

**Digital editions of canonic texts:
do such projects affect the canon?**

Péter Hajdu

Hungarian Academy of Sciences

pethajdu@gmail.com

The existence of a complete critical edition of an oeuvre used to be a reliable indicator of the given author's central canonical position, even if critical editions are mostly used by professionals or at least students of literature. Moreover a critical edition is supposed to be the canonical (or the most canonical) version of the text. In the digital age, however, these two aspects of canonization (that of an author and that of a text variation) seem to be separated. A text that cannot be found on the web hardly has the right to count on a canonical position today. The grand majority of my students simply refuses to go to the library and read books, and if they come to a class without having read that I had asked them to read they think it is a legitimate excuse that it could not be found on the net. And the suggested solution for copyright problems for example is that I should scan everything and make accessible in the university's closed network (co-space). What cannot appear on the screen does not seem to exist any more.

Digitalisation and on-line publication of the written heritage therefore seems simultaneously an attempt to make cultural goods accessible to everyone in the manner of enlightenment, and a conservative move to save the canonised culture. What is ironic in the latter aspect is that the transcription into the new medium means now to use a much less durable carrier. It is astonishing to see how fast digital data carriers age. Do you remember those 18" flexible floppy discs or zip drives? They are unreadable now, since nobody has hardware for them. And something similar can be said about the formats.¹ The problem of endurance does not seem so obvious with on line publication, but that is only because of the false impression of the net's immateriality, since the data are still carried by physically existing tools, which are not eternal at all.

Nevertheless some people think that digitisation is the only way to conserve the cultural heritage. A major scholar in my country, who was a pioneer of using computer techniques in literary research and text edition, is propagating now a plan that the "whole" Hungarian cultural heritage should be digitised in a portable format and every member of the national community should be given a copy. All the Hungarian books along with an "international minimum" should be in the pocket of every Hungarian.² Conservation of the cultural heritage and democratisation of the access to cultural goods go hand in hand in such a plan. And even if it is

¹ I could only briefly refer to a much discussed topic. Cf. *Guidelines for the Preservation of Digital Heritage*. 2003. <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>. 2003.

² Horváth, Iván, 2010, "Számítógép és irodalomtudomány" *Alföld* 6(3): 18–19 (9–27). Horváth not only advertises such a plan, but also described it officially as a patent in 2004.

emphasised that the entire cultural heritage should be conserved, canonisation may play a role. It is obvious in the selection of “international minimum”, but it also can haunt as the question of what is “cultural”.

An old fashioned critical edition, however, has a sophisticated technique to promote a canonised version of the text, and the attitude is double-faced. A text exists in versions—at least that was the situation before the digital age. A critical edition displays all the varieties. Some people regard critical editions strongly canonical, because they create a main text and subordinate the varieties exiling them into the footnotes, the *apparatus criticus*. Nevertheless it depends on the reader’s decision; of course many readers simply disregard the variations and accept the main text as the only one. And publishers of “popular editions” are supposed to accept the main text as the basis of proliferation. The professionals had done the job and produced “the best text”. But the critical edition still can be read as offering all the varieties. Variations are present and they challenge the exclusive validity of the main text. A reader may or should keep in mind that the main text is no more than the editor’s personal suggestion, which should be checked continuously through a comparison with the other options contained in the notes. If a reader chooses this strategy, they will regard the priority of the main text as a false impression created by print limitations; a book displaying all the varieties equally could be very big, heavy and expensive. Theoretically there are no such limitations for on-line editions.

However, I cannot say that the digital age has brought with it more democratic editorial strategies. In the contrary. Data bases offering huge bodies of canonical texts simply do not hint at the existence of variations. They tend to heavily rely on the one text/one version fiction. And in addition, those versions tend to be old. Let me first refer to the Persius project publishing ancient classics. Obviously for copyright reasons, Persius publishes more than 100 years old text editions, but (probably for simplicity's sake) without critical apparatus. Textological problems and debates are not referred to at all.

