

P0325 環境分析におけるデータ管理とワークフローの効率

○平城里香、佐藤信武、小西泰二、江崎達哉（日本ウォーターズ株式会社）

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

はじめに

環境分析におけるデータ管理の課題とそのソリューション

- ◆ 多様な分析機器
機器の種類とメーカーが異なることによるデータフォーマットとレポートフォーマットの異なるデータの保管/管理の煩雑さ
- ◆ 長期間にわたるデータ保管の必要性
通常5年以上のデータ保管に伴う保管スペースのコスト的負担、と管理のための人的負担
- ◆ 多くのマニュアル処理
最終結果算出までに、表計算ソフトへの数値の手入力やチャートの切り貼り等、手作業による作業による生産性の低下

これらの課題を解決するために、ウォーターズは分析機器の種類やメーカーに依存せず、コンピュータが作成するあらゆるファイルと印刷物のイメージを取り込み、管理するためのNuGenesis® SDMS（Scientific Data Management System）システムを開発し、環境分野での実用化のためのシステム構築例を報告する。

SDMSへのデータの取り込みとインデックス化

オリジナルファイルと印刷イメージの取り込み

SDMSへのデータの取り込みはファイルキャプチャとプリントキャプチャの2種類がある。

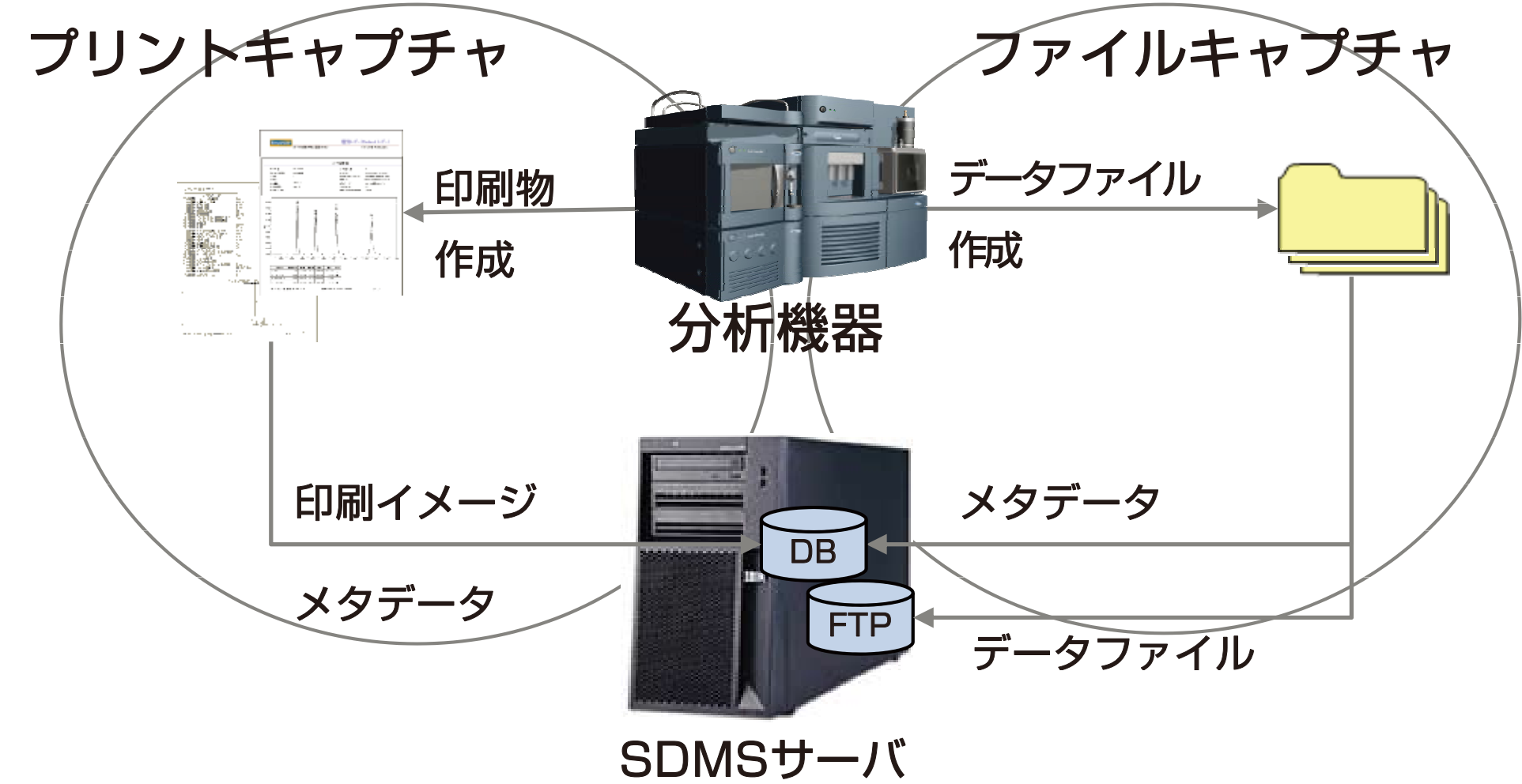


