the development of extremely powerful technologies which facilitate editing and copying texts and (via the internet) allow for a world-wide exchange of information. The dissemination via the internet not only gives people nearly unlimited access to information, but it also causes more and more people to publish their work online. Thus the available amount of texts, papers and information increases drastically. These facts, combined with the possibilities offered by modern data processing make plagiarizing much easier and, thus, also enable more people to plagiarize. Despite the fact that further developments like software tools can help to detect plagiarism, the topic of plagiarism becomes more relevant for more people, and its facets become ever more complicated.

5.1.1 Motivation

The motivations why people do plagiarism are quite different. On the one hand the plagiarism is done accidentally by the lack of knowledge about correct citations, or sometimes the reason is simply missing care to mention a work or a citation. On the other hand plagiarism appears also as something done with purpose. Here the background often is the desire to enhance own work by presenting new ideas and results of another person to get some sort of benefit.

The concrete reasons why students plagiarize arise in similar categories. The background sometimes is a missing motivation for a theme or a study, laziness or the lack of time at the end of a project. The plagiarist wants to finish a text without doing the work. Also the factor of stress to get a good exam is for many students an example for a motivation to plagiarize. In the scientific work there is a popular misconception of using results from supervised work. In the "internet culture" one finds the opinion "there is no intellectual property": Everybody can use all ideas without identifying the sources.

5.1.2 Types

According to [5], [6] and [1], there are different types of plagiarism. The first type refers to whole texts or text fragments of another author without a citation. Also the replication of ideas or concept of another work is an example for such a plagiarism if the citation is not given. Another form of plagiarism is the wrong usage of sources. Sometimes a work will have (some) sources which are not used in the text to suggest that the article has a higher standard. For example the author writes a source of a famous scientist in his research field in the list of references without using it. A usage of different sources for one text passage without a citation of all used sources also is possible. So a usful principle is: Always try to quote the original results. Programs like Citavi or EndNote help to run a literature search and to organize the sources.

If someone reuses his own ideas or results without citing one speaks of self-plagiarism. Self-plagiarism is controversial in science, because it seems difficult to assume that one can steal one's own property. Löwer [2] describes that it is not possible to do a self-plagiarism.

Another type of plagiarism is ghost-writing: A person employs another person to write the work under the name of the employer. This is also a difficult legal issue when the actual writer is paid for his/her work.

5.2 Ways and Methods of Plagiarism

The field of plagiarism is a really large one, there is a vast variety of scenarios where one can speak of plagiarism. The differences are based on the details of who plagiarizes, what is plagiarized, in which way. This basically results in different methods to plagiarize. The most common ones are presented in the following, based on [6].

5.2.1 Copy-and-Paste

As the name already indicates, here people take an existing text, copy it (either completely or in parts) literally and paste it into something they denote "their"

work, without mentioning the original author. Having two texts and the suspicion that one is copied from the other, it is easy to check this by comparison. However, sometimes the plagiarists include small changes, e.g. by changing the order of enumerations. There is usually no certainty whether a passage was copied. Nevertheless, some hints might be: Spelling mistakes (as some people even copy those, rather than doing proofreading), Changing formats or singular occurrences of unusual formats (as some people also copy style elements without adapting them to their own writing). Another sign of plagiarism is automatic formatting, like the underlying and blue text color that appears when one pastes an URL into an office document (this common mistakes happens when copying text from a web page, where the text contains a hyperlink).

5.2.2 Plagiarism by Translation

Here, the plagiarist takes a text written in another language and simply translates it. Translating does not only prevent an instantaneous detection of the plagiarism. Often, language skills are not perfect, so that reformulations of sentences are necessary and this, again, diminishes the similarity between source and plagiate. Nowadays, in the time where automated translation is easily possible (e.g. via babelfish.com, a free online translation website), this method of plagiarism may be considered to be a very impudent one. Proving a translated text to be copied is much harder than in the case of copy-and-paste. Detection, however, can be based on strange and uncommon formulations. The occurrence of peculiar words can be another hint.

This method of plagiarism is quite popular with scientist, e.g. when writing a thesis in German and copying from papers written in English, or vice versa.

5.2.3 Shake-and-Paste

Again, the plagiarist copies text written by another author and uses this text in his work. In contrast to copy-and-paste, here several sources are used such that the result is a text consisting of several passages, each copied from possibly another source. As with the copy-and-paste method, plagiarism can be proven by literally comparing words. Hints that a text might be produced by shake-and-paste are changing styles of writing as a consequence of different sources. Also, different spellings of one and the same name or technical term as well as changing formatting from one passage to the other can evoke the suspicion of plagiarism.

5.2.4 Adopting the structure

According to [6], this is a very contended way of performing plagiarism, but setting up a structure of a text is a mental effort as well as writing a text. However, the intensity with this method may vary within a large range. Where the first plagiarist might only acquire another author's idea of how to structure a text, the second takes the whole structure of another work and only rewrites the content with his words or uses his results. This kind of plagiarism is hard to detect since finding the same words is not sufficient, but the reader really has to understand the content and recognize somebody else's work.

Despite the efforts to discover and fight plagiarism, one has to keep in mind that a similar way of thinking or working may also result in similar structures of text. Hence, even if an argumentation resembles another, this is not necessarily plagiarism.

5.2.5 Software plagiarism: Renaming variables

This method is relevant for people working in the fields of computer science, mathematics and other disciplines involving programming. In programs that are intended to compute a certain solution, looking at their output on the same given input does not give helpful information on their relation to each other. To identify plagiarism, a look at the program code is necessary. And here, names of variables, procedures, functions, interfaces and so are dominant. Hence, changing these apparent names makes the code look different – even if a plagiarist took the code from another programmer and only renamed the most important elements. For people working in these fields, thus, it is important to randomly check the code of different programs in order to avoid plagiarism. And when checking, a look at the structure, the interrelation of variables, functions and classes is important, where the names should be widely ignored. However, especially in medium or larger programming projects this task is hard; and always a certain understanding of what could be copied in which way is absolutely necessary.

