10 Years ECCC

A Digital Library Being Used Worldwide

Volker Klotz, Christoph Meinel
University of Trier, Germany

10 Years ECCC
A Digital Library Being Used Worldwide

Volker Klotz, Christoph Meinel
University of Trier, 54286 Trier, Germany

Summary

Over the last couple of years, many scientists have discussed the advantages and possibilities which the new media like Internet and WWW offer for the scientific electronic publishing. [1] As a matter of fact, the requirements of using new information technology cannot always be compared to other applications when it comes to scientific publications. For example, the scientific results must not only be published quickly, but also the used technology must hold up to the traditional methods with regard to quality assurance or quotability. These requirements are being met by a survey process and a paper-bound archivable publishing procedure. However, these procedures are very time consuming and often cause a delay of the publication of latest scientific results which is problematic especially in informatics with its rapid growth of essential new knowledge. By means of the new media, the "Electronic Colloquium on Computational Complexity – ECCC" attempts to bridge this gap between the traditional requirements of scientific publishing – quality assurance and quotability – with the desire for quicker and vaster distribution of scientific works – here within the field of Computational Complexity.

ECCC [2] is a digital library offering a fully equipped virtual workspace to scientist from the field of Computational Complexity on which current research papers on Computational Complexity, which have been carefully selected by a renowned international Scientific Board, are being made available. In doing so, ECCC offers – similar to a real "colloquium" - the possibility for discussion so that publications are being reviewed publicly and, thus, "live on". Furthermore, based on a strict version documentation and version control, ECCC offers a revision mechanism allowing authors to constantly submit improved versions of their papers without the bibliography consistency of published papers going astray.

This report describes the well-established compromise made 10 years ago between the direct electronic publication by the author (without quality control and bibliographical stability) and a conventional (connected with long waiting periods) publication in a scientific journal or on a conference; it also describes the experiences with ECCC so far.

1. Introduction

The invention of the World Wide Web (WWW) offered completely new possibilities of quicker access to scientifically relevant information and the faster distribution of new results also to the Scientific Community [3]. At the same time, the WWW is an unstructured collection of materials which makes the search for relevant information on a certain subject even more difficult. The person looking for the information often has the choice between quickly and easily accessible articles and information sources which are difficult to access, however, might be of higher and more reliable quality. The publications from the first mentioned, unstructured sources are not filtered and, therefore, unsecured with regard to their quality; publications from the second source often stand out negatively due to the long waiting period between submission of the article and its publication, comparable to traditional journals.

---

[1] Zentrum für wissenschaftliches elektronisches Publizieren (WEP), Universität Trier. [On-line material]. Available: http://wep.uni-trier.de

[2] Electronic Colloquium on Computational Complexity (ECCC). [On-line material]. Available: http://www.eccc.uni-trier.de/eccc/.

[3] Christoph Meinel. Elektronisches Publizieren im World Wide Web – Eine neue Dimension in der wissenschaftlichen Kommunikation. In *Forschung und Lehre*, pages 122 – 125.

The "Electronic Colloquium on Computational Complexity – ECCC" [4, 5, 6] is a digital library which was specified and implemented following long discussions in 1994. ECCC is a compromise of both extremes of typical scientific publication methods, and combines the advantages of both. A large group (working voluntarily and without any payment) of internationally renowned scientist – Scientific Board of ECCC – secures by means of a newly developed selection process – the Screening – the content-related quality and topical relevance of publications whereat, thanks to the scientists involved, the publication standard of ECCC is very high. The usage of the word "Screening" shall, thus, rule out the possibility of confusion with the applied survey process of scientific journals and conferences. Submissions which have been accepted for publication by the Scientific Board are being published as a ECCC-report. As the screening of a submitted paper is being done within the very short period of two months at the longest, and a submitted paper after its acceptance is available in the form of a ECCC-report directly via the internet, the ECCC meets the need for quality-assured scientific publications within a short period.