Fig.1 SDMSでのデータ取り込み概略図

- ◆ ファイルキャプチャ
分析機器毎に新しいデータや変更されたデータが存在するかどうかをモニターし、そのデータファイルを自動的にバックアップする。
 - － 機器毎のファイルの所在の指定
 - － モニターの頻度
 - － ファイルの取込みのタイミング（例：作成後1ヶ月、最後の変更後1か月）
 - － ファイルの削除のタイミング（例：取り込み後すぐ、最後の変更後3ヶ月）これにより、バラバラな法則で管理されていた生データを確実に、かつ統一して保管することができます。
- ◆ プリントキャプチャ
分析機器のソフトウェアで印刷を行う際に、その印刷物のイメージを自動的に作成し、データベースに取り込む。プリンタドライバのように働くため、印刷時にプリンタにSDMSを選択する。印刷した紙のデータの代わりとして保管すれば、保管スペースを大幅に削減することができます。

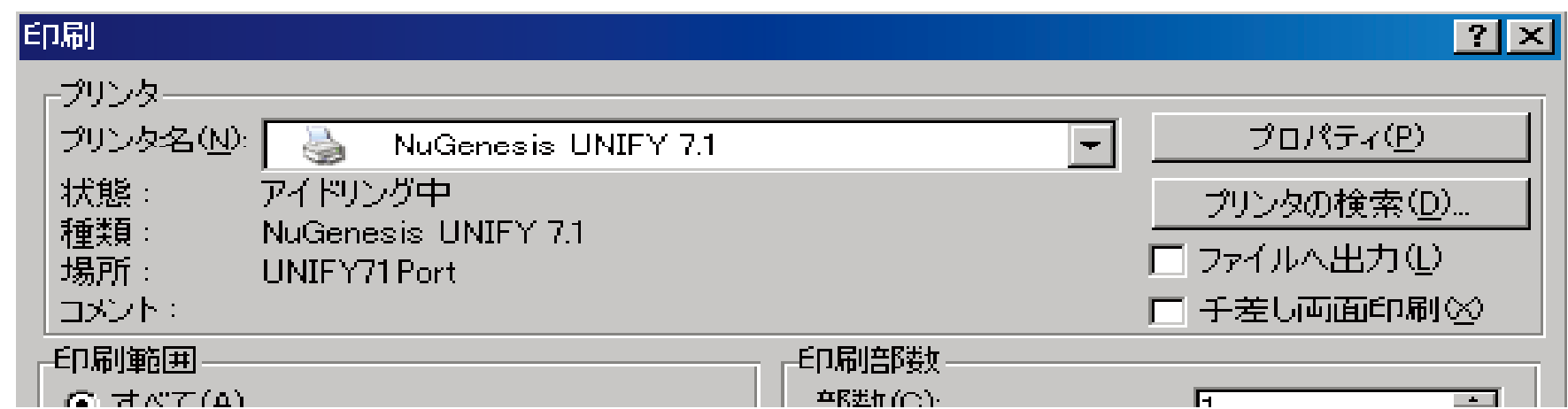


Fig.2 プリントキャプチャ時のプリンタ設定

メタデータの取り込みとインデックス化

プリントキャプチャでもファイルキャプチャでも、データのキャプチャの際には、そのデータを説明するデータ（メタデータ）も同時に取り込み、SDMSデータベースにインデックス化する。メタデータには「サンプル名」「分析日」「分析条件」「分析者」等が挙げられる。どのメタデータをインデックスとして使用するかは、使用者が分析機器毎に指定することができる。

