

Case Study: The Use of RealObjects PDFreactor® to meet statutory Archiving Obligations (PDF/A) at the Internet Branch of the Sparkasse Savings Banks

About Finanz Informatik

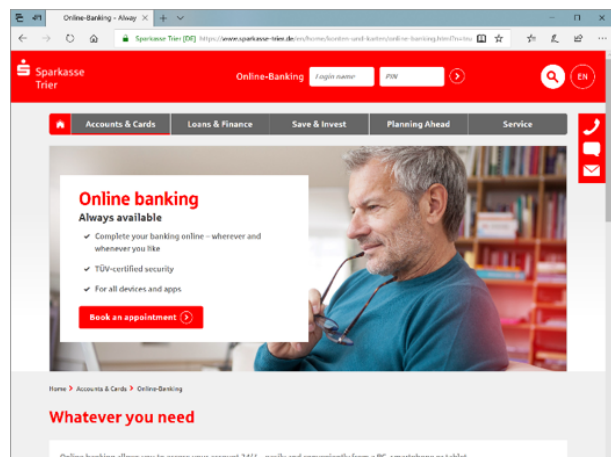
Finanz Informatik, headquartered in Frankfurt am Main, is the central IT service provider of Sparkasse Finance Group. Its customers include 384 Sparkasse savings banks, six state banks, DekaBank, eight state Landesbausparkasse state building societies, public sector insurers and further companies within the Sparkasse Finance Group. Finanz Informatik offers a complete IT service – from application development through infrastructure and data centre operations to consultancy, training and support. In OSPLus the company is now providing one of the leading overall banking systems for the German market.

Finanz Informatik is undertaking the service for 117 million bank accounts; more than 128 billion technical transactions are executed on the computers and systems every year. At the end of 2018 the company employed 3,560 members of staff and the sales revenue was approximately 1.9 billion Euros in the 2018 business year together with the subsidiaries.

Subsidiaries such as Finanz Informatik Technologie Service, Finanz Informatik Solutions Plus, inasys and finally Star Finanz supplement the Finanz Informatik IT portfolio with the services they offer.

The Issue

According to Directive 2004/39/EC of 21.04.2004, released by the European Parliament and the Council regarding markets for financial instruments (MiFID, abbreviation for Markets in Financial Instruments Directive) banks are obliged to archive advertising content presented to their customers for certain types of products for a period of five years and to submit this in the event of legal disputes. This also applies to advertising information shown online via the Internet branches of the Sparkasse savings banks. Realization of archiving that conforms to the directive (conforming to PDF/A) presents substantial challenges to Finanz Informatik as the operator of the Internet branch.



As an online banking portal the Internet branch of the Sparkasse savings banks is a complex and highly dynamic web application and not a collection of static pages. For the Sparkasse savings banks it is the central hub for service and dialogue with its customers within the digital multichannel system. Thus all the product, consultancy and service offers that are relevant for the final customer are integrated into the Internet branch and are reviewed constantly and developed as necessary. The Sparkasse content and banking pages are accessed approximately one billion times a month – more often than any other German web presence. This puts the Sparkasse site ahead of sites like Bild.de, T-Online.de or eBay Germany.

According to MiFID content that needs to be archived can potentially be shown on different pages and page areas and in different contexts. The almost 390 Sparkasse savings banks also have editorial freedoms that are used for the extensive, institution-specific adjustments and additions to centrally provided content. Both central and institute-specific content is updated frequently, sometimes daily. The program code for the application is also subject to constant development with a standard release cycle of six months.

Finanz Informatik has examined various approaches to how the MiFID requirements can be fulfilled under these conditions. Emphasis was also placed on legally compliant archiving by supporting the long-term archiving format PDF/A, the format for the long-term archiving of digital documents that was normed by the International Organization for Standardization (ISO) as a subset of the Portable Document Format (PDF).

Examined Solution Approaches without PDFReactor

One straightforward approach would be to submit archived web information in the event of an MiFID request, i.e. if information was demanded, in order to reproduce the dynamic page in question with the exact content that was valid at the requested time. The page would have to be regenerated based on the archived content and using the code status of the application that was valid at the given time. The constant further development of the application code would mean that every code status that was ever productively utilized would have to be stored in an operable manner on a system or newly installed at the time of the MiFID request. This kind of approach proved to be very elaborate to realize given the complexity of the application and taking into account the systems and frameworks used.