5.3 Dealing with Plagiarism

5.3.1 Where Plagiarism can be Found

Plagiarism is a phenomenon which can be found in various areas of scientific life. First, we mention people, usually students or examinees, who need to work out some task or exercise and, afterwards, hand it in to be corrected by some assistant or supervisor. More generally, some task will be evaluated in order to determine a grade. Especially when dealing with exercises that need to be completed to pass a course it is necessary to take good care: Plagiarizing often is easy as the same or at least a similar exercise might have been posed in former years or at other universities, and usually there are many students in the same course that need to do the same exercise. Thus, the "material" from where one can copy, is extremely large. Also, the inhibition to plagiarize usually is not that high since students assume the exercises' correctors do not work very carefully.

Besides students plagiarizing when performing assignments, also professors can plagiarize. This is, for example, the case when a professor publishes or uses the results from master, diploma or PhD theses he supervised, without disclosing the actual author or, at least, contributor. If the former student already left academia those cases possibly never come to light. But even if they emerge, the professor's position is strong if the student is still in a relation of dependence (e.g. in a dissertation project).

A third field where plagiarism can be found is in research when researchers, working in the same subject use the other one's work without permission and without announcing. This can happen when one reviews somebody other's work and exploits this access by using the results. Or a researcher just copies from other author's papers and hands the results in on other conferences or journals.

Especially when these are smaller and less renowned ones, their review process often is not thorough enough to discover plagiarism.

5.3.2 Why Plagiarism should be assumed

As it turned out above when discussing different methods to generate plagiarism, it is important, or at least helpful, to nourish a certain initial suspicion in some situations. One of them is when a perfect writing style of writing is found with an author of lower quality. Depending on what kind work one is dealing with, an extraordinary good style of writing should attract the reader's attention. For example, when reviewing student homework it is unlikely to get a paper which presents results in perfect scientific style. In general, the choice of words as well as the readability are suitable indicators. Especially when dealing with formal work, as it is the case in computer science and mathematics, the quality of introducing formal notation, setting definitions and conducting proofs should be considered when asking the question whether the existent quality of text is in accordance with the writer's competence level.

As written above when presenting different methods, a change of the writing style may be a hint on text taken from another author. This style change can include a sudden broad use of subject terms, when this use was not visible before. A change can also be found in the style of writing headlines: Where one author prefers concise ones, another possibly choses longer and more comprehensive titles. Textually, also the average length of sentences might be a criterion.

Other significant changes might occur in the style of formatting the text. Different line pitch or typeface are quite obvious and easy to discover. Besides significant changes between different parts of a work, further hints can be remarkable grammatical errors, which might be a consequence of using an automated translation tool. No reason not to nourish suspicion is when one finds a name or an established technical term written in different ways, since this is a strong hint to the Shake-and-Paste method, as described above. Usually, there is no reason for a single author to use different ways of spelling the same thing; so differing spellings might very well be caused by copying from others.

5.3.3 What to do Against Plagiarism

Before thinking of consequences one has to make sure that a work is plagiarism. And with this usually the question of how much of a text needs to be copied to speak of plagiarism arises. Some consistency might be coincidence, some similarity in a text's structure might be the natural result of the same underlying problem. But iextreme cases indicate that there is intent rather than coincidence. This can hardly be measured and definitely depends on the subject and the situation.

How to fight plagiarism is not easy to answer and heavily depends on the field one is working in. In the case of undergraduate students copying their exercise solutions, plagiarism is relatively easily verified, and methods of punishment are obvious (by bad grades or failure to pass the exam). For the more serious case of theses and research papers, however, the situation is more difficult. Here one cannot limit the sources that are usable for plagiarism and one has to resort to software tools for automated search in the web.

But here a difficult question arises, namely whether a policy of distrust is appropriate towards all thesis work and research papers, and if applied for selected cases, one has to justify this selection. It is yet open whether the systematic use of software for plagiarism detection is desirable – at least in disciplines like computer science, mathematics and engineering.

5.4 References and Links

[1] Gerhard Fröhlich. „Plagiate und unethische Autorenschaften“. In: *Information Wissenschaft und Praxis* 57.2 (2006), pp. 81–89. URL: http://eprints.rclis.org/7416/1/plagiate.pdf.

[2] Franz Himpsl. „Können Doktoranden aus Versehen plagiieren?“ In: *Die Zeit* 06/2014 (2014). Interview mit Wolfgang Löwer. URL: http://www.zeit.de/2014/06/promotion-plagiat-versehen/komplettansicht.

[3] Bernd Klein. *Geistiges Eigentum – Abgrenzung materielles Eigentum*. URL: http://www.geistiges-eigentum.eu/eigentum.php (visited on 07/23/2014).

[4] Bernhard Kuntz. *Vorsicht Plagiat! So schützen Sie sich vor Content-Diebstahl*. URL: http://www.unternehmer.de/marketing-vertrieb/157650-vorsicht-plagiat-schuetzen-sie-sich-vor-content-diebstahl (visited on 07/23/2014).

[5] Christian Schwarzenegger and Wolfgang Wohlers. „Plagiatsformen und disziplinarrechtliche Konsequenzen“. In: *unijournal* 4/2006 (2006), p. 3. URL: http://www.rwi.uzh.ch/lehreforschung/alphabetisch/schwarzenegger/publikationen/unijournal-2006-4.pdf.

[6] Debora Weber-Wulff and Gabriele Wohnsdorf. „Strategien der Plagiatsbekämpfung“. In: *Information - Wissenschaft & Praxis* 57.2 (2006), pp. 90–98. URL: https://komm-in.uni-hohenheim.de/fileadmin/einrichtungen/agrar/Studium/Plagiate/strategien_plagiate.pdf.