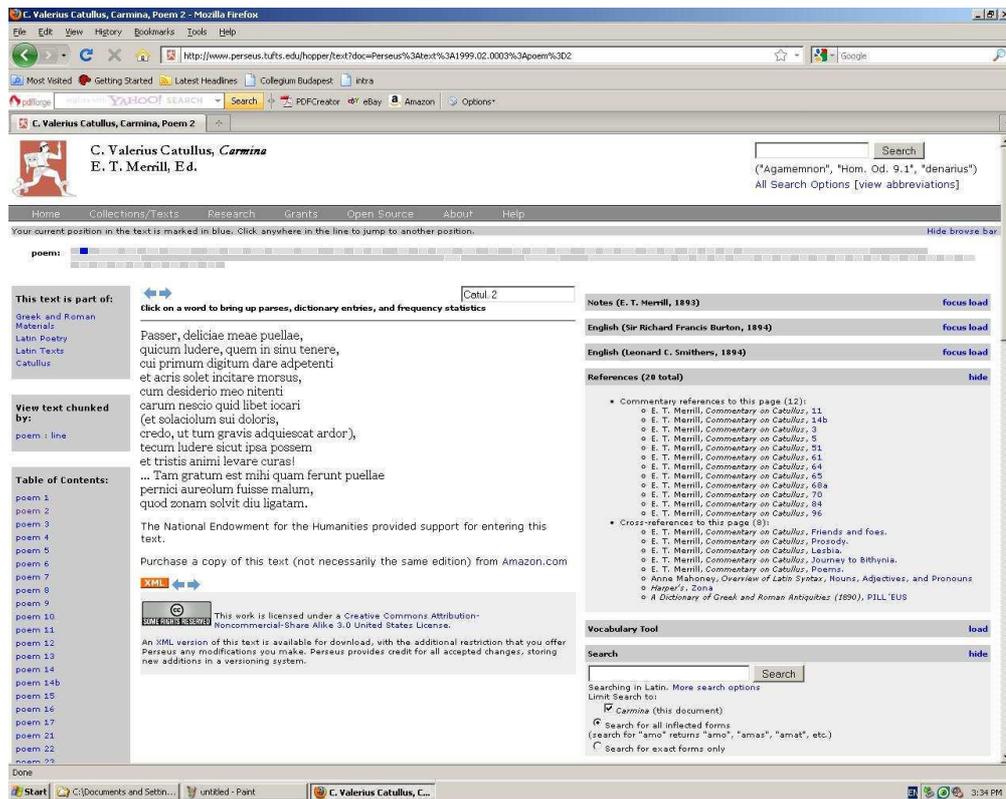


Figure 1

The conservative aim of saving the cultural heritage and the civilising aim of providing free on-line access to it go hand in hand in the project MEK (Hungarian Electronic Library), which works as part of Hungary's

national library. The national library collects all the books published in Hungary (actually a law enforces the publishers to submit some free copies of every book), and this electronic department does its best to make accessible a digitised on-line version of as big proportion of the textual heritage as possible. A canonising intention seems neither evitable nor to be looked for. At least when it goes for selecting the texts. As for the canonised versions, the situation is more complicated. But actually the history of *MEK* (see <http://www.mek.oszk.hu/html/torteneteng.html>) suggest a dynamic relation towards canonization.

In the spring of 1994 a few enthusiastic librarians came forward with the idea of a future Electronic Library, and the first test-collections were established on the gopher servers of the University of Miskolc and the University of Economics in Budapest. In 1995 the actual work started in the context of National Information Infrastructure Development Program. As it seems, a collection of national textual heritage was originally an intellectual challenge for information technology. In the second half of the 1990s the team inevitably created the collection development policies: "the MEK collects resources concerning Hungary or the Central European region, in the fields of culture, education and academic research. The collection consists of text-based resources, but does not exclude other formats such as maps, sheet music etc. The Hungarian Electronic Library also collects periodicals and journals, and links to other relevant resources, services and documents concerning libraries." This idea of collecting everything about Hungary is rather similar to the basic mission

of a national library, therefore it was just logical that the national library decided to provide a new home for the *MEK* project in 1999. *MEK* became a department within the national library. This, however, does not result in a strong canonizing position, since "Technically a civil movement has formed around it, due to the fact that anyone, even with the simplest of methods and means can contribute to the development of the collections: individuals can offer digitized works (their own or by other authors) with due respect to copyright-issues. Also, users can make propositions concerning the archiving of material found on other servers." This opportunity that practically anyone can offer texts for the *MEK* project means that it is like a battlefield of competing canons now. The *MEK* team do their best to guarantee a quality from the viewpoint of editorial issues, like typos, spelling etc. But their situation is not easy.