Over the last 10 years, several electronic scientific journals were published on (and disappeared from) the internet [7] whereas most have merely changed the user interface but not the filter and acceptance process towards traditional journals. Thereby, they could make available to the user an easier and easily accessible interface, however, did not change the core for the long waiting periods for publications – the reviewing process.

However, in ECCC not only the filter process was changed, but also new possibilities of electronic communication were used. Thus, ECCC offers a discussion mechanism for the individual ECCC-articles and gives authors the possibility to publish updated or improved versions of an already published version. Explicit report numbers and mechanisms for the version control, thus, grant the bibliographic consistency of the papers.

Furthermore, ECCC offers a web portal for the international community of complexity theorists including distinct electronically books, multimedia tele-readings (online or on demand), scripts, theses, professorial dissertations and many other interesting links.

2. **Traditional Methods of Scientific Publications**
Scientific processes are primarily being pushed by the publication of scientific results. Further research is being done on the basis of published results with the publication granting the author's rights. Due to various requirements, independent scientific publication forms have emerged over time serving different purposes. The three main forms are:

Journals: Complete, extensive works are being published in scientific journals. The submitted papers are being reviewed very thoroughly regarding their thematic and content-related quality and are then being published – usually after an iteration of revisions. This reviewing process is very time-consuming and inevitably leads to a delay of the publication of current results. With the publication of papers, scientific journals also offer an accepted and approved method of archiving and distinct naming of scientific results whereby the articles can be *referenced* clearly and unmistakably.

---

4 Jochen Bern, Carsten, and Christoph Meinel. ECC – ein elektronisches Kolloquium in Internet. In *Informatik-Spektrum*, pages 230-231, 1996.
5 Jochen Bern, Carsten Damm, and Christoph Meinel. The Electronic Colloquium on Computational Complexity (ECCC): A digital library in use. In *Proceedings of ECDL*, 1997.
6 Jochen Bern, Christoph Meinel, and Harald Sack. Electronic Colloquia: Idea and Practice. In *International Symposium of SIGDOC*, 1998.
7 Digital Library Federation. [On-line material]. Available: http://www.diglib.org/.

Preprints: Especially in technology-driven science areas like IT, important results are being achieved which can quickly go out of date due to the further enhancement of technology. In order to prevent the delay of typical publication methods, more and more authors and their employing institutions tend to publish their results as preprints ("Technical Reports", "Research Reports" or the like), and to distribute it in hardcopy form and / or electronically via the internet. Thus, the publication is often only controlled by the author or his / her institution which on the one hand make a *direct distribution* possible, on the other hand, however, due to simple correction possibilities and missing external control, it becomes almost impossible to distinctly use such publications as reference over a longer period.

Conferences: Conferences also offer the possibility to publish research results within the conference proceedings. Conferences can, therefore, be seen as a compromise between the two already mentioned publication forms which is one of the reasons for their growing significance. Submissions to a conference are being reviewed and selected by a program committee. In contrast to a journal, the filtering process is restricted by a number of specified general regulations. Typically, the review and selection has to be finalized within 2 – 4 months. At the end of this process, some of the submitted papers are then being selected for a presentation and publication in the conference proceedings.

For an assessment of the advantages and disadvantages of the aforementioned typical publication methods, the following three characteristics are relevant: the *quality assurance*, the *timeliness* and the *quotability* of published papers.

Quality assurance includes the content-related quality as well as the topical relevant of a paper. Especially in times of increasing quantity of publications, the reader becomes more and more dependent on a reliable and professional pre-sorting of information sources. The most common and proved method of quality assurance is a professional's review. Here, scientific journals are representative: Each submission is being assigned to several reviewers – all of them being specialists on the relevant subject – who then verify the paper with regard to originality, correctness and its general quality. The reviewers must be given sufficient time which can lead to the aforementioned delay of publications. On the other hand, due to the high quality standard, a publication in a renowned journal can be seen as the biggest acknowledgement within the scientific community. The review process for conferences is being organized quite similar, however, with stricter deadlines which in return leads to the fact that not the same level of quality assurance (as for journals) can be reached. Preprints are usually being published without any further quality filter by the scientific community.

