

# Proposal Of A System For Computer-Based Case And Evidence Management

Prof. Dr. Fritjof Haft, Pascal Hassenpflug, Hans Lecker

Normfall GmbH  
Bismarckstraße 22  
80803 München, Germany  
imf2007@normfall.de

**Abstract:** Evidence secured from computers created the need for computer-based evidence management. But there are other good reasons: Presently a growing number of major cases with a large amount of electronic and paper evidence creates a massive overload for the attorneys and judges involved. Hence a growing demand for better organisation and new tools in this area is evident. Leading auditing companies are offering forensic and dispute services, fraud services, forensic integrity services etc to their clients. But there is a big gap between their results and the everyday needs of criminal lawyers. The authors, one of which is a long-time university professor and practitioner of German criminal law and legal informatics, claim that there is significant potential to increase quality and lower cost through a fully computer-based case and evidence management process for large and complex legal cases. They define the process and discuss how Electronic Document Discovery (EDD) and Document Management Systems (DMS), which are commonly used by forensic services as well as innovative Case Management (CM) systems, can be combined to support it. Their company Normfall GmbH is one of the first to develop CM tools in Germany, which have already been used successfully in major cases of white-collar crime.

## 1 Overview

IT Forensics is the term applied to the area of securing evidence of IT security incidents. In a broader definition, of securing evidence from computer systems, which includes computer-based management of the evidence gathered. The next step is as natural as it is necessary – to include the management of any kind of evidence by computer systems.

Whereas standards for electronic document exchange with juridical authorities are finally being established, cases and paper evidence are still often managed in a non-electronic fashion. The U.S.-American market for legal software is more developed than the European, but not all products can be easily adapted to suit European needs.

This paper deals with the design of a computer system to support document-intensive cases such as in commercial law.

We will first present a motivation from the point of view of a legal practitioner experienced with the current situation in Germany. We will then lay out how work on a large case can be done in an efficient, cost-effective way under the aspect of process organization. To support this process, several kinds of software tools can be used. We will present the main categories of existing tools, then formulate requirements for an overall solution and finally discuss ways of combining existing tools to fulfill our requirements.

## 2 Motivation

On February 25th 2005 the German Parliament („Deutscher Bundestag“) enacted the public statute concerning the use of electronic communication in the judiciary („Gesetz über die Verwendung elektronischer Kommunikationsformen in der Justiz Justizkommunikationsgesetz – JKomG“, BGBl I 2006, S. 837 ff). It regulates the electronic relations and the use of electronic files in all areas of the German legal system. But no thought has been given to the question what substantial improvements of legal procedure are possible when using IT technology. Although there is worldwide research in legal informatics, German courts and lawyers don't go much beyond production of electronic documents, the exchange of emails and the use of the Internet.<sup>1</sup>

This defect produces damages especially in complex criminal procedures, which occur more and more in many fields of criminal law (i.e. penal law concerning business, fiscal fraud, deceit in the medical field, corruption in companies etc.) Judicial evidence often consists of many thousands of files. The principle of oral presentation and the lack of a written protocol in all cases brought the county courts („Landgerichte“) to overestimate the testimonial evidence, even though this is the least valuable evidence of all. The factual truth that is hidden in the files often cannot be found simply because nobody – neither the judges nor the attorneys – is able to deal with the overabundance of files.

Electronic documentation offers some progress like OCR and full text search. But this gives little help if the searched descriptors are not very specific. Furthermore, it takes too much time. If somebody appears as witness in court and she or he is asked questions, the questioner must be able to retrieve relevant documents within seconds. No known document retrieval system was designed to enable judges and lawyers to do this. As a result presumption often were (and still are) the bases of the court decisions although certainty was (and is) in reach.

---

<sup>1</sup> Legal informatics pertains to the application of informatics within the context of the legal environment. It encompasses two conceptual areas: IT in law and law of IT. While the latter (issues such as privacy law, copyright law etc.) is well developed in Germany, the former (issues such as information retrieval, artificial intelligence and even such practice issues as applications which help lawyers in their day-to-day operations) is less developed.