Chapter 6

Organizing a Workshop or Conference

Scientific work cannot be done just by studying and writing papers. Scientific progress needs personal contacts through talks and discussions. An ideal way of getting into such contacts with a larger community of fellow researchers or established scientists is the visit of conferences, workshops, and research schools. The success of these events depends largely on good planning and organization. In this chapter we summarize some issues that are relevant for efficient planning and organization of scientific meetings. This summary represents what we learned in the research training group AlgoSyn when organizing two major events in Aachen (the Aachen "Fall School on Algorithmic Game Theory" in September 2013, and the international conference "Frontiers of Formal Methods" in February 2015). Both events were completely organized by the doctoral researchers of AlgoSyn, with some help by the advisors regarding the invitation of keynote speakers. (The work of the program committee was also taken by the advisors since doctoral researchers were explicitly invited to submit and conflicts of interest were to be avoided.)

We proceed in a rough time schedule, with early tasks, tasks before the conference, and tasks during the conference.

6.1 Early Tasks

6.1.1 Overview

About one year before the event, a number of important issues has to be decided; and in some sense, these decisions are the most important for the success of the event:

- Decide on the title and profile of the event.
- Decide on the venue (and ensure that the venue is available and suits the purpose, not only regarding the talks but also regarding reachability and availability of "lobby space" where the coffee breaks take place) – do the room reservation and (if several lecture halls are needed) ensure that they are easy to reach from each other.

- Decide on the date, and prepare a schedule of all "important dates" (announcement and call for papers, submission deadline, date of notification of acceptance, registration deadlines)
- Make a first planning of the budget, and organize funding.
- Most important: Decide and check availability of keynote speakers and program committee – note that this information has to enter the announcement and the call for papers in order to create attraction of the event.
- Following this, formulate the announcement and the call for papers and distribute them via the web.
- Create the conference website.
- Design a nice poster and send it to selected institutions.
- Make pre-reservations at hotels and other locations for accomodation in order to ensure that enough accomodation is available when participants register.
- While doing all this: Distribute tasks of organization among all members of the organizing team, valid until the event takes place.

Let us add some general remarks:

6.1.2 Organizing Team

Usually some institute, some university chairs or a couple of chairs, or another unit (such as a research training group) take the organization of the event. The organizing team will often consists of about five to ten people (research associates, secretaries, and others). It is helpful to have some experienced people among them; others do not have any background in these matters and have to jump in at the deep end when they organize their first event. Hosting a conference of course means lots of hours of work, but for a doctoral researcher it is a good example of how to train soft skills during the doctoral studies. This is in contrast to industrial corporations where such jobs would be outsourced to professional event managers (who are possibly better at the economic aspects but would not know well how to serve the respective scientific community best).

6.1.3 Funding

Many larger events as conferences or symposiums are supported by industrial sponsoring. It is very important to arrange the sponsoring contracts with the business partners early enough to have a safe budget before the planning of the event goes into details. In particular, the registration fee that has to be mentioned in the announcement and call for papers will depend on the funding that is available. Therefore, potentially interested industrial partners must be found and contacted in a professional and appealing way. Especially the first interaction with a company is crucial as the counterpart must see the benefits they can take from supporting such an event. Hence, there should be a concept providing possible sponsoring packages. In larger events one can distinguish several classes of sponsoring (often called *platinum, gold, silver*). In this case, the

corresponding promotion services need to be designed carefully so that organizing team can make an attractive and appropriate offers.

6.2 Before the event

About half a year before the event, all the early tasks are done, and in particular the call for papers is well under way; in fact the deadline for submissions deadline will be around this time.

- Detailed Scheduling: Several questions have to be settled.

  When shall the participants arrive? When will they leave? It is not a good idea to plan the closing session too late on the last day as the participants will likely want to return home on the same day.

  Are there social events? Most conferences offer a dinner, a guided "touristic" tour or a combination of both. Sometimes also a "get-together" (on the evening before the start of the conference) is offered, with snacks and drinks.

  Will there be one, two or even more tracks? In most cases there are several sessions per day (some subsequent, some in parallel). Note: The more often the schedule is changed the more difficult will the organization of the conference be.

- The payment of conference fees has to be organized; the convenient payment by credit card is sometimes difficult to implement in universities

- Reservation of venues and entrances for all social events.

- Organization of catering for coffee breaks (and possibly lunches): The traditional coffee breaks between the sessions are a vital and maybe even the most important element of every academic event. Experience has shown that a huge number of valuable contacts have been made during the breaks and that many research-related ideas have been born meanwhile. Accordingly, planning an appropriate number of sufficiently long coffee breaks is mandatory for every conference organization. Therefore it is best to employ a professional caterer as there are many things one can possibly miss or forget and nothing would be more embarrassing than having cold or too less coffee in the break. The professional suppliers know best how much coffee will be needed and how to arrange tables, coffee, tea, cookies, cake etc. Nevertheless the organization team is in charge of providing a room that offers enough space and a convenient atmosphere for a break and that is located near the auditoria.

  At many conferences there is also a lunch or at least some snacks and a soup. Especially if there is no attractive gastronomy in the surrounding of the conference venue, this is essential. If the guests are supposed to go for lunch by themselves, the hosts should provide a list of possible locations including a description of how to get there. Note, that many especially international guests will be doctoral students without monthly salary and that those will most probably be unwilling to spend so much money on their lunch. In contrast, if the lunch is included in the registration fee of

the convention, it will be paid by the participant's home university in most cases. The same holds for the social event, especially for the dinner.

