Steve --

Here are a few quick thoughts about possible referenceware.
[This has many personal opinions, and obviously is not for general consumption, etc. etc.]

First, one should see what makes things more useful on a computer than they would be in printed form. Here are two characteristics:

1. You can search for all items, not just what is in an index. It would be particularly nice if you could search for words that are CLOSE to a given word, but do not necessarily start with the same letters. (e.g. for looking up words whose spelling you do not know)

2. You can take pieces of reference data, and insert them in your work. So, for example, you could take scientific data, and use the numbers in Polymath. Or you could take a quotation, or a form letter, and insert it in your document.

Scientific data:

The American classic for scientific data is:
CRC Handbook of Chemistry and Physics.
This is a mixture of many different sources; it is not (so far as I can tell) in computer-readable form.

The enclosed book:
Kaye & Laby: Tables of Physical and Chemical Constants
is much smaller, but is much better presented, and is probably in computer-readable form.

There are endless more specialized scientific handbooks. Here are three that would probably be easy to get and perhaps have some special scientific interest:

1. Tables of Particle Properties. The definitive list of elementary particles and their properties. Produced by the Particle Data Group at LBL (Berkeley). The main tables are about 40 pages. I believe it is a public domain document.

2. Astronomical data. Positions etc. of astronomical objects. There are many sources for this.

3. DNA Sequence Data. Genetic codes (none yet complete, except for some viruses) for various organisms. If I remember right, the whole database so far is about 10 million base pairs (1 base pair = 2 bits). You might perhaps get the pieces for E. Coli (a very common organism, extensively used in genetic engineering). There is a version of the database maintained at Los Alamos; there are also commercial versions.

Mathematical handbooks:

You might think about getting mathematical handbooks. However, we hope that Polymath will make these largely obsolete.

There are some tables of mathematical formulae that are not yet completely included in Polymath. However, they are not in computer readable form, and they need quite extensive work to make them useful. Over the next few years, I hope that many of these will be converted to Polymath form.

If you had the space, the

Encyclopeadic Dictionary of Mathematics (MIT Press) [4 vols.]
would be very nice.

Dictionaries:

I think it would be very nice to have several dictionaries that get searched together. As well as the obvious foreign language dictionaries, a dictionary of science and technology would be nice. Many of the publishing firms (McGraw-Hill, Academic Press, Chambers, etc.) put out such dictionaries. Most of them are pretty useless.
Of the ones they had in the library (probably, but not definitely, a complete set), the best seemed to be:
McGraw-Hill Dictionary of Scientific and Technical Terms [1800 pp]

Computer programming:

Presumably you will have the language specifications for:
C, PASCAL, Fortran 77, possibly Common LISP, Ada, APL, Scheme;
also PostScript, possibly TeX (or, specifically, LaTeX).

I think you should also have a book with documented code fragments that people could splice into their programs. There are several candidate books. The best I could find is:
K. Jamsa: The C Library (some information enclosed).
You might ask Jon Bentley (CMU) for more suggestions on this.

Numerical computing:

Again, I hope that Polymath will make most of this material obsolete. The obvious book to get in this area (at least if Polymath did not exist) is:
W. Press et al, Numerical Recipes (Cambridge University Press).

General information:

I believe you should definitely have an almanac. It seems that the best one is:
The World Almanac and Book of Facts (Pharos Books, New York, NY).

You might also try to get the Guiness Book of Records.

You asked about census data etc. The relevant book is:
Statistical Abstract of the United States, (US Dept. Commerce, Bureau of Census) [1000 pp].
Most of this book seems amazingly spurious. I would stick to a good almanac: I think it has the data in a much more useful form.

Another general document that reference librarians tell one about is:
US Government Manual.
Again, it looks to me as if most of the useful information from this is in the almanac.

Yet another such document is:
Gale's Dictionary of Associations [2000 pp].

Technical data:

Some additional sources of technical data:

Medical: Medical dictionary (e.g. Blakiston's Pocket Medical Dictionary);

Merck manual; Physcian's Drug Reference...

Electronic: Electrical engineer's handbook; TTL databooks, etc.

Engineering: It would be good to have some engineering handbooks; there does not seem to be a single definitive one.

Document preparation aids:

It would be good to have some form letters etc. on line, for people to import into their documents. [Actually, if people don't have the imagination to write their own letters, I don't feel too much like helping them, but....]

Standard business letters
Standard legal documents
Standard resumes
Possibly Tax forms etc.

[there are many books of these things]

Some related documents:

Chicago Manual of Style
Webster's Secretarial Handbook

Small but useful databases:

Worldwide telephone country/area/city codes + time differences [including exit codes from other countries etc.]

Calender (partially provided by UNIX cal program, but should include worldwide list of holidays etc.)

Units conversions (UNIX has a program, Polymath will have a notebook, to do this).

Some geographical data would be good to have. Some of it will be in the almanac. Maps would be great, but are probably too hard to get. I have a copy of a list of US airports, with the exact longitudes and latitudes -- the coordinates are quite useful in computing distances. Thinking Machines used to have a computerized map of Boston, but they had a very hard time getting it. The USGS does have some maps (I suspect in the form of satellite data) in computer-readable form. There is a nice Macintosh program that plots routes between US cities on interstates.

Academic data:

The most complete work at a sensible level seems to be:
World of Learning: Worldwide directory of universities [1000 pp].

There are more detailed books with faculty at major US universities listed in full -- amusing for some of your potential customers, but probably not particularly useful.

More later .... but now I have to get back to writing Polymath code....

Stephen