The German law system, like many other continental European law systems, follows the tradition of the Roman law. So a short look at the history of law might be helpful for understanding the way a modern tool for computer-based case and evidence management has to be designed.

Originally, Roman law was, like modern Anglo-American law, case law. The capabilities of Roman jurists were demonstrated by their capacity to find elegant solutions for specific cases, many of which have lasted up to the present. These solutions were recorded in the so-called *responsa*. In the *responsa* the case and its solution were formulated and presented as briefly and lucidly as possible.

As the centuries passed, thousands and thousands of such *responsa* came into existence. This created two problems. How could one find existing precedents when dealing with new cases? And, even more importantly, how should one organize any meaningful law education in the face of such an abundance of legal material?

The answer to this was found by the Roman jurist Gaius (around 160 A.D.). He made use of the insights formulated in the philosophy of science by the Greeks, whereby terms have both content and scope; a reciprocal relationship exists between the two. A lot of content means little scope and vice versa. The Greeks used this to create hierarchically ordered and structured pyramids of terms which were used to try and explore the essence of things. The Romans were more practically minded – and used this concept to create something which could be applied in practice. Gaius made use of the insights of the Greeks to develop a system of hierarchical, structured legal terms and definitions. The legal systems of continental Europe have developed out of this tradition. Hierarchies of juridical terms have been developed for almost two thousand years. In the 19th century a so-called *Begriffsjurisprudenz* was created in Germany, whereby every individual case could be subsumed under a codified legal term; leading German jurists were of the opinion that the construction of legal terms was the sole task of the study of law.

The interesting thing is that the structural systems are so complex that – although they are formulated using specific terms - they can no longer be adequately represented by language alone. We can only practice solving a case by training the use of the system on the basis of legal cases. Just as with typological terms, certain forms of knowledge exist which are located in language but which cannot be comprehensively described in language. Unlike typological terms, however, there is no everyday experience to help us out. Most people will know whether something is good or bad on the basis of their daily lives. Only the trained jurist will “know”, if this is in accordance with the rules of law without him being able to grasp this knowledge fully and put it in words. This knowledge is embedded in the system and constitutes a specifically juridical variation of the sense of justice.

This finding is particularly interesting from the point of view of behavioral theory and has as yet not been explored in any depth. We consider this to be an interesting field for further research and we are convinced that it should be possible to use computers to create tools for investigation, which will be able to help us cope with the challenges mentioned.

This background, which is familiar to every German lawyer, induced the creation of our software tool – the Normfall Manager – that has been used successfully in several criminal cases.

### **3 A Process Organization To Support Large Cases**

In our vision of how large cases will be managed in the future, the key element is the organizational step after the evidence has been secured: The evidence must be evaluated and organized. This can be done by a team other than the people actually working on the case. It can even be outsourced. In the past, highly paid professionals often spent their expensive hours trying to extract the information important to them from a large amount of paper evidence. This is still a reality for many, but some have already managed the step to use DMS systems (in conjunction with EDD systems and scanners) to enable full-text search of documentary evidence. However, this creates the illusion that computers can take the task of interpreting the evidence off our hands. But whereas full-text search is a good tool, the evidence must still be viewed and evaluated by legal professionals in its entirety. Computer systems must support this evaluation process and provide a means of making the result accessible to everyone who is working on the case.

The outline of this organizational process is as follows (Figure 1):

#### **3.1 Capture and Discovery**

*Capture and discovery* refers to the phase where evidence is discovered and copied from electronic storage devices and paper evidence is made available digitally through scanning. This process needs to be integrated with the archive layer so that each copy of the evidence can be traced back to the original piece of evidence.

At this stage, technical expertise is needed rather than legal expertise. Electronic Document Discovery (EDD) systems should be used to gather documents that may be relevant as evidence. All printable document formats can be used as input to the next stages, but content in binary or other form that is not easily printable or viewable outside the system it was authored on (e.g. a database) must be analyzed by EDD experts, and their findings submitted to the next stages.