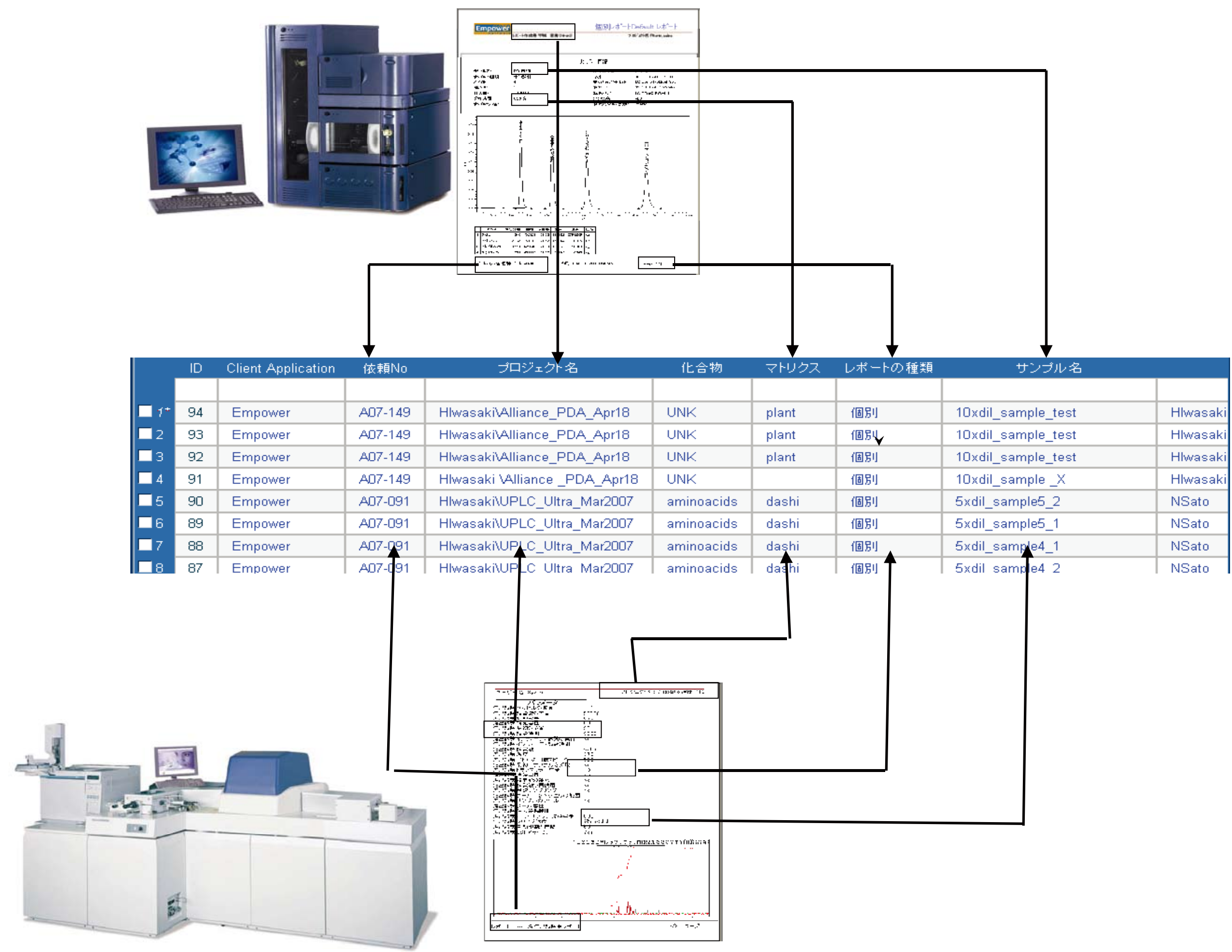


Fig.3 メタデータの抽出とインデックス化概要図

データの検索

メタデータの検索と印刷イメージ中の内容の検索

インデックス化されたメタデータにより、データを容易に検索することができ、長期保存中のデータも所在があらやになることはない。印刷イメージの場合は、メタデータに定義されていない内容をいわゆる全文検索のように検索することもできる。

印刷イメージの元になったファイルデータの検索

更に、印刷イメージの元になったファイルデータを検索することもできるため、過去のデータを再見直し、再解析する必要が生じた場合に、速やかにデータを探し当てることができる。

