nature genetics

Online Methods

Most of our readers access our articles online, in formats that deal well with increasingly complex research methods and the growing requirement for increased precision of citation. These considerations have now led us to publish Methods online.

Starting this month, readers will notice a new section called Online Methods in our Letters, Articles and Technical Reports. Material previously published as Methods and Supplementary Methods is now combined, fully edited and hyperlinked in the new format that will be present on the journal's website and reprints, and can be downloaded in PDF format. Readers of the monthly print journal will now be directed to find the Methods online.

Online Methods allows us to remove print-imposed limitations on numbers of cited references with the benefit that we can enable correct citation of primary research papers for the data, conclusions and methods they contain. Of course, reviews can still be cited to provide readers with background they may need to understand the paper fully, and in cases where the review presents a concept for the first time. However, citing primary sources preserves essential information and provides a more accurate record of attribution and priority.

We are aware that readers' time is valuable, so we will continue to edit and ask that authors remain concise, keeping the Methods within a limit of 1,500 words, excluding references. This limit accommodates the typical content of Methods and Supplementary Methods we previously published. So far, we have found some exceptions that can make the methods section longer than usual and these cases are instructive, giving clues as to what the research paper of the future will need to recognize, database and present.

Acknowledgments, author names and affiliations should be put in the main paper in order to give proper attribution. Collaborators who are not authors should be acknowledged, but if very numerous (thousands), these can be listed in a Supplementary Note. Extended biobank and cohort descriptions are probably best handled by submission to a public database and citation via an accession code. Protocols can be submitted to *Nature Protocols* and cited.

In general, the journal is extending its existing policy of encouraging authors to submit datasets of all kinds to public databases and to cite the dataset explicitly within the paper via a unique accession number. This practice clarifies the semantic structure of the publication and increases the utility of the paper as a research tool. For many data types (for example, DNA sequences and expression profiles), the journal mandates submission as a condition of publication.

Supplementary Information provides a useful temporary solution for relatively unprocessed supporting data, but unless databased, formatted and uniquely identifiable, this data is restricted in its utility, and so does not help the reputation and citation of the associated paper. Authors may not know where to put their unique data because each journal may have seen only a few examples of similar datasets, but in contrast, the curators of public databases may have handled many similar sets and may even be able to accommodate brand new types of data using existing storage and identification systems.

The scientific paper of today is a chimera of conventions established in the era of publication on paper and those enabled by web links. It is evolving toward a mix of natural language and formal concept structures. Unique data identifiers are therefore the building blocks of semantic publications in which sets of data referred to by their accession codes underpin all of the concepts and assertions established by authors. In most cases, explicit correspondence can be established between data multiples and the conceptual constructs that they support, although natural language is needed to expand and explain the formal concepts being established. However, it is not clear that formal and natural language will be able to coexist forever within a unitary text. The concept architecture of the paper may eventually need to be extracted so that the human-readable text does not become overburdened with signifiers. ∎