Journals are being considered as the most respected publication form, however, the published articles do not always reflect the current status of research. In reputable journals, the waiting period between submission of a paper and its publication can take years which has a negative effect on the timeliness of the research results. This also applies to conferences which have a shorter waiting period, however, the waiting period is still at least 6 months. The timeliness of results is of specific interest particularly in rapidly developing research areas like IT; that is why preprints are often used as the publication medium as they reflect the status of research due to their direct distribution.

Besides quality assurance and timeliness of the published research results, the aforementioned typical publication methods clearly differ regarding the quotability of the article. This characteristic is the main reason for a paper publication in hard copy. Scientists must be able to reference the used sources clearly and distinctively. Data sources like a preprint server which

are being administered by an individual or an institution, do not always offer this. In order to grant the quotability of a paper, the following characteristics of an information source are essential:

- *Long-term availability*: Secure availability of published information over a long period of time;
- *Clearness*: The possibility of a unique identification of a paper and the immutability of the information over a long period.

Typical journals which are being released in hardcopy, completely fulfil the characteristics of papers' quotability. The long-term availability and the clearness of journal articles are being granted by the archiving of the journals in libraries and the nature of printed word. Printed paper cannot be amended and remains legible even after several 100 years. The specific description of a journal, its number as well as serial number allow for the distinct identification of the published articles. Printed conference articles offer a comparable, however, normally slightly lower distribution of papers, a lower quotability. Electronic preprints offer only a very limited quotability while preprints in printed form which are being passed on to interested people, offer a certain level of quotability due to this distribution. At least, the possibility for subsequent changes of the content is limited.

The long-term availability of electronic publications is being influenced by the rapid development of storage technologies as well as the limited time of usage of the relevant software. After just a few years, a computer system and the data formats of electronically publications can be outdated and useless. Not even frequent backups are a way around, and they do not guarantee the long-term availability which is why electronic preprints are particularly unsuitable regarding a well quotable information source.   Developments of emulators of obsolescent computer systems and the establishment of standard formats for electronic publications are one attempt of working against the problem of electronic distribution [8]

3. ECCC – Electronic Colloquium on Computational Complexity
3.1. The Concepts of ECCC

In 1994, a group of leading scientists from the field of Computational Complexity discussed the possibilities to use the developing internet for making an attractive "virtual working space" available to the community. In doing so, both - the possibility of faster distribution of research results, and the assurance that the submitted information meet a certain quality standard and the visitor finds the information (s)he is interested in a filtered form – should be given.   Furthermore, this virtual working space should enable for a continuing public discussion on the published material. As a result of these discussions, ECCC was specified, implemented and launched in 1994.

ECCC is a compromise between the aforementioned typical publication methods and can be classified between the unfiltered scientific publication methods (preprints) and the electronic journals which have to undergo an extensive reviewing process. In order to use the advantages of the individual methods, new concepts were launched which have proven to be successful in the 10 years since the beginning of ECCC.

In order to assure the timeliness of a publication, the time span between the submission of a paper and its publication must be as low as possible. The timeliness, however, must not be

---

[8] Elektronische Publikationen. [On-line material]. Available: http://de.wikipedia.org

realized at the cost of a working filtering of the publications based on thematic and content-related quality. Therefore, compared to the very time-consuming review process in traditional journals, an optimized process of quality assurance must be launched. Before this selection process will be described in more detail, the ECCC structure shall be introduced below:

Upon submission of research papers at the ECCC, there are two statuses: the "waiting status" and the status "published". Upon submission, a paper is automatically put into the waiting status. Articles in this status can only be reviewed by a small group of selected scientists, the Scientific Board of ECCC. Only a member of this board can transfer a paper according to the selection process to the "published" status where the paper will then be available to the public, and can be accessed via the internet by every visitor. The selection process, called "Screening", is being executed as follows:

1. Every week, all board members are being informed by email about the current submissions which are in waiting status, and the members are being asked to screen submissions. Every member of the board can see all submissions and access them accordingly whereas this is done on a voluntary basis and depending on individual interests, the specialization, personal preferences and the available time. After a member has selected a certain paper for screening, its being reserved for him/her exclusively by another email. Compared to journals or conferences, the assignment of papers to the board members is not being done but rather left for the members to decide for themselves.