- The "Survival Guide" has to be written, with informations on
  - travel information, local traffic
  - local history
  - restaurants
  - the full program
  - maps on how to reach the venue
  - information on how to access the internet during the conference
- Early enough there should be information on the conference website which accomodation is recommended.
- The proceedings volume has to be compiled and printed.
- Conference Bag: One has to decide on the items to be put into the conference bag. If possible the bag should be produced with the conference name on it. Usual items are: The survival guide with program, the proceedings (hardcopy or USB stick or other), local tourist information, a pen with a (paper-) notebook, some little souvenir.
- Internet connection: Sometimes, it is useful to supply a way to connect to the internet in addition to eduroam.
- For the invited speakers, hotel reservation should be made. One should also check that the program is compatible with the travel restrictions of the invited speakers.

6.3 During the conference

This is of course the most exciting time for the organizing team, and the best strategy here is to keep calm. A non-exhaustive list of relevant items is the following:

- In the neighbourhood of the venue: Placement of signs pointing to the conference venue
- Equipment the registration desk with office material, internet connection, etc., and schedule for the people serving at the registration desk (at any time somebody should be there)
- Organization of who chairs which session – there should be clock and signs (3 min / 1 min / STOP) for the session chairs
- Technical support and equipment: It is a good policy to have at any time one local person in each lecture room who can help when technical problems arise. In particular, this person is needed for an efficient switch from one speaker to the next. The session chair will not be able to do this.

Every conference room needs at least an LCD projector. Notice, that the projector is likely to become broken when it is required most. So the organizers should provide enough backup devices. Having an overhead or dia projector is not necessary in nowadays computer science conferences anymore, unless it was announced to provide one. It is up to the hosts if the slides shall be shown on a dedicated presentation laptop or if every speaker should bring her own one. In the latter case that should be communicated explicitly. However, a backup laptop should be provided anyhow. Be aware of incompatible plugs, broken adapters, etc. All components should be tested before. For the speakers it is nice if they make use of a presenter, which allows them to switch slides while walking around and point on the screen with a laser. Power plugs at any seats are nice to have but not mandatory. Nowadays laptop batteries reach for six or more hours. (A positive side effect is that a participant with an empty battery is forced to follow the speaker instead of surfing on the internet....)

- Photos: Somebody should take photos during the whole event; these (or selected) pictures are then put on the conference website as a memory. This detail is appreciated very much by the participants.

6.4 After the event

The most interesting point here is the payment of all the invoices (from the caterer, restaurant of the conference dinner, etc.) and the conclusion whether the budget planning was good.

But more important is the echo of the participants. In fact, a good sign that the organization was perfect is just the lack of complaints. The ideal conference organization is one that just runs efficiently and invisibly.

Chapter 7

Elements of Doctoral Studies

In this chapter, selected aspects of the phase of doctoral studies are discussed. Many more issues than those treated here are important. Examples of such omissions are "The art of presenting science" and "How to write a project proposal". But on these topics many references exist which we do not want to duplicate here.

7.1 Motivation

Like for any other major project, it is important to have good motivation in order to pursue the doctoral studies in a productive way, with realistic expectations. We point out possible reasons to start doctoral studies; in fact, the goals of doctoral candidates might differ considerably. We also discuss the character of doctoral studies which in general is somewhere between working as a regular employee and being a student.

The core motivation should be to participate in true research – i.e. participating in the efforts of the scientific community to obtain progress in science or engineering. Without such "drive" and without some fascination it will be difficult to pursue doctoral studies with success.

So the stereotypical doctoral candidate usually has the honorable goal to be an excellent researcher, sometimes with additional expertise – for example in teaching. But one observes that in reality, the actual reasons for starting doctoral study may also be of different nature.

Upon finishing his or her master thesis, some students simply want to spend another few years at the university. They may not want to leave their social environment, they may be shy of entering industry where they may be subject to stricter rules and more stress, or they may simply not know what to do next. This also explains another interesting fact: Once a student started working for some company, he or she very rarely returns to the university to start doctoral studies.

A strong interest in the research topic sometimes is of secondary importance, as the selection of the research topic is often influenced significantly by the funding. The prospect of an academical career as a permanent research assistant

or even a professor is so uncertain that it is not wise to take this as the first motivation.

In contrast to the anglo-american academic world, being a doctoral candidate is usually much more like being employed than being a student. A doctoral candidate is part of the scientific staff and as such the duties range from researching and cooperations with other chairs and universities (and even industry) to teaching and administrative tasks. This in turn is reflected in the financial situation. A doctoral candidate in computer science may be payed on a full position – which can of course be very different in other subjects.

The education of the doctoral candidate besides the research activities is usually done in separate and special courses offered by a university (in Aachen: by the Center of Doctoral Studies). This is comparable to in-house education programs in large companies.

7.2 Teaching

Teaching represents a large part of the responsibilities of a doctoral student. Many professors will try to equally distribute the teaching workload amongst their students, so nearly everyone will be responsible for teaching at some point.

While some doctoral students will see teaching as a duty, it is in fact a valuable experience and a great opportunity to (among other things) strengthen didactical thinking, presenting skills, and also the ability to handle criticism and stay calm, which is all of great benefit to ones personal skill set.

The scope of duties is quite broad and involves duties such as giving exercise classes or supervising seminar reports, practical projects and theses. In the following we will go through some of the most common teaching activities and point out some of their main aspects.

7.2.1 Organization of a Course

When it comes to teaching the organization of a course is one of the most common and demanding tasks. Giving a course is typically a lot of effort which professors will usually delegate to their doctoral students. The variety of tasks here is broad and we will present some of them in the following.

Exercise Classes

Probably one of the most common duties is the organization of exercise classes, as most lectures have such an exercise class associated with them. The intended effect of exercise classes is that the students reinforce and deepen their knowledge on the lecture material, become familiar with it and exercise performing routine tasks which are assumed to be common knowledge in the further course of the lecture. The responsibilities in organizing an exercise class involve:

1. Coming up with exercises.
2. Coming up with solutions to the exercises.
3. Correcting and grading the exercises turned in by the students.
4. Presenting the solutions of the exercises to the students.