検索でヒットしたデータファイルは、権限に応じて、任意の場所または元の場所にダウンロードすることができ、元のアプリケーションに戻して再解析等の処理を行う。

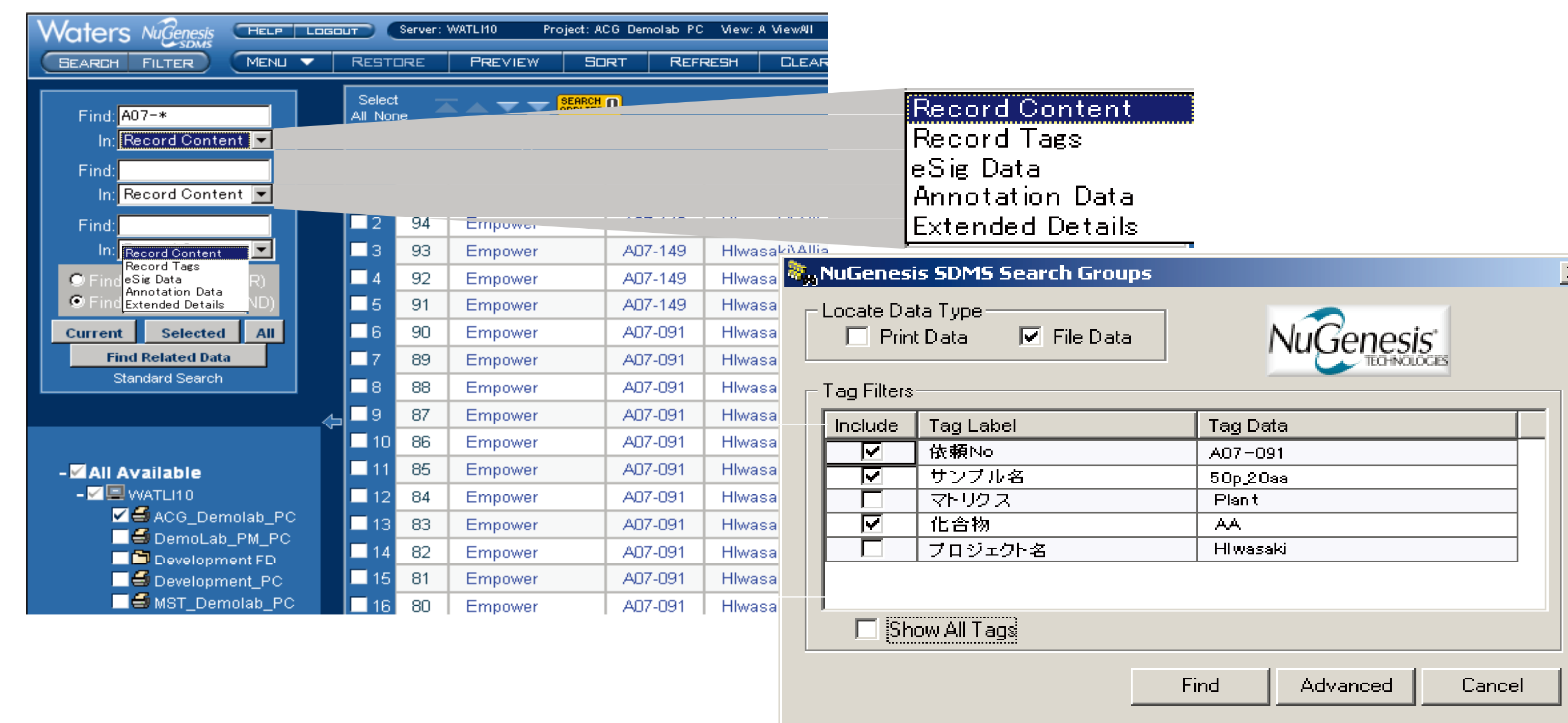


Fig.4 イメージ中の内容の検索と関連するファイルデータの検索

データの二次利用によるワークフローの効率化

環境分析におけるマニュアル処理

◆他のアプリケーションへのデータ転送

ラボのPCはOSのセキュリティパッチやウィルス対策ソフトへの制限から、社内LANに参加していないケースが多い。そのような場合、ワードプロセッサや表計算ソフトへのデータの入力の際、データをUSBメモリやCDにコピーして移動したり、機器からの印刷物を見ながら、その数値データを手入力でワークシートに書き写したりしている。