2. The board member decides whether or not the submission is in line with the required publication standards. It is being verified whether the submission fits the thematical frame of ECCC, whether the results are new and of interest, whether all evidences are included and whether the paper is being presented in a legible and comprehensible form. By email to the ECCC, the scientist can accept the paper, reject it or – in case (s)he does not want to make a decision – release it for other members of the board.

3. If a board member accepts a submission, it will directly and automatically be published as ECCC-report whereas the published paper is being furnished with distinctive bibliographical data (ECCC, year, report no.) in order to make the paper referenceable. As soon as a decision was made on the submitted paper, a message will be sent to the authors either containing the exact number of the report (in case the paper has been accepted), or (in case of refusal) the reasons for the refusal of the submission including the anonymized reasons given by the reviewer.

The entire selection process is being executed by a member of the board which leads to a substantial reduction of time for the filtering process. The experiences from the last 10 years and the acceptance by the scientific community have shown that despite the simplified filtering process, a surprisingly high quality standard can be reached. The success is based on the following factors: On the one hand, the scientific committee consists of more than 40 internationally renowned representatives from the relatively small community of Computational Complexity whereas the majority are already members in many conference committees and editors of relevant journals. Thus, these scientists have a very good overview over the current works and trends in the research area. Another factor for the high-quality standard is that the submissions are being introduced to all member who then in return chose the papers for screening which are most relevant for them according to their experience and personal interest. This organization of the Scientific Board shows that one board member's opinion reflects the opinion of many scientists of the Computational Complexity community.

Another important reason is the following regulation: Submissions which require particular attention due to e.g. sensational statements and, therefore, have to undergo an extensive reviewing process, are not being considered but rather relegated to typical journals as in those cases the implemented selection process is not the appropriate method of quality assurance.

By combining the selection process with the continuous publication of papers (a paper is being published directly after having been accepted), the time between submission and publication can be kept relatively short which can have very positive effect on the timeliness of the paper. If after two months a paper should not have been picked for the screening process by a board member, the submission will be refused due to lack of interest (a clear sign that the paper does not fit the content-related frame of ECCC). In order to prevent that a paper is being refused after two months, the board members are being informed on a regular basis about submissions which are about to be turned down. At first, ECCC started with a deadline of 3 months, however, after a short time, this deadline could be reduced to 2 months without loss of quality.

As submissions are being reviewed by a board member only if his / her time allows, papers are usually being accepted or refused shortly after screening. Only in very few cases does the selection process take several days. In these cases, the scientist is automatically and regularly being reminded by the system that a decision is due. The experience of many years has shown that most submissions are being published after just a few days.

A basic characteristic of the publication method is the quotability of the article. As ECCC-articles are electronically being published via the WWW, some arrangements had to be made to grant the long-term availability and clearness of the paper references:

- A backup of data and system is being run regularly.

- An archive-CD is being compiled every year containing all contributions of the previous year. This CD is being sent to several libraries and all members of the board, and can be purchased at cost by any interested individual. On the occasion of the 10th anniversary of the ECCC, an anniversary CD was compiled containing all (700) contributions of the first 10 years (1994 – 2003).

- All electronic contributions are being printed regularly and paper-archived at several locations, e.g. at the library of the University Trier.

- ECCC was furnished with the ISSN 1433-8092.

Especially by paper-archiving the contributions in libraries, a long-term availability of the ECCC-articles can be granted. In order to be in a position to restart operations within a short period of time after a crash of systems (in the first 10 years, there has been no such crash of ECCC), electronic backups are necessary.

