Copy Protection Via Plagiarism Search

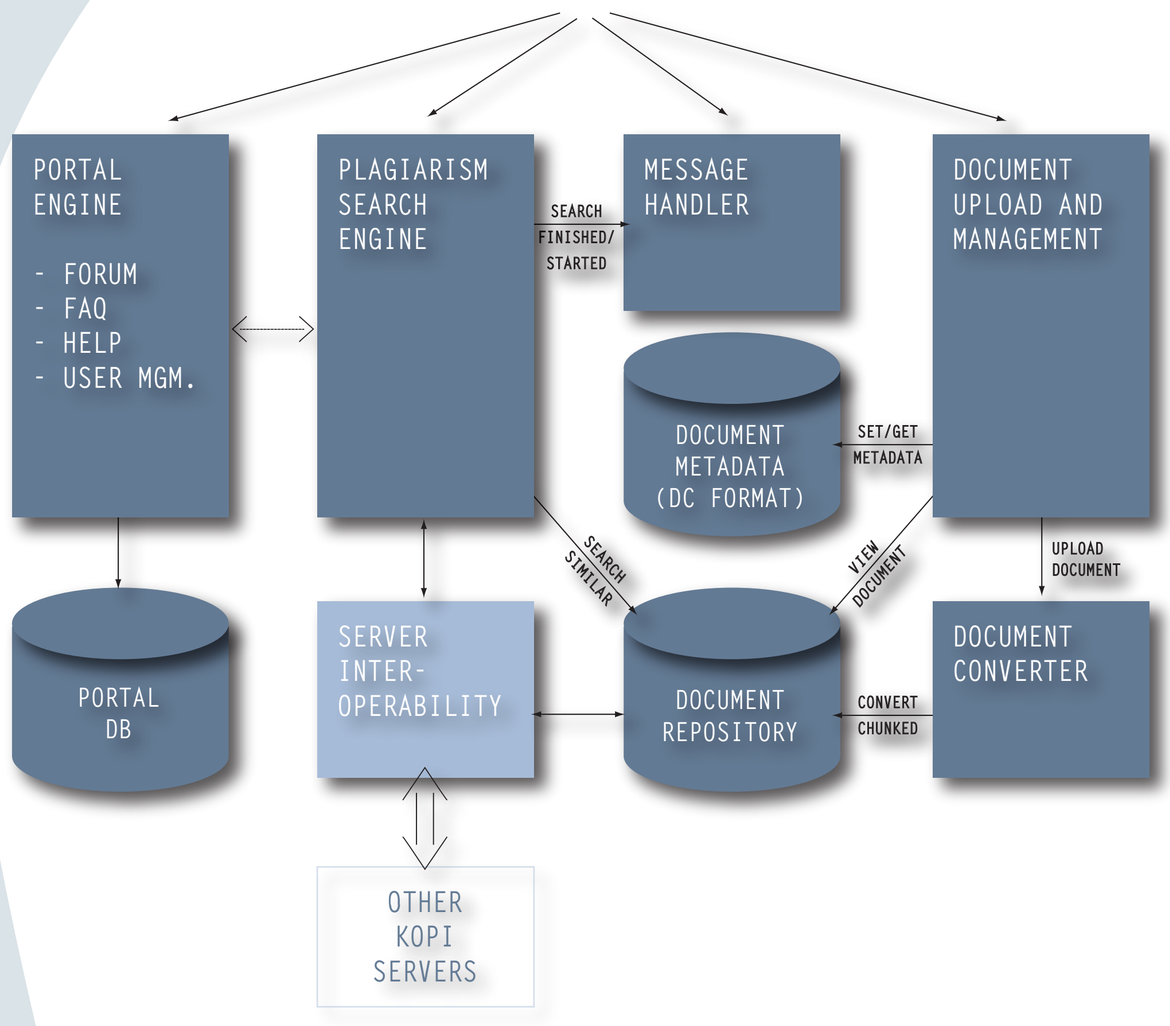
Public accessibility of digital libraries (DLs) highly depends on the characteristics of the works the respective DL contains. Recent works, publications and theses are rarely accessible for the wide public, or in case they are, they use some kind of copy protection (like protected PDF files, Java Applets or even proprietary client applications) to avoid plagiarism. However, preventing unauthorized copying and, at the same time, ensuring that authorized people can easily access them is very difficult. With ready-made tools available on the Internet free of charge, most copy protection mechanisms can be easily circumvented. Other mechanisms are harder to break, while the official use is also made too complex, in many cases users have to install special programs or tools, which may not work in all systems or which may take too long to get through and would discourage people to further attempt access. Also, other users may face access difficulties when using special tools, such as mobile phones or home page readers.