There is usually already a pool of old exercises one can draw from, so a mixture of old, well proven (slightly modified) exercises and new exercises is a good fit for a new exercise sheet.

It is furthermore important to synchronize the contents of the exercises with the contents of the lecture. Both should thematically proceed at the same pace so that they are at the same level. In order to ensure this it makes sense for the doctoral student to either consult with the professor or simply visit the lecture themselves. Occasionally doctoral students are also asked to organize exercise classes for lectures which they have not attended themselves before during their own studies. In this case visiting the lecture during the course of the semester is especially recommended.

After the students have turned in their exercises, the solutions to the exercises should be presented to or discussed with the students. It is a good idea to prepare solutions to the exercises *before* handing the exercise sheets out to the students. Oftentimes when preparing the solutions one finds small errors in the exercises which at that point can easily be fixed. This is a lot harder once the exercise sheets are handed out to the students. While errors in the exercises, though undesirable, are less of a problem in an advanced course, they tend to pose much confusion and lack of understanding amongst the students in undergraduate courses.

When coming up with solutions one might also find that a particular exercise is in fact to hard or even to easy to solve. A priori insights into the hardness of preparing a solution also help in coming up with an appropriate grading scheme.

The organization of exercise classes, especially for undergraduate courses with hundreds or sometimes even thousands of students, is typically a huge amount of effort. Therefore there are usually multiple doctoral students assigned to one course. Furthermore, for large courses there are generally financial means available to hire student assistants which grade exercises, give the exercise classes, and discuss the lecture material with the students in smaller groups. A well proven concept are weekly briefings in which the doctoral students present the exercises and the solutions to the student assistants. During these meetings it is often the case that small errors or misleading formulations in the exercises are revealed. The feedback of the student assistants is valuable also because they are typically closer to the student life.

It is important to note that most student assistants (as well as doctoral students and professors) have their own style of teaching which evolves and develops over time. It is therefore not advised to micromanage the student assistants but rather have a hands–off approach and give advice through positive reinforcement.

Preparing and Grading Written Exams

Professors usually ask the doctoral students who already organized the exercise classes to prepare tasks for the exam (and possibly for a repeat exam). This basically amounts to the same task as preparing exercises. However, even more care and attention has to be dedicated to this task than to preparing normal exercises. Errors in the task of an exam are *extremely* undesirable, lead to confusion and nervousness amongst the students during the exam and might entail an unpleasant aftermath.

Also special attention should be paid to an appropriate balance of difficult

and less difficult tasks in the exam. Too many difficult tasks might lead to an undesirably high failure rate whereas too few difficult tasks make it impossible to distinguish between good and excellent students.

After the students took the exam, the exam needs to be graded. Only in some circumstances the student assistants are allowed to help in grading the exam, so usually this is the duty of the doctoral students. Although officially only professors themselves may grade exams, the doctoral students are allowed to perform a pre–grading which is then revised and signed by the professor. Typically all doctoral students are required to help with the corrections of the exams in order to complete the grading as quickly as possible and in order to balance the workload amongst the doctoral students. In case of exams with a huge number of participants, some chairs will impose a ban on leave for all doctoral students during the correction phase which in these cases may last several weeks.

After the exams have been graded, it is mandatory to allow the students to inspect their exams. On these occasions students may suggest that an error has been made when correcting their exam and the correction can be revised.

Assisting in Oral Exams

When the form of the exam is an oral exam, the duties of the doctoral student are typically a bit different, although some professors will still ask the doctoral students to prepare tasks for the oral exam. The most common duty, however, is to record the oral exam in writing. This need not be a literal stenography of the entire oral exam but rather a record containing which questions were asked and how satisfactory the answer of the student was. After the exam, the professor usually involves the doctoral student in the discussion on how to grade the student.

7.2.2 Supervising Practical Projects

During their studies most students have to complete at least one practical project. Supervision of such a practical project shares some aspects with organizing exercise classes, namely that the doctoral student has to prepare practical tasks for the students to solve.

Even though the tasks should in principle be solvable, a certain solution to the tasks need not be predetermined. Instead the idea should be to allow students to be creative in coming up with their own, potentially innovative solutions to certain problems. The focus in practical projects should be more on a discussion of the idea behind the developed soultions and what advantages and disadvantages they have compared to the other presented solutions.

The grading should consequently also consider the effort that the students have put into developing an original idea for solving the task. Even though a solution might not be working or might not be so mature at the end of the practical project, the idea might still be quite sophisticated and good.

7.2.3 Supervising Theses and Seminar Reports

Supervising theses of any sort is another important duty of doctoral students. Theses represent a particular challenge for both students as well as for their

supervisors. Beside scientific expertise, an effective project management is also necessary. The aim is to design a meaningful, structured and focused cooperation between the supervisor and the student, which represents an important contribution to the successful completion of the work.

At any point in time the supervisor should be very clear about the objectives that he or she expects the student to achieve. Only through that an atmosphere of transparency and fair grading can be established which in the end makes for a good working atmosphere for both the student as well as the supervisor.

An important aspect is also the frequency of the meetings between the supervisor and the student. Some supervisors like weekly updates, others want to only be concerned with milestone results. The same holds for students: some like to present big as well as small advances frequently, others are more comfortable to present results not until they have completed a milestone. In the beginning of the thesis project, the supervisor and the student should sit together and communicate to each other which "type" they are. Having that clarified, one very common source of misunderstanding is eliminated right from the beginning.

Sometimes bachelor or master thesis projects generate such good results that one may think about publishing them. This of courses poses a win–win–situation for both the supervisor as well as for the student: In general, supervising a thesis involves many discussions with the student on a scientific topic, so the scientific results achieved in a bachelor or master thesis are typically the result of a close joint work of the supervisor and the student. Naturally the doctoral student is interested in publishing results to which he or she has contributed.