Every accepted article is being furnished with a distinct number which is automatically been added to the electronic paper upon publishing (similar to a water mark). In order to guarantee the distinct quotability of articles, some regulations were met and the following conditions are being connected to a publication. After publication of a paper, it is no longer possible to make any changes of the paper. Modifications or corrections of public papers can only be re-submitted by an additional revision. The revisions are being marked accordingly and are being published in addition to the original paper.

3.2. ECCC as a Colloquium

The electronic platform of ECCC offers the visitor the possibility of discussing current research papers and supports the statement on published papers. The discussion mechanism is implemented similarly to the submission of articles: A discussion thread is being sent in PostScript Format to ECCC where it is being clearly marked und published on the website of the referring report. The discussion is moderated in order to prevent abuse and to guarantee a high standard.

In addition to the discussion mechanism, ECCC offers a revision mechanism. Here, it is possible to continuously publish improved, extended versions of an already published paper. In order to guarantee the clear identification and, hence, the quotability of ECCC-articles, the original paper is not being replaced by a revision, but rather the revision is being published in addition to the article. In order to also make revisions quotable, they are being furnished with a unique revision number which consists of the original report number and a relative counter. The revision mechanism is implemented similar to the submission mechanism with the only difference being that a revision is being published directly – i.e. without additional screening. An additional screening is not necessary as only authors of a published paper are allowed to submit a revision.

3.3. ECCC as Web-Portal

Besides the publication possibility, ECCC offers a web portal for the complexity researchers. Thus, over the years, ECCC became one of the first places for scientists to go when looking for relevant information on Computational Complexity.

Currently, the portal offers links on books, lecture notes, online lectures, summaries on professorial dissertations, diploma and graduate theses. Furthermore, scientists can add the link to their homepage to ECCC provided that this website contains relevant research information. Over the years, this lead to a substantial collection of the researchers' private websites connected to Computational Complexity. A newsletter and various additional information resources, such as conference announcements, library servers, books and journals, complete the ECCC's offer.

3.4. Additional Characteristics of ECCC

As the development of ECCC dates to a very early stage of the WWW, and should always remain accessible also to researchers with lower technology standards, ECCC has a very easy user interface whereas the content is being presented within simple static HTML-pages. Each ECCC-publication has its own HTML-page containing the name of the author(s), the title and the summary of the paper. Should comments or revisions of the paper be available, they are being presented additionally on this page, along with the article. All papers (including comments and revisions) are being offered in PostScript or the popular PDF-format for download. In addition to the static pages, dynamic components, such as keyword search within the author names, in the title or the summary, are being offered.

Authors can submit their papers, comments and revisions via a WWW-interface or by email whereas the papers must be in PostScript-format.

All ECCC-articles (including revisions and comments) are being labelled by a distinct number which – like a water mark – is being included to the published papers. Thus, every paper contains all information in order to be referenced explicitly.

After extensive discussions with authors of ECCC and publishers, it was decided to put no copyright on the ECCC-publications. By doing so, the barrier for submissions to ECCC is being kept as low as possible, and the author is free to transfer the rights to a publisher. If such a situation occurs, the paper must be removed from ECCC – which is, by the way, the only accepted reason for removal of a ECCC-paper. As the removed article is being replaced by the source of the new publication (due to the very likely existing quotability of the source following the transfer of rights), and the number of the article cannot be re-used, the distinct quotability of the paper remains intact.

4. Acceptance and Usage

In its 10 years of existence, ECCC has become a highly-accepted channel for the publication of research results within the Computational Complexity whereas the quality of the papers is high (which was, at first, unexpected), and can be compared to the top conferences and renowned journals in this area. Over the last 10 years, more than 700 articles were published on ECCC. Every day, the ECCC-articles are being accessed approx. 300 times. The following chart will show the accesses in the period October 2001 until mid-April 2004, whereas a database cleaned from crawlers and robots has been used. The second column shows the average daily accesses within one month. The columns 3 and 6 refer to the total accesses within one month. As described before, the published articles can be accessed in two different formats. The columns 4 and 5 show how often the articles were accessed in either PostScript- or PDF-format.