#### **3.2 Document Evaluation**

Lawyers work together with a team of qualified assistants, whom they instruct to evaluate the documents.

#### **3.3 Case Management**

The findings will be entered into the case management system, in which the lawyers that work on the case organize all the information pertaining to it.

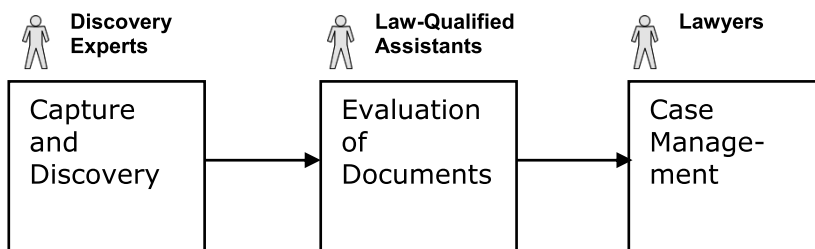


Figure 1: Three stages of preparing documentary evidence

## 4 Existing software systems

We will now present different categories of software systems that are used in case and evidence management.

### 4.1 Electronic Document Discovery systems (EDD)

Electronic Document Discovery (EDD) refers to the process of retrieving documents from electronic data storage devices as is necessary for using the information as evidence in a trial.

### 4.2 Document Management Systems (DMS)

Document Management Systems (DMS) manage storage and retrieval of documents. They can have very different characteristics, depending on what they are used for. The main characteristics of a DMS used in evidence management would be:

- Read-only archive that controls access to the documents
- High stability and good data backup capabilities
- High performance, file-format independent full-text search engine
- Document origin tags that point to the original piece of evidence.

### 4.3 Case Management / Litigation Software

Case management software, also known as litigation software, lets lawyers organize the information about a legal case. As with DMS, the particular features may vary depending on the intended use. In this document, where we assume large and complex cases, the aim is to build an informational structure of the case, that helps users understand the case, communicate about the case and access (evidence) documents using the case structure as a key.

In contrast to the United States, German attorneys and prosecutors rarely use case management tools – yet. We think that this has mainly to do with the pace at which German lawyers embrace technological innovation. The need for such tools seems obvious to us, especially in large and complex cases. Normfall GmbH, the authors' company, is one of the first to offer this kind of tools in Germany.

The main idea behind case management is that all information pertaining to a case is organized in a systematic manner, so that it can quickly be retrieved, and the people working on a case can easily get an overview of it. For cases with large amounts of documentary evidence, it is especially important to organize this evidence around the case issues. We propose a tree structure of the entire information in a case, e.g. of the legal issues, where all correspondence that is found in the evidence and corresponds to a given issue is attached to the structure node representing that issue. We have also created a tool for this which is currently marketed as the Normfall Manager.

The following are typical information categories for case management, which could be represented as nodes in a tree structure, or as records in a table:

**Facts:** Claimed facts of both sides, confirmed or unconfirmed, organized e.g. by

**Issues:** Organizes the legal battleground – what are the legal or other (e.g. moral) issues in the case.

**Cast of Characters:** Which natural or juridical persons are involved, and what are their roles and relations to each other?

**Evidence:** Lists all the evidence belonging to the case.

**Investigation Questions:** What information has yet to be investigated?

**Law Research:** What is the law relevant to this case?

### 4.4 Trial Management (Process-Enhanced Case Management)

Although not yet implemented, this is another area for possible computer support in the future.

The so-called *Relationstechnik* („Relation“) is a juridical method which is used in German civil cases for the assessment, arrangement and appraisal of complex contradictory facts. It is used by judges as well as by lawyers. Every German jurist is training this method during his education at universities, courts and law firms. For the judges, *Relation* is the quickest and cheapest way to find a correct judgement. For the lawyer it is an proven method as well.

The *Relation* consists of six possible steps: Sufficiency of motion, logical coherence of action, relevance of answer, contestation and plea, evidence and judgement.

