Reviewer's report

Title: Some methods for blindfolded record linkage

Version: 2 Date: 26 January 2004

Reviewer: Ross Lazarus

Reviewer's report:

General

I enjoyed reading this paper which appears to me to be well written, original and potentially very useful and important. The subject is of general interest because of the difficulties faced by researchers and administrators wishing to maintain the privacy of individuals while adding value by record linkage of large and important public and private data sources. This is a complex social, legal and political topic, no less in Australia than here in the United States. Many potentially important public health, pharmacoepidemiological and clinical research issues can only be effectively tackled through safe, efficient and reliable record linkage. The approach appears to offer a practical way to securely automate a very tedious and insecure manual process for large scale linkage projects.

The paper addresses an important subset of technical problems in record linkage and suggests some additional applications. It is of necessity a highly technical paper and although most readers are likely to be, like me, familiar with only part of the broad range of literature covered, it is very well written and makes the technical issues very clear. The novel application of n-gram scores with layers of PKI and distributed processing appears to me to be very well thought through, pragmatic and offering sufficient security for potential real world usefulness. As the authors indicate, it is not without limitations, but these (computational and data transfer loads and some as yet uncertain loss of accuracy compared with manual methods) are honestly described and seem minor compared with the inherent security problems and costs of manual and other currently available methods.

In addition, the authors have moved beyond the purely theoretical by provided a working proof of concept. Their Febrl project will no doubt prove to be a valuable resource for practical applications as it continues to evolve and grow.

My only regret is that I have nothing to offer the authors in terms of suggestions to improve the work - they've done an excellent job in my view.

Discretionary Revisions (which the author can choose to ignore)

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

What next?: Accept without revision
Level of interest: An article of outstanding merit and interest in its field

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:

None