Month	Daily Accesses (Average)	Accesses	PS-Format	PDF-Format	Total Storage (Mbytes)
Apr 04	348	5571	933	4638	1100
Mar 04	292	9070	2106	6964	2700
Feb 04	297	8620	2851	5769	1900
Jan 04	222	6893	2056	4837	1500
Dec 03	255	7930	3442	6043	1600
Nov 03	256	7690	1555	6135	1700
Oct 03	229	7106	2427	4679	1600

Chart 1: Monthly access to articles published in ECCC.

Overall, the ECCC has approx. 3000 hits per day, and transfers approx. 3000 Megabyte data per month.

Chart 2 describes the relation between the "age" of a publication (year of publication) and the access figures. The period 01.01.2004 through 16.03.2004 was examined in which the access on publications of a certain year within the referenced period were calculated; the total figures are shown in the second column. Column 4 shows the average access per article (total access divided by the number of articles which were published in the relevant year).

Publication Year	Total Access (01.01.04 – 16.03.04)	Number of ECCC-Reports	Average Access per Article
1994	939	27	34,8
1995	1296	63	20,6
1996	1418	67	21,1
1997	1339	61	22,0
1998	1666	78	21,4
1999	1267	48	26,4
2000	2275	91	25,0
2001	2881	104	27,7
2002	2160	74	29,2
2003	3382	87	38,9
2004	1457	18	80,9

Chart 2: Accesses to ECCC-articles sorted by publication year. The data derive from the period 01.01.2004 through 16.03.2004.

It shows that the articles become less important the older they get, however, are being searched for with the same level of interest. (The exceptional situation of the launch year 1994 can be explained by the fact that at the beginning of ECCC, the Board Members had published their own important papers in order to show the aspired quality level of the platform). The chart clearly shows the meaning of timeliness of scientific publications.

ECCC is being visited and used by colleagues worldwide. For example, in March 2004 accesses from 94 different countries were registered, whereas users from Germany and the USA lead the statistic. At the same time, ECCC is not only being visited by an international community, but also being used for publication. The following chart shows the number of countries (sorted by year) from which the publications (traceably) are being submitted.

Publication Year	Nations (total)	Reports (total)
1994	8	27
1995	11	63
1996	12	67
1997	13	61
1998	13	78
1999	12	48
2000	16	91
2001	16	104
2002	19	74
2003	19	87

Chart 3: Number of countries from which the publications per year are being submitted.

The chart clearly shows the number of different countries from which the articles are being submitted has grown over the years. For example, in 2002 the authors of ECCC-papers came from 19 different countries, whereas in 1998 only 13 different countries with the same number of ECCC-reports could be counted.

Besides the reports, ECCC offers a mailing list for monthly announcements of new ECCC-publications. At the moment, 435 users have registered for this service. As mentioned before, researchers can register their homepage on ECCC which is being used by 180 scientists from the field of Computational Complexity. The collection of links to the private researchers' site is, by the way, a very popular place for the research community of Computational Complexity and is one of the 10 most visited sites of ECCC.

The Scientific Committee

Miklos Ajtai	*Janos Komlos*	*Vojtech Rödl*
Eric Allender	*Nathan Linial*	*Steven Rudich*
Noga Alon	*Richard Lipton*	*Michael Saks*
Sanjeev Arora	*Michael Luby*	*Claus Schnorr*
David Barrington	*Wolfgang Maass*	*Peter Shor*
Richard Beigel	*Kurt Mehlhorn*	*Michael Sipser*
Mihir Bellare	*Christoph Meinel*	*Madhu Sudan*
Jin-Yi Cai	*Noam Nisan*	*Mario Szegedy*
Merrick Furst	*Moni Naor*	*Ingo Wegener*
Oded Goldreich	*Christos Papadimitriou*	*Avi Widgerson*
Johan Hastad	*Pavel Pudlak*	*Uri Zwick*
Stasys Jukna	*Jaikumar Radhakrishnan*	
Mauricio Karchmer	*Alexander Razborov*	
Marek Karpinski	*Rüdiger Reischuk*	

Acknowledgement