These steps or *stations* can be reproduced in software, so that each case's current station can be captured. The claims of each of the opposing sides can be tracked with their status, e.g. which side needs to prove the claimed fact, whether it is disputed or not, and what both sides have said about it.

The trial management tool could be combined with features supporting electronic document exchange with the courts as have been specified by the German government.

## **5 Functional Requirements**

We will now outline the functional requirements for the combined case and evidence management system. These are only the requirements that are relevant to the combination of systems (EDD, DMS, case management) whose properties are already well-known.

### **5.1 Capture and discovery**

Documents created by Electronic Document Discovery and paper scanners must be tagged to identify their original source. The tag will be maintained throughout the entire system.

### **5.2 Original Documents**

If digitized evidence documents will be changed or converted during the case management process, their originals must be maintained on an archive system.

### **5.3 Universal viewing capabilities**

After the capture and discovery stage, all documents in the system must be easily viewable, independent of their digital format (at least for all digital formats that are designed to be printable).

## **5.4 Full Text Search**

Mainly during document evaluation phase, users must be able to perform a full text search of all documents.

## **5.5 Commenting**

The system needs to support commenting of the documents by multiple users. Ideally, the comments should be visible in the document viewer, but not be stored within the document files themselves.

## **5.6 Structuring and case management**

To prepare for the case, lawyers need to organize the information around the case issues.

## **5.7 Multi-user access**

On a complex case, lawyers from different firms may work on a case, each contributing their expertise in a special area. The system must therefore support multiple users.

## **5.8 Remote access**

Given that the users may come from different firms, they need remote access to work with the system from their office.

## **5.9 Offline availability**

A lawyer should be able to bring all case information to court on a laptop computer.

## **5.10 Ad-hoc installation**

We assume ad-hoc collaboration of lawyers with different kinds of expertise. This implies the need for ad-hoc installation of the client software, taking into account the various network and software security policies that may be in effect at each lawyer's office.

# **6 Composing the system**

Existing software systems can and should be used, leveraging time-tested technology. The challenge is to combine them to actually support a distributed team of law professionals working on a large and complex legal case.



## 6.1 Document Sources

At first, we need to look at the sources we want to use for our evidence management – these are mainly electronic data storage devices such as hard disks, and paper evidence. The paper evidence is captured using a scanner. Typically, it is enriched by *Optical Character Recognition (OCR)* to enable full-text search. Probably the most popular format to combine picture and OCR data in one file is the *Portable Document Format (PDF)*.

In contrast to purely electronic documents, scanned documents always contain a picture layer which puts certain restraint on the way the document can be handled in a scenario of remote access, as we will see later.

From the start, the DMS needs to keep account of the origin of a document. Users must be able to trace it back to the original piece of evidence.

## 6.2 Original Live Systems

The software used to retrieve data off electronic storage devices is often referred to as Electronic Document Discovery (EDD) software, as mentioned before. The main challenge with data retrieved by the EDD is the file formats – they can be anything. Now one approach is to try and rebuild the live system a document base was used on, say on a virtual machine. Whereas this may in some instances be justified, for example to use document retrieval functions unique to this system, it puts great constraints on our possibilities to organize the documents in a case management software (we would have to run the case management software on the system that we built).

## 6.3 Universal Viewing

Our preference is to have the information “dead”, on our own system, but still be able to view it. There are two ways to achieve this: By using a universal document viewer, or by using a universal document converter, that converts all documents to a common format such as PDF. As we will see, the conversion scenario proves to be advantageous when dealing with offline access.

Either way, we still want to be able to access a document in its original file format if necessary. Therefore, we need a DMS system that simply stores all of the files in their original format. Note that there are special considerations for viewing files in their original form.

## 6.4 Protection From Malicious Code

Many documents such as Word documents may contain malicious program code that must be restrained from executing on the machine of the evidence management system’s user. Even document formats that do not support executable code may contain code that is brought to execution on a particular viewer using buffer overflow techniques.