My main example to be discussed here is the case of a major Hungarian prose writer, Kálmán Mikszáth (1847–1910). A complete critical edition of his oeuvre started being published in the 1950s. This gesture suggests a central canonical position, since he was one of the two prose writers that were given such an honour in Hungary. The enterprise, however, was really challenging. As far as we can judge, the plan was to publish his writings in 100 volumes. Nice books with critical text, textual apparatus and commentary. By 1999, when I got involved in the critical edition, only 76 volumes had been published. The first one appeared in 1956 (Volume 1, the first volume of the sub-series "Novels"), the last until then in 1992 (Volume 86, volume 36 in the sub-series "Articles and

Sketches”), which meant that the gigantic series contained all the novels and the complete available correspondence, but the short stories just until 1892 (except the year 1888), and the non-fiction until 1897.

Hungary is a small country, and there is actually one enterprise that tries to sell CD-ROMs with texts or with digital versions of older, expensive books or series. It is called Arcanum, and one of its first publications was *The complete work by Kálmán Mikszáth*, which is a funny title, since that CD-ROM was based on the incomplete critical edition, and it contained only the material published in it until 1997. Decades of shorter writing are missing, but they thought it easier to use the title of the *Complete Work* still incomplete. But this is only the first half of the problem. Scanners and text recognising programs were not satisfactory that time. The CD-ROM was produced in a rather conservative (or backward) way, which Arcanum stopped using years ago. The main text of the critical edition was digitised through typing then a printed copy was proofread on the basis of the original books. No care about text variations or commentary. The result is a plain text full of typos, which is, however, very useful if you need a complex search.

After selling as many CD-ROMs as the limited Hungarian market seemed to need Arcanum allowed the text to MEK. The text is less useful without Arcanum’s special research engine, but it can be downloaded for free. Parts of this corpus appear on the web as belonging to *The complete work by Kálmán Mikszáth* (even if it is incomplete), and a logo suggests that the text is identical with that of the critical edition, although it is not.

Users need to do some investigation to realize that it is the Arcanum text, full of typos.