On the other hand, his or her thesis becoming a scientific publication is quite a rewarding experience for the student (especially since this is usually his or her first scientific publication) and may encourage the student to consider a further academic career. In case that the publication was submitted to a conference, there are in some cases special undergraduate funds available which provide financial aids for the students to attend that conference, which is another great and exciting opportunity for students to dive into the scientific life.

Supervising seminar reports, although they are smaller projects than bachelor or master theses, amounts to the same tasks as supervising bachelor or master theses and is also of great benefit to the supervisor. Apart from learning how to manage such a project and create a fruitful and constructive working atmosphere, one can also encounter excellent students and approach them later e.g. regarding a student assistant position or a bachelor or master theses project.

7.3 Life in Chairs and Institutes

The academic staff at a German university is almost organized in terms of professorships, also called *chairs*, where each chair is held by one professor. Often there are several chairs aggregated to one *institute*. The tasks related to an institute or chair cover a bundle of things as researching, teaching students, acquire their own finances, organizing academic events and so forth. In nearly all of these aspects a professor is supported by a staff of assistants, called *research associates*. In Germany the majority of doctoral candidates, except in medicine, are employed as research associates and vice versa. That means that the most doctoral candidates are deeply concerned with the administrative

questions a chair has to deal with. Consequently, they get in touch with a lot of administrative processes and spend a significant amount of time besides their own research work. On the one hand this may delay the time of graduation but on the other hand it gives to a researcher a lot of additional experiences and qualifications that would probably not have aggregated by just doing research. The experience shows that the amount of work that is not directly related to the candidate's research topic significantly differs between institutes of the engineering disciplines on the one hand and those in natural science on the other hand. As computer science is traditionally located somewhere between these and therefore is influenced by their cultures, respectively, the situation of computer scientists working in public research has aspects of both. This also has impact on the average duration of doctoral studies, which is rarely less than five years for an engineer while most physicists hand in their dissertation after three years.

Here are some exemplary issues a research associate has to deal with in everyday life.

Organizing Teaching Program – Providing a lecture is of course one of the central tasks of a professor and his staff. Besides the actual teaching activities as holding a lecture, creating exercises and exams, etc., there are many tasks that have to be organized "in the background".

Finances – A chair receives public funding for every course the professor offers. As providing lecture content, slides and exercise sheets and finally giving the lecture costs a lot of time, most institutes are strongly reliant on that money. Therefore they have to document the time the professor and any research associate have spend on each course and attest the overall effort to the university administration at the end of each semester.

Scheduling – Most professors have a very full calendar where the lecture schedule has to fit in. Their reserach associates support them by allocating properly sized lecture halls and matching the time table onto the students' curriculum. In this context, conflicts with other teaching units have to be resolved, also involving the university administration.

Participating in Commissions – In the university there are many opportunities to participate in the administration and the strategic development (e.g. when vacant professor positions are to be filled). To take part in this kind of work may be time-consuming, but it is also most rewarding by the experience one obtains in this activity.

7.4 Funding

An aspect of research which is often not considered when thinking of the tasks of a researcher is funding. In this respect, being a researcher has some similarity to being self-employed. Most researcher positions (including virtually all positions for doctoral candidates) are of limited duration and especially in Germany there are many cases where the duration of the initial position does not cover the whole time from starting the doctoral studies to the defense. Thus, a researcher needs to find sources of funding to continue his or her scientific work. In addition to the salary of the researcher, research might also include other tasks for which

financial support needs to be acquired (e.g., experiments and traveling). In the subsequent subsections, we list funding sources for a researcher's salary, we address duties of the researcher depending on the source of funding, and we treat benefits related to the kind of funding. This information needs to be taken into account when applying for positions and searching for financial support.

7.4.1 Sources of Funding

The choice of how to fund one's position is of course important in the beginning of doctoral studies. But since the initial position might not cover the whole time until the defense, this choice might re-appear after a while. The longer the doctoral studies take, the more probable it is that a change of the funding source is necessary [2]. In the following sections, we discuss several kinds of funding sources together with possibilities how to acquire them. Note, however, that we do not distinguish between full and part-time positions (according to [2], 60% of the doctoral candidates in computer science and electrical engineering with an assistant position are full employed, while the remaining 40% are part-time employed).

Funding by a University as Part of Advisor's Contract

The classical position for doctoral candidates in Germany is associated directly to the advising professor, who has a certain amount of funding for research assistants in his or her contract. A study from 2012 reported that approximately 27% of all doctoral candidates in computer science and electrical engineering had such a position [2]. Some professors publish their vacant assistant positions and let candidates from all over the world apply for these positions. Others select promising candidates among their master students and offer a position directly to them when they are about to finish their master studies. While some other countries impose strict rules on the application process, professors in Germany have almost all freedom to choose any candidates they deem to be suitable and just hire them. In other words, the application process for a doctoral position is to a great extent not formalized [2]. Positions based on the advisor's contract usually have a duration of one or two years, but can repeatedly be extended for another such period up to a maximum of six years in total (usually without additional formal applications). The limit of six years is due to German law, which allows a duration of at most six years for a non-permanent position for scientific employees (§2 WissZeitVG). In theory, a professor could extend such a position beyond six years, but then the position must be made permanent. This does not happen in practice.

University Projects

Some universities offer positions for doctoral candidates which are not directly connected to a certain advisor, but to a topic. Sometimes such projects are joint with other universities, preferably from other countries to increase the international experience of the doctoral candidate. Still, in most cases a certain advisor or a group of possible advisors will select suitable candidates for such positions. However, these positions are usually announced publicly and one needs to apply for them in a formal application process. Project positions usually have

a fixed duration between one and three years. Some of them can be extended after this period is over, but they usually have a maximum total duration of no more than five years (most projects only allow a duration of three years in total). For extensions, additional applications are necessary.

