Chemistry Preprints

Henry S. Rzepa

Department of Chemistry, Imperial College of Science, Technology and Medicine, London, SW7 2AY, U.K.

Received November 19, 2001

Scientific preprints have a long history, one that is perhaps forgotten in their more recent and often thought provoking if not controversial electronic incarnations. Certainly, prior to the revolution brought about by commercial and learned society scientific publishing in the early 1950s, the preprint had a substantial role to play. Discoveries were often initially presented in the form of "papers" read to meetings of scientific societies by senior members, who in effect acted as peer reviewers of the articles. During this period, departmental faculty libraries were actually much more of a rarity (it was part of the function of the professor of chemistry to maintain this library as his/her personal property, although one might imagine that access rights might not always have been equally granted to all), and so the importance of personal contacts and exchanged preprints was high.

Scientific publishing grew both exponentially and became much more of a commercial and perhaps less personal process in the second half of the 20th century. With the equally rapid growth and importance of the reviewing and secondary and tertiary publishing processes, issues such as copyright ownership rights and the intrinsic costs of subscribing to and searching of the growing volume of information came to the fore. Authors became more concerned with turnaround times for their publication, whether the peer review processes were the fairest way of selecting for high quality and novel science, and what the impact of their articles might be, not least on their promotion and grant prospects. Almost overnight, starting around 1995, it had become apparent that electronic publishing would certainly augment and perhaps replace conventional paper-based dissemination within a short period (which we now know to have been around 5 years!). Incorporation of the conventional issues of cost, speed, review, and ownership into this electronic framework proceeded at a much slower pace, perhaps reflecting that the dynamics of human interactions often changes more slowly than technology. People, for example, still refer to "the paper" in very much its 19th century rather than its potentially new electronic meaning. The e-publishing revolution brings with it the chance for radical rethinking not only of the communication process but also of the social issues.

Following this introduction, we see a review of one of the fruits of the changes in e-publishing, the first chemistry electronic preprint server (CPS), and how some of the issues above are addressed. In one sense, it represents a return to the original ethos and concept of the scientific article of the 19th century, but in another it might be seen as perpetuating the conventional publishing process, in that worthy articles are expected to rematerialize in more permanent and conventional form elsewhere. Will the preprint server be another catalyst for accelerating change in the very nature of publishing or perhaps an evolution in the commercial models for this information handling? Intrinsically, processing the underlying intimate mixture of data and meaning in the scientific document remains not only almost the exclusive preserve of the human reader but also an error prone, very expensive, and of course often highly creative procedure. In this regard, CPS is surprisingly conservative, perhaps reflecting its authorship and audience. The articles are presented, as with most other journals, in a form (Acrobat) designed predominantly for printing, and the supplemental files when present (less than half of the 20 most accessed articles had any such information) seem rarely to enhance the article in the sense of presenting readily reusable and semantically rich data and information.

From a personal point of view, I would hope that services such as the preprint server move us in the direction of helping humans to do what they are best at, the creative and discovery process, while enabling software agent and other new tools to handle those aspects that can be or should be automated. Murray-Rust has coined the concept of a "datument" as replacing what we currently understand as the basic published object. We are both convinced that technologies and environments which capture these concepts, such as XML, currently represent our best hope for continuing and very much extending the remarkable progress that has already been achieved in a few years. The CPS has yet to manifestly encourage its audience to move in these directions; this is the challenge that we set out here to them and to the community!

CI010389P