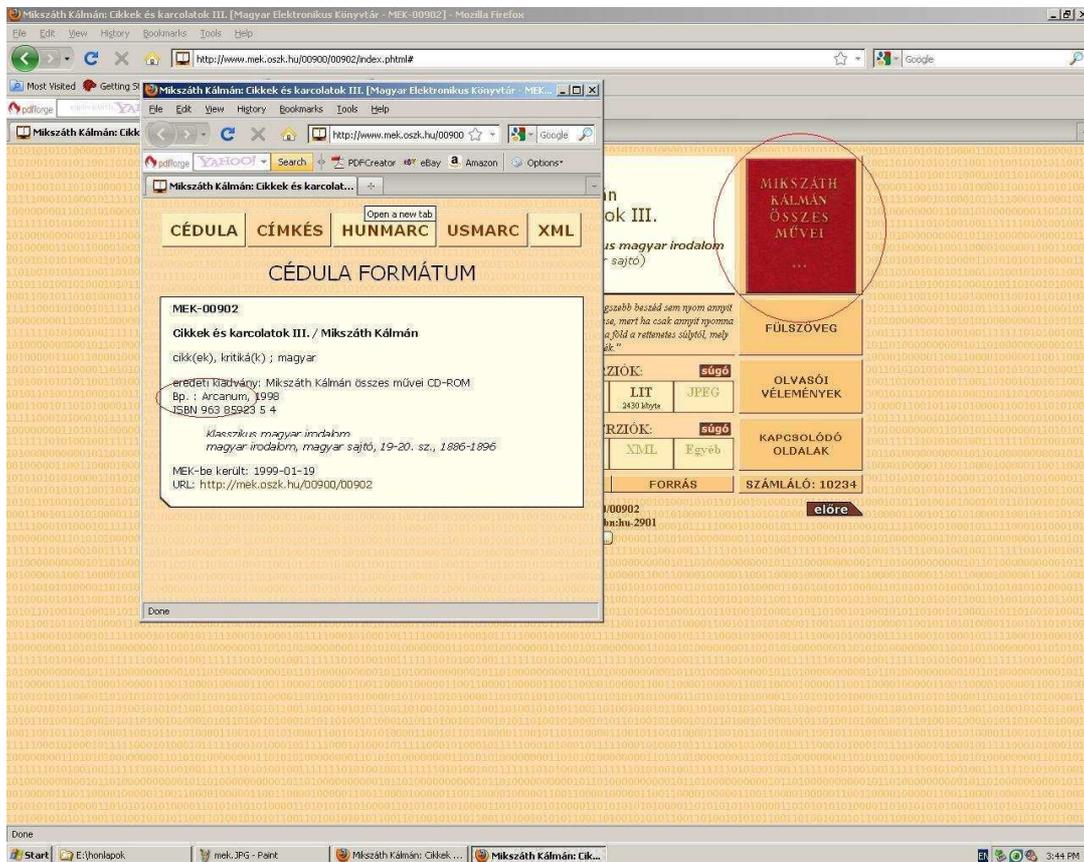


Figure 2

A version of the text appears in the disguise of the most canonical version, moreover it increases its privileged position to exclusivity. The critical edition still contained the list of text variations as part of the notes, while this internet version does not offer a clue about the very existence of alternatives.

Two years ago with a small team we decided to finish the Complete Work series. But a paper publication of journalism with full-length commentary seemed too expensive and superfluous after all. What was important, was collecting late journalism. Nobody knew how many political

and literary essays Mikszáth wrote in the last 13 years of his life, and many of those papers were practically non-existent, since they were published in newspapers and never re-published. Many others, however, appear repeatedly in various collections of his essays. Therefore a possibly full collection of the late journalism seemed more important than finishing the book series. And it was just obvious that *MEK* is the place where they can appear.

But I wanted to publish the texts with or rather in all their variations. Let me refer to one single example. For more than 20 years, Kálmán Mikszáth was the editor of an *Almanac*, which contained a calendar, an introduction by him and an anthology of Hungarian short stories published in the previous year, and a lot of ads. In 1898 in the place of introduction he put an essay called "Franz Joseph as short story writer", which analysed two texts the king-to-be of Hungary wrote as a child when he was learning Hungarian. Probably as a kind of promotion this essay was published in a newspaper which Mikszáth himself the edited as the editor in chief, simultaneously with the *Almanac's* appearance on the market. And I have found it in 5 other journals in November 1898. As for the other journals, the differences are not so important, since given the copyright situation of then Hungary journals could take and publish the introduction even without letting the author know, and they might have thought that the *Almanac's* publishers were grateful for the promotion. But the variations Mikszáth's own newspaper offer might be intentional as well. At the end of 1903 a very ambitious newspaper was established (it was

simply called *Az Ujság*, which means 'the newspaper'), and it presented all its subscribers with a luxury edition of Mikszáth's selected journalism for the Easter of 1904. It is called *My contemporaries*. For that purpose Mikszáth modified the text in many places, first of all because he realised that in 1898 he made a mistake when giving a wrong first name to the king's Hungarian tutor. The collection *My contemporaries* was published in a second enlarged, but less luxurious edition in 1907, when Mikszáth inserted a footnote into his essay on young Franz Joseph.

We can see that this text exists in at least four different versions, and it is too easy to say that the last one is the final and most important version. Each version has its own context and its own right for existence, and therefore it seems unfair to canonize only one of them and to exile the others into footnotes (as critical editions usually do) or into nothingness (as internet publications tend to do). Digitised publication theoretically offers various opportunities to make readers realise the simultaneity of variations. Limits of print costs or the physical limits of a book cannot prevent us from publishing all the variants near each other, or to make variant-shifts possible at every place where they differ. Nevertheless I personally had to make a compromise, since *MEK* authorities stuck to a concept of accessibility. Accessibility, of course, is a very important concept to make the life of handicapped people easier, and in my country it is especially important. In the period of communism public buildings, restaurants, underground stations, even streets and buses were designed as if everybody was young and fit. After the

revolution of 1989 accessibility slowly became an issue of politics, and the general view of the Hungarian cities is changing. Internet accessibility, however, is a bit different. The argument says that an online publication should be designed a way that also blind people could have access to it. They cannot see the screen, to be sure. But they can listen to loud reading, which should be therefore completely linear. If the loud reading leads someone into the maze of variants, the blind will be lost. Internet accessibility therefore requires rather monolithic texts and linear proceedings.

I cannot help regarding this concept a pretext to reinvent a very conservative type of reading and text presentation, and to turn back to all the ways internet reading and surfing undermine traditional hierarchical reading habits. My online publications of Mikszáth's late journalism became therefore also conservative. I put the list of variants in endnotes. The variants, the multiplicity of text variations is presented, nevertheless there is a main text, i.e., an offer of a single canonical version at the top of the hierarchy. And those interested in the comparison of variations must work hard to understand the situation, although a computer could make it rather easy for them.