Non-University Research Projects

According to [2], 27% of the doctoral candidates in computer science and electrical engineering had a position based on a research project not funded by the home university. Foundations or companies offer money for conducting research in a certain area or even on a clearly specified topic. Some positions are offered directly by the funding organizations and the doctoral candidate additionally needs to find a suitable and willing advisor from any university. However, most of these positions are acquired by professors, who select suitable candidates as for positions included in their contract. As for university projects, such positions usually have a fixed duration of up to three years. Extensions (if possible at all) require additional applications.

Scholarships

Instead of regular positions as research assistants, one can also apply for a scholarship. Scholarships are usually associated with the doctoral candidate instead of the advisor and allow for a switch of the advisor without losing the scholarship. However, most scholarships offer less money (net, cf. [2]) than a full position and since the money is paid directly without taxes or insurances, one does not obtain any pension claim and additionally needs to pay for health insurance. Scholarships can be offered by all kinds of institutions or even single persons. They usually require a formal application, but especially scholarships from single persons might be offered under individual conditions. In [2], 27% of all doctoral students in computer science and electrical engineering had a scholarship as their primary funding source.

Individual Funding

It is also possible to complete the doctoral studies without a position as a research assistant. Then the doctoral candidate needs to find funding on his or her own. This can be done, e.g., by working for a company while studying in one's free time or by living from savings. If the doctoral studies are taken without a research assistant position, one only needs to find professors willing to review the thesis and the defense. In [2], 15% of the doctoral candidates in Germany are reported to fund themselves by external jobs.

7.4.2 Additional Duties Associated to Funding Sources

Depending on the source of funding, several kinds of duties might apply to doctoral candidates in addition to their research. We list the most common ones and discuss which funding sources are usually associated with these kinds of duties.

Teaching

Teaching is the most common additional duty for a doctoral candidate. It may include the creation of exercises and exams for lectures, giving lectures in Vertretung, presenting solutions of exercises and exams, and advising students on seminars and practicums as well as bachelor and master theses. Most positions based on the advisor's contract include a certain amount of teaching (up to 50% of the working time, see also [2]). While most other funding sources do not officially include teaching obligations (or even explicitly disallow them), some professors distribute their teaching obligations evenly among their assistants independently of the individual funding sources. This might be a condition by the professor on advising the thesis even for doctoral candidates who fund themselves. A doctoral candidate should analyze the situation carefully, since refusal to do teaching might lead to problems between doctoral researcher and advisor; since in Germany, the doctoral candidate highly depends on the advisor (cf. [1]).

Research Proposals

Research proposals are used to apply for funding of research. They are the most common way for a professor to increase the financial resources beyond those guaranteed in the contract. In particular, they are used to obtain funding for assistant positions. Thus, a doctoral candidate might be asked to support the writing of research proposals to get the funding of his or her future position. In a research proposal, the topic and goal of the research needs to be specified along with ways to measure the success of the research work. If accepted, these parts of the proposal lead to further duties (see the following sections), including reports on the state and success of the research work.

Research on Specific Topics

If the position is based on a research proposal, the doctoral candidate might be obliged to focus his or her research on a specific topic, even if this does not belong to the research interests of the doctoral candidate and cannot be used for the thesis. The same might also happen if the advisor wants the doctoral candidate to work on a certain topic. This kind of duty might occur with all kinds of positions, but positions based on projects or advisor's contracts are more likely to involve research on specific topics outside the interests of the doctoral candidate.

Publications

Similar to research on specific topics, another obligation by research projects might be to publish at certain venues or in certain journals, even if the research interests of the doctoral candidate and those of the venue or journal have nothing in common. Again, this duty might also be imposed on the doctoral candidate by the advisor. However, as far as we know it is a rare occasion that a doctoral candidate is obliged to publish at a certain venue or journal where there is no overlap in research interests.

Patents and Products

The need to register a patent or develop a commercial product can also be a part of research projects, especially ones funded by companies. But if the advisor is involved or interested in patents or products, also other kinds of positions may include this duty. Still, it is not very common that doctoral candidates are asked to register a patent or develop a product on their own (to the best of our knowledge, this kind of duty is not very common among doctoral candidates at all). Usually, if such an obligation exists, it is part of a bigger project where many people are involved and where the doctoral candidate is only responsible for a small part of the overall process.

Non-Research Work for Funding Organization

External research projects might include a certain amount of work for the funding organization which is not related to research at all. While not necessarily being useful for the doctoral studies themselves, practical experiences gained by this kind of work might still be very beneficial for the future career. However, this kind of obligation is not very common among doctoral candidates.

Administration

Virtually all assistant positions contain a certain amount of administration work. Usually, administrative tasks are not fixed in contracts, but are just part of the daily work since most research groups organize many things like their IT infrastructure themselves. Further examples of administration work done by assistants are hiring students, local organization of workshops or conferences, and publication of technical reports.

Travel

An important aspect of research is the dissemination of results. This usually includes presentations at renowned conferences taking place all over the world. In particular, to publish a paper at a conference, at least one author is required to attend the conference and give a talk about the paper. Thus, traveling is also a very common obligation for doctoral candidates. In addition to giving presentations, there might also be obligations for research visits or internships at certain institutions or organizations. Since it is often the case that some people in a research organization like to travel, this duty can sometimes be passed to someone else, e.g., by letting co-authors give the talks at conferences. This is not the case for positions which include research visits or internships as obligations.

Courses

Finally, some positions are associated with a certain curriculum of courses to pass by the doctoral candidate (sometimes in addition to courses being part of the doctoral studies). This kind of obligation is most likely to be associated with scholarships.

7.4.3 Benefits of Funding Sources

Some sources of funding offer certain benefits in addition to the salary of the researcher. We discuss some common benefits and what kinds of funding sources usually offer them.