The most practical ways of achieving protection are:

- Accessing the documents on an isolated execution engine (stand-alone computer or virtual machine).
- Converting all document types to a single document viewer format such as PDF. The converter must not be vulnerable to malicious code within the documents or must be isolated.
- Making sure that only the universal viewer can access the documents, and that the viewer is not vulnerable to malicious code within the documents.

## 6.5 From DMS To Case Management

Our next step is to go from a simple DMS to a case management system. The documents must be viewed before they can be attached to the case structure in the case management system. This might be a good point to do the conversion as well: A universal document viewer might come with conversion capabilities. So after a document has been viewed, it is converted and transferred to the case management system. If PDF is used, the document origin tag may be translated into embedded metadata. The information stored with the document in the case management system must also enable us to find the copy of the document in the DMS efficiently if needs be.

An alternative is to do batch conversion (the question here is whether user assistance might be necessary in the conversion process), and to store the converted documents in the DMS. This might be a good solution if the DMS also supports project-based offline caches (see below).

## 6.6 DMS / Case Management Integration

The case management system can either refer to the documents by their ID in the DMS, or keep its own copy, such as in the scenario where we generate PDF versions of all the documents we want to use in the case structure. As the case management system does not need advanced DMS functions (such as file format independent full-text search), it may store the documents in a project folder in the file system, giving us an easy implementation of an offline cache: simply copy the folder.

To sum it up, the options are:

- **Viewable documents stored in DMS:** The case management system only stores document IDs from the DMS.
- **Viewable documents stored in case manager:** The case management system keeps its own copy of each document in a standard format.

## 6.7 Remote Access

We need to ask ourselves: At which stage is remote access needed, for what tasks, and what are the requirements? Preferable, we do not want to transfer original documents because of the security considerations stated above. That leaves us with two options:

- To deliver documents in a “harmless” viewing format (e.g. PDF).
- Remote Desktop access to a safe environment such as a virtual machine.

The weak point in all remote desktop scenarios is the use of scanned documents because they consist of image data. We created a test setup where we accessed documents over a VPN tunnel. With an underlying Internet link of 2 Mbit/s, which should be a realistic value for many remote workplaces, we could scroll through a document with considerable flickering - still workable, but not comfortable. Once an area of the document was outside of the screen, it had to be reloaded completely the next time it was scrolled into view. The bandwidth requirement on the server is particularly high, because new image data is requested any time a client scrolls.

If the actual document is to be transferred to the client, the performance for large documents can only be acceptable if partial transfer is supported. E.g., when using PDF over the HTTP protocol, the Adobe Reader supports a feature known as *byteserving*, where only individual pages are transmitted and not the whole document. There is an important option in the Adobe Reader software which is set to load all pages in the background when byteserving is used. If this option is turned off, loading individual pages over a 640 kBit/s link is a very smooth experience and puts much less stress on the server. It is no problem to use this feature from a Web application to request an individual page of a document.

We believe that a Web application is the best document viewing solution, but we have to weigh it against architectural issues (on the evaluation stage, we may not have the documents in a Web-friendly format), and implementation issues (it may not be practical to develop a high-end, powerful case management tool for the Web).

## 6.8 Remote Client Installation

Deployment of the remote client is an important issue. That is because, as mentioned above, heterogeneous groups of lawyers may work on a project in an ad-hoc manner. They don't all use the same infrastructure. The deployment solution should ideally work without the need for an administrator, particularly for updates. Whereas the Java applet framework was once designed for this purpose, thin client solutions such as remote desktop clients and Web browsers are far more popular with system designers today. Thin clients eliminate the need for updates completely because the actual application logic runs on the server.

We need to consider network access in our deployment scenarios as well. Most naturally, with sensitive data, the connection would have to be encrypted. Whereas it is generally no problem to establish an encrypted HTTP over SSL connection for a normal user at any organization, remote desktop protocols or VPN tunnels may not be routed.

The design of the case manager plays another role here. If it is a complex application, it may be difficult to port to the Web.

