Gutenberg II: 5 million pages and more than 27,000 books available at www.freearchives.ch/demo

Pfaffhausen, 20 May, 2009: In 2004, we needed data for a test database. During the resulting search we became aware of **Project Gutenberg.** This is a project that involves old (copyright-free) books being scanned and re-typeset. The results are published on the internet as free downloads. The project was started back in 1971, but unfortunately the user interface isn't exactly pretty. In the best-case scenario, the texts can be viewed as HTML files, and a free-format text search doesn't exist.

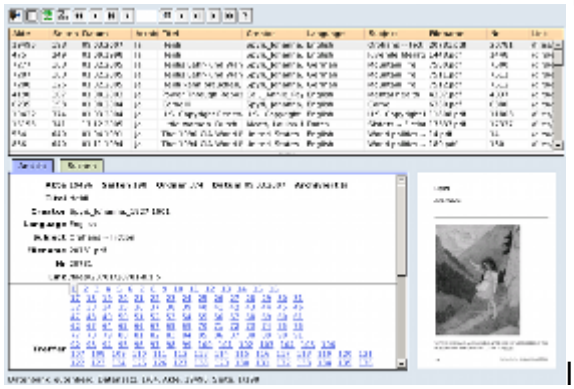
Gutenberg I: An first attempt in 2004

For this reason, in 2004 we created a test database with approx. 2 million pages and made the books available for download. You can find more information about the project in the **position paper from that time.** The results at the time were somewhat conflicting. While the books were received well and happily downloaded, more active collaboration among the community did not materialise.



Anyway, the test archive later served as a valuable test environment during further development of the ArchivistaBox. Unfortunately (as the ArchivistaBox system was a priority), we never migrated the then-existing Gutenberg database to the final ArchivistaBox. Furthermore, the solution at the time had the disadvantage that only the text from the books (without the illustrations) was published.

The freearchives.ch Association implements Gutenberg II



In 2006, I started the association freearchives.ch as one of its co-founders. The association's purpose is to encourage the archiving of long-lasting data. Quite a number of members have now joined the association over the last two years and it has so far supported several OpenSource projects (e.g., exactimage).

Thanks to a lot of free labour, the association freearchives.ch can now present a completely new version, i.e., Gutenberg II, which is located on the association's computer. Thus, you can now find more than 27,000 books with approx. 5 million pages at **freearchives.ch/demo.** In a time when the market for digital books is almost exclusively dominated by one player, 27,000 books may not

appear to be particularly outstanding. However, I'd like to list the following benefits of **freearchives.ch/demo.**

This is why freearchives.ch/demo is an open solution that is free of charge

a) The solution is based on OpenSource. The site **freearchives.ch/demo** uses a standard ArchivistaBox. I think this point is neglected nowadays. While the large players repeatedly assert how important OpenSource is to them, they rarely talk about publishing their own sources. But how is a solution supposed to function in the long term if its availability depends precisely on one person or legal entity and/or the stockholders behind such an entity?

b) The solution runs on a single small box computer. In an age when data centres are mushrooming almost everywhere, this may not appear to be very impressive. But if we consider the effort required to keep data centres and/or clusters alive, it seems that a simple method of operation providing a long-term solution is of crucial importance.

c) Ad-free content: indeed, plenty of stuff is available for free on the internet nowadays, including the constant advertising. But if you look more closely, you will quickly discover that free doesn't really mean free. What does non-commercial content mean exactly, and why can I only obtain the content as an individual? All these restrictions shouldn't exist, and don't have to exist. Naturally, **freearchives.ch/demo** will also cause some small costs but these are already covered for the next few years by existing membership fees.

So, we hope you enjoy Gutenberg II. As mentioned, instead of having to suffer disruptive ad flashes, you can read the books online entirely without plugins. Or, let's be honest, if you really want to read a book, you'll probably print it out offline as a PDF file. Regarding this, we have a printing tip: the books have been created in such a way that you get the best font size if you reduce them from A4 to A5 format.

And one more thing: if you want to receive the data on a USB hard drive, we invite you to become a **member of the freearchives.ch Association.** As soon as you receive the membership confirmation (which can take a little while – up to a few weeks) you can obtain all the data on a USB hard drive (>400GB). Or, even better, you can help the freearchives.ch Association to expand the online archive.