The KOPI Plagiarism Search System developed by the Distributed Systems Department of the MTA SZTAKI proposes an interim solution to protect DLs against plagiarism. Here the protection is twofold: Firstly, if a work is copied, the system can tell whom it was copied from. Secondly, ubiquitous access, the widespread use and wide familiarity of the system can prevent people from presenting others` work as their respective work, as nobody would risk being exposed to be a plagiarist.

The academic society can have the greatest advantage of the system, as information (theses, papers etc.) could be freely circulated among students and professors without worrying about mass plagiarism. This way students may build on the knowledge and achievements of the others, may use appropriate references and would most probably make better achievements in their work. If digital libraries at the universities also comprise the theses and other works of the students freely available to the public, then companies and enterprises might search for future employees there as they could have an insight into the theses in their area of interest and could make a "pre-selection" based on the profile and the quality of the work.

Future Work

During the last 4 years, we gathered a lot of useful information from our own experience and from the feedback and comments of our users regarding the system and the way to make it more effective. Based on these, we would like to implement an external interface for automatic document upload and plagiarism search (e.g. SOAP), as this way KOPI could be easily integrated to existing systems. Also, some universities would require a system which can be operated by themselves, and so they could upload also sensitive data into it. KOPI would be installed at different institutions, yet, the systems would be able to initiate searches in one an other's database without giving access to their documents. The implementation of the distributed plagiarism search system would be the next step to encourage a widespread use of the KOPI System at Hungarian universities.



Architecture of the KOPI System

Copy Protection *Versus* KOPI Protection

Copy Protection	KOPI Protection
☑ It makes copying more difficult and complicated	☒ Does not protect against making a copy of the document
☑ Protection level can be adjusted (against copying or printing etc.)	☒ People have to use it to discover copy violation
☑ If it is a proprietary work, the copyright holders get more income	☒ A couple of larger plagiarism search systems have to be used to be really effective
☒ Most copy protection mechanisms can be easily circumvented (Ready-made tools available on the Internet)	☑ Documents can be freely distributed
☒ Some are harder to break, but the official use is also made too complex	☑ Can be used together with any other protection mechanism
☒ If once broken, the document is unprotected	☑ Protects also parts of the document
☒ In some cases legal users also have to circumvent it or have to ask for removal of the protection (e.g. blind people)	☑ Any type of copy (digitized or paper) is protected
☒ People using special tools (mobile phones, screen readers) cannot access them	☑ Can tell the source of the work plagiarized
☒ Web crawlers cannot index a lot of the protected documents	☑ Exposes plagiarists when submitting others' work as their own
	☑ Cannot be automatically circumvented

sztaki kopi
http://kopi.sztaki.hu/

DSD DEPARTMENT OF DISTRIBUTED SYSTEMS MTA SZTAKI

László Kovács
Laszlo.Kovacs@sztaki.hu

Máté Pataki
Mate.Pataki@sztaki.hu

MTA SZTAKI, Hungarian Academy of Sciences, Computer and Automation Research Institute Department of Distributed Systems
Address: H-1111 Budapest, Hungary, Lagymanyosi u. 11. Phone: +36 1 279-6269, Fax: +36 1 279-6200 Web: http://dsd.sztaki.hu