## 6.9 Offline Clients

We still need an offline client. There are two offline client scenarios, depending on where the viewable documents are stored:

- **DMS:** The case management system only stores document IDs from the DMS. The DMS has integrated offline capabilities, so the whole document base for a project can be exported to a machine and used with an offline client of the DMS.
- **Case manager:** The case management system keeps its own copy of each document in a standard format. It can then implement the offline functionality, for example by simply copying the project folder to the offline machine.

We see the first scenario as the more elegant one because the boundaries of competence of each system are better defined. We would only prefer the second scenario if the documents managed by the case management software would somehow differ from those in the DMS, such as when the content is enhanced by comments of the users working on the case, which brings us to our next consideration:

## 6.10 Comments and Collaboration

An ideal commenting system supports comments that refer to certain passages in the document, and the system is able to show the document with a marker at that passage. The comment itself should be shown outside of the document viewing area, and other users should be able to add their own comments to the primary comment.

The system should also be able to show documents without the comment markers, and comments would best be implemented as data structures outside of the actual document file, so the file is kept in its original form.

Commenting might be a feature of the DMS as well as the case manager.

If the DMS provides an commenting system, the case management system should be able to “call” the comment by an ID.

In a simple solution, PDF annotations would be used as comments. They can be linked to by the case management system directly through an ID.

## 7 Proposals of an evidence management system

Building on the previous discussion, we present the following proposals (see also: Figure Figure 2). Note that PDF can be substituted by an equivalent format.

### Alternative 1: PDF on the DMS side

- Documents are converted to a universal format (PDF) at an early stage.
- PDF copies and offline cache are both managed by the DMS.
- DMS comes with remote viewing client and offline cache client.
- Case management system links to comment IDs in the DMS. Case management system has its own offline and remote access solution.

### Alternative 2: PDF on the case manager side

- Documents are converted to a universal format (PDF) at an early stage.
- PDF copies and offline cache are both managed by the case management system.
- Case management system comes with remote viewing client as a Web application, and an offline cache client.

### Alternative 3: Original documents / universal viewer

- Documents are left in their original state.
- Documents and case management software can only be accessed from within a virtual machine.
- With online access, virtual machines running on the server may be an alternative.
- With offline access, virtual machine including document offline cache must be installed on the offline computer.
- Can be combined with document management through a dedicated DMS or through a case manager tool (in the table below: “Access by DMS / CM”).

		<b>Alternative 1: PDF / Access by DMS</b>	<b>Alternative 2: PDF / Access by CM</b>	<b>Alternative 3: Original Format</b>
	Access binary data	No	No	Yes
	Viewer	PDF	PDF	Universal Viewer
	Comment Display	PDF Viewer or DMS or Case Manager	PDF Viewer or Case Manager	Universal Viewer or DMS or Case Manager
	Comment Management	DMS or Case Manager	Case Manager	DMS or Case Manager
	Viewer Security	Doc Conversion	Doc Conversion	Virtual Machine
	Special Capabilities	DMS: Manage PDF version	Case Manager or PDF export: Manage origin tag	Alt. 1 or 2
	<b>Overall document availability rating</b>	<b>good</b>	<b>medium</b>	<b>optimal</b>
<b>Remote</b>	Protocol	Web	Web	Terminal (RDP)
	Speed (Text)	good	good	good
	Speed (Image)	good	good	medium
	Remote delivery server (HTTP)	DMS	Case Manager	Alt. 1 or 2
	Client Installation	Alt. 2 + possibly DMS client (not if case manager generates access URLs)	Possibly case manager client (not if Web application)	RDP Client
	<b>Overall remote access rating</b>	<b>good</b>	<b>good</b>	<b>medium</b>
	Case Manager Offline Capability	Case data	Case data + documents	Case data + documents
<b>Offline</b>	DMS Offline Capability necessary	Yes	No	Alt. 1 or 2
	Installation	Case Manager + DMS	Case Manager	Virtualization Software
	Document Availability	All	Case Manager Documents	Alt. 1 or 2
	Offline Export	Case Manager + DMS	Case Manager	Alt. 1 or 2 to dedicated file system
	Offline Changes	Comment + case manager must support it	Case manager must support it	Alt. 1 or 2 + file system must support it
	<b>Overall offline access rating</b>	<b>good</b>	<b>optimal</b>	<b>medium</b>

Table 1: Overview of the three alternative solutions proposed

Remote access: If the case management tool integrates document delivery and case management into one Web application, we see the *alternative 2* architecture as superior. In all other cases, *alternative 1* seems the better integration option.

