
Subject: History of DHS Data Access

Posted by [Trevor-DHS](#) on Wed, 11 May 2016 22:31:52 GMT

[View Forum Message](#) <> [Reply to Message](#)

For those interested in a little bit of the history of DHS Data Access I'm attaching a paper presented exactly 20 years ago (to the day) that describes the plans then to distribute DHS datasets online! See how things have changed over the last 20 years!

File Attachments

1) [Prospects for Distribution of DHS Datasets via the Internet 1996.pdf](#), downloaded 1327 times

Subject: Re: History of DHS Data Access

Posted by [Reduced-For\(u\)m](#) on Thu, 12 May 2016 21:07:52 GMT

[View Forum Message](#) <> [Reply to Message](#)

Trevor,

That was fun! Thanks. Some highlights for the tl;dr crowd:

1. Bernoulli Cartridges

Currently, DHS supports data distribution on diskette, and on seven different types of Bernoulli cartridges for use with PC-compatible computers, and on several magnetic tape formats for mainframe computers. DHS also has the capability of producing custom CD-ROMs for large data requests.

2. The Internet

Additionally, the Internet has appeared as a global means of communication. It provides possibilities unimaginable a few years ago. With communication from the US to Australia taking seconds, and transfers of fairly large quantities of information over the Internet taking just minutes, new ways of distributing data are becoming available.

3. Prescience, or Commitment Device:

DHS is proposing the establishment of a DHS user forum, where any researcher can post messages for DHS staff or other researchers to react to and comment on.

4. #NotARealProblem

The major benefit to charging for datasets, other than to cover some costs, has been to discourage people from requesting data that they are unlikely to make use of. Most researchers who request data are keen to make the fullest use of the data. While there is much less rationale for charging to cover the costs of distribution of datasets over the Internet, the secondary benefit of discouraging users who are not serious about using DHS data should not be ignored.

5. #UsedToBeARealProblem

This is particularly important, in relation to concerns at DHS about bandwidth.

6. Nielson's Law

With a 28.8K baud modem connection using America Online as the Internet network service provider, the file transfer took 11 minutes and 40 seconds. Downloads over a system that is connected directly to the Internet take less time, but downloads over a 9600 baud modem, over noisy telephone lines may require as much as an hour for the same size file. While this file size is about average, there are certain data files that are considerably bigger and may require as much as 12 hours in an extreme situation (assuming that the connection can be maintained for that long).

7. Better Late than Never

Any suggestions on the services DHS hopes to provide and any feedback on the ideas presented in this paper will be appreciated.