In a further approach it was examined whether the requirements could be fulfilled with the mechanism that is already present in the editing system and used to archive the editing view of pages. This also proved to be unsuitable because the performance of the editing application is reduced too far if the data quantity is stored retroactively at all times for a period of five years. Apart from this one of the application codes upon which the development is based is already required for the creation of the editing view and the same problem would occur as for the first approach.

Realized Approach with RealObjects PDFReactor

As the final approach a procedure was investigated whereby a static representation of every page that can be generated dynamically and is relevant for MiFID is created and archived after every edit modification to the content or development of the application code. This means that no dynamic regeneration of the pages is necessary at the time of a later MiFID request and the above problems in the provision of operable historical code statuses or large content quantities that need to be managed in the productive system are resolved. This approach also presented difficult technical challenges but it was possible to master these well by using the PDFReactor product from RealObjects. Thus this approach was substantially simpler and less expensive to realize than the alternatives that were investigated and it was selected for realization.

Finanz Informatik chose a PDFReactor Corporate Application License because the Internet branch is operated using a group concept (50+ groups, 400+ physical and/or virtual servers) and this also covers the future expansion of the server infrastructure. In the following we will look at the technical challenges that arose in connection with this approach and how it was possible to solve them using PDFReactor.

Technical Details

One fundamental challenge in this approach is finding a suitable static representation of the dynamically created pages. A purely text representation is not sufficient because MiFID-relevant content may also be contained in images (e.g. promotionally effective interest rate information). The spatial arrangement and interaction of text and image on the page can also be important for the conveyed content and its relevance to MiFID. Therefore the archived representation of the page must correspond visually as exactly as possible to the original image rendered by the browser. PDF documents are thus suitable as a static representation. However the challenge then lies precisely in the creation of a PDF document that corresponds visually to the browser representation.

This challenge can be mastered particularly well using the specific features of the PDFReactor code library, which stands out from many other libraries for the creation of PDFs. The library is based on a PDF service in each of the server groups. A crucial feature is that PDFReactor takes the rendering of the PDF document (PDF/A 1b) directly from the HTML representation of the page that is to be archived independently. In contrast to many other PDF libraries the programmer does not need to insert texts and images into the PDF that is to be created using code retrieval, explicitly inputting the positions, colors and sizes. Rather it is sufficient to submit the URL of the page to be archived. PDFReactor then works like a browser, i.e. it makes the necessary requests to the server, including the requests for embedded resources such as images and CSS, renders the visual from the HTML representation and then produces it as a PDF. This means that every dynamically generated website that can be called up using an URL can be converted into an archivable PDF immediately with a single library retrieval using PDFReactor, without any intermediate steps that need to be programmed separately. This key feature of PDFReactor alone means that the selected approach to the fulfillment of the MiFID requirements can be realized in the Internet branch so cost effectively.

A further challenge was presented by the correct realization of the PDF representation of the kind of complex CSS layouts that are designed for responsiveness and are used in the Internet branch. It was also possible to easily master this challenge because of the high quality and the standard conformity with which the PDFReactor takes CSS formatting into account.

The Internet branch also uses Javascript frameworks such as JQuery and Plugins that undertake formatting with Javascript after an HTML page is loaded in a browser. Javascript needs to be executed in order to take these layouts into account correctly in the created PDF, as would also be the case in a browser. PDFReactor also offers this feature. Of course it is not possible to display actually dynamic elements such as a rolling opener in a static PDF. This is true in principle and is not related to a limitation in PDFReactor. In the project this weakness in the approach was resolved simply by displaying dynamic elements in the MiFID archiving in an alternative static representation. Thus, for example, all the forms of a rolling opener are shown under one another.

The integration of the archiving functionality in the overall application is another challenge. Alternative PDF libraries that require a browser environment or are based on PHP or other script languages may have caused problems in this regard because the target environment is based on Java and OSGI. However simple and stable integration is guaranteed because PDFReactor is available as an OSGI bundle.

Conclusion

The use of RealObjects PDFreactor allowed an approach to be realized for the Internet branch that guaranteed the fulfilment of statutory archiving requirements (PDF/A compliant) that saved a lot of expenditure compared to the alternative approaches that were investigated and that would not have been easily realizable with other PDF libraries. One Finanz Informatik developer commented: *„From the point of view of the developer this is a great library that is easy to integrate and has enormous scope. The high quality of the support is also noteworthy.“*

Finanz Informatik
Theodor-Heuss-Allee 90
60486 Frankfurt am Main
Germany
E-Mail: info@f-i.de
www.f-i.de