Note that we do not present a total overall rating for the three alternatives. The choice may depend on the tools available. We see alternative 1 as having the best overall architecture, alternative 2 as the easiest implementation and alternative 3 as the best availability if combined with features from alternative 1 (i.e. the documents are delivered by the DMS).

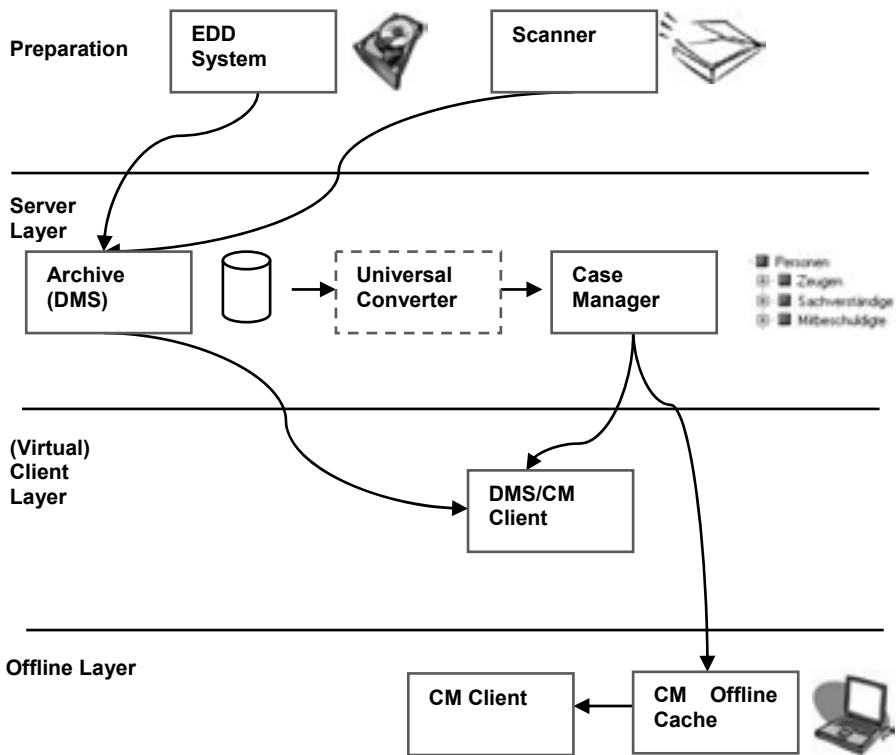


Figure 2: Proposal of an evidence management system, including EDD, DMS and Case Manager (alternative 2)

Of course, the splitting of the architecture into DMS and case management tool is somewhat arbitrary. The idea behind it is that most DMS will not have sufficient case management functions and most case management tools will lack DMS features. Both could be combined in one tool, but to us, this doesn't really reflect the current market situation.

## **8 Conclusion**

In this paper, we have proposed a process organization for managing large cases, emphasizing a step where evidence is evaluated by law-qualified professionals.

We have discussed how a system for computer-based case and evidence management supporting that process might look like. We have presented three alternative solutions for the architecture of such a system. Ad-hoc collaboration and remote access by independent lawyers were especially considered.

A solution involving terminal services and machine virtualization was considered best for document authenticity and availability. Two solutions involving conversion of documents to a universal file format (PDF) were considered best for usability.

We believe that the ideas presented in this paper may help to advance the discussion about electronic support for evidence and case management in German law practice, an area in which innovation, although essential, has stagnated for some time now. We are committed to continue work in this area, in theory as well as in practice through our company's Normfall Manager product line of case management software.