Attending Workshops

Most workshops serve as informal platforms to exchange and discuss ideas and work in progress. Especially for young doctoral candidates, they offer the possibility to present their current work and get feedback independent of their advisors. While the costs for attending workshops often need to be covered by the doctoral candidates themselves, some positions (mostly scholarships) include a certain amount of funding for visiting workshops.

Attending Conferences

In contrast to workshops, presenting a paper at a conference might be mandatory for the publication of that paper. Hence, attendance of conferences where a paper is to be presented is usually funded (except for individual funding sources). However, it might be the case that the advisor attends the conference instead of the doctoral candidate. Attending a conference without the need of giving a presentation is similar to the funding of attending workshops.

Research Visits

To establish a collaboration or to disseminate research results, it can be useful to visit other research institutions. Whether such research visits are covered by the funding source of the doctoral candidate highly depends on the importance of the respective visit in the eyes of the funding organization and/or the advisor. For scholarships it is usually easier to obtain funding for research visits while doctoral candidates on positions based on their advisor's contract need more convincing arguments to get their visits funded.

Internships

Another beneficial experience for a doctoral candidate is an internship at another research institution or a company. Especially projects funded by organizations also offering internships may include an internship as part of the project. Apart from such projects, many positions offer the possibility to be paused for the duration of an internship and continued thereafter. Since most internships include a salary, this is a general possibility for all kinds of positions.

Studying Abroad

International experience is an important aspect in the curriculum vitae of a doctoral candidate. Thus, studying abroad can be beneficial, also to broaden the horizon by learning about topics not taught at the home university. Some scholarships and university projects include a certain time of studying abroad as part of the funded position.

Inviting Researchers

Like visiting other research institutions, inviting other researchers to their home university might facilitate collaborations for doctoral candidates. Depending on the prestige of the invited researcher, such invitations might be funded by the university or by the funding source of the doctoral candidate.

Research/Working Material

Finally, some funding sources offer additional money dedicated to working material such as IT infrastructure or prices for participants of surveys. While most positions are equipped with working material like paper and ordinary IT infrastructure, special material like high-performance computers usually need additional funding. In research proposals, one can also apply for such kind of funding.

7.5 The Main Goal: The Dissertation

The dissertation is the written document that shows the scientific achievements of the candidate as a whole and together with the examination concludes the doctoral studies. In this section we briefly highlight some aspects of a dissertation.

7.5.1 Shaping a topic

It is often believed among undergraduate students that the topic of a doctoral research study is precisely fixed beforehand. The candidate is then given some years of time to "execute" the required studies in order to meet his or her research goal. In reality we see that the way to the doctorate is less stringent and in fact no one can really know what exactly will be the outcome of the studies and what questions are feasible to answer. In other terms, if you could precisely plan research then its results would not be really new and thus the whole project would be rather useless.

A candidate typically will agree on a research area with his or her supervisor and sometimes formulate a *research statement* where currently open problems are stated and the intention of their solution is formulated. At the very beginning, a candidate will lack the overview of the field and the related fields of research. But as the candidate dives further into the research field, new or related questions will emerge along with a feeling for which questions might be particularly interesting. The successful or failed attempts at answering them will guide the direction of further work. Further influence on the research will come from the exchange of ideas with other researchers. Occasions for such interactions usually occur when attending talks, workshops, conferences, summer schools or going away for research stays. In essence the search for the dissertation topic is an evolving process which eventually converges. Over the years the candidate becomes an expert in a narrow domain of research in which he was able to contribute his work. Finally, when the dissertation has to be written one finds that the rather broad formulation of the original research statement can be refined to a precise dissertation topic.

7.5.2 Doctoral thesis

Usually a doctoral thesis in computer science will have the size of a book with up to 200 pages. It will clearly state what are the scientific contributions that the author has made to the research community. Moreover the thesis is expected to be self-contained which means that it has a general introduction to the topic, surveys related work and recapitulates all notions and theories used by the author which go beyond common – i.e. an undergraduate's – knowledge. Another aspect is that it is consistent. This means that even though the individual publications of the author may use different terminology and notation, in the thesis they have to be unified. Finally, there are no strict limitations on the length of the dissertation which allows for more detailed treatment of the subject and more examples than the papers which the thesis is based on. Of course, one may also discuss unpublished work in the scope of his thesis and give an outlook on new interesting problems that emerge from the researcher's work.

7.5.3 Thesis by publication

As an alternative to the self-contained dissertation thesis there is the so-called *thesis by publication* or *cumulative thesis*. There the doctoral student is "only" required to write an introduction chapter (around 50 pages) that give the reader a motivation for the research topic and an overview of the challenges and solutions attempts taken by the candidate. Then at least three of the strongest or most recent publications are attached as they are to this introduction. Particularly in Anglo-Saxon countries candidates join a PhD programme which often is fixed a priori to three or three and a half years. There the cumulative thesis allows the candidate to write up quickly and finish in time. However in Germany the standard, self-contained dissertation is the norm.

7.6 References and Links

[1] Kalle Hauss, Marc Kaulisch, and Stefan Hornbostel. „Neue Strukturen gleich bessere Betreuung? Eine erste Bestandsaufnahme aus dem Promovierendenpanel ProFile". In: *Promovieren heute. Zur Entwicklung der deutschen Doktorandenausbildung im europäischen Hochschulraum*. Ed. by Margret Wintermantel. Hamburg: Edition Körber-Stiftung, 2010, pp. 110–124.

[2] Kalle Hauss et al. *Promovierende im Profil: Wege, Strukturen und Rahmenbedingungen von Promotionen in Deutschland. Ergebnisse aus dem ProFile-Promovierendenpanel*. Tech. rep. iFQ-Working Paper No. 13. Institut für Forschungsinformation und Qualitätssicherung, Nov. 2012. URL: `http://www.forschungsinfo.de/Publikationen/Download/working_paper_13_2012.pdf`.