SDMSに取り込まれたデータは、そのデータを作成したオリジナルのソフトウェアがインストールされていないPCでもデータを表示することができる。そのため、データを外部メディアにコピーすることなく、二次利用することが可能となる。グラフィックとしての利用のみならず、例えば表になっているデータは、その表の形のまま、数値は数値、文字は文字という認識で表計算シートに貼り付けることもでき、手作業が多い、分析後の作業を効率化することができる。更に、他のアプリケーションに転送したデータは、そのデータのSDMS内での出所をハイパーリンクやURLで示すこともできる。これにより、データの信頼性の向上にもつながる。

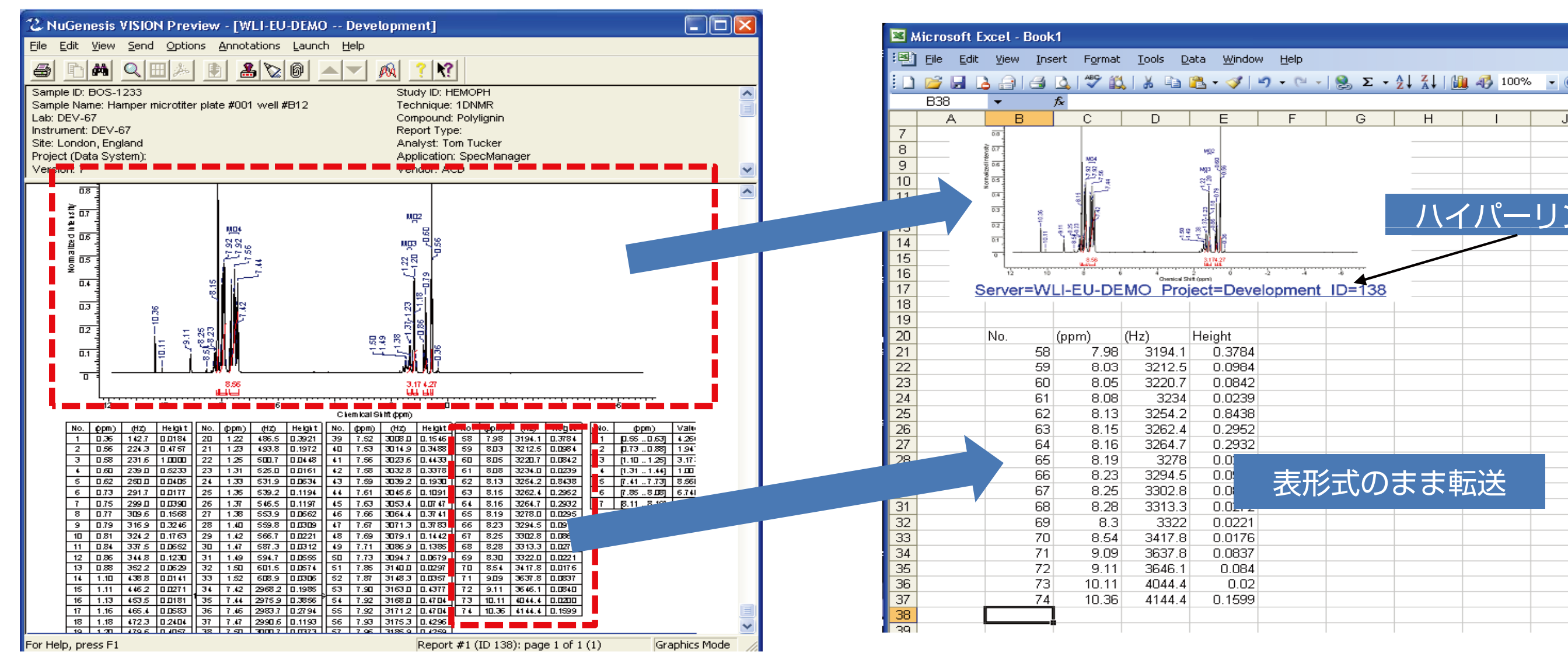


Fig.5 データの二次利用の効率化

◆ラボのPCと事務用PCとの通信

分析機器室でデータを取り込み、社内LANを利用して事務所のPCで報告書を作成するような場合に、社内LANとラボ用LANの間にSDMS用のサーバーを置くことで、分析機器のPCが直接社内LANと接続する必要がなくなり、安全に使用することができるという構築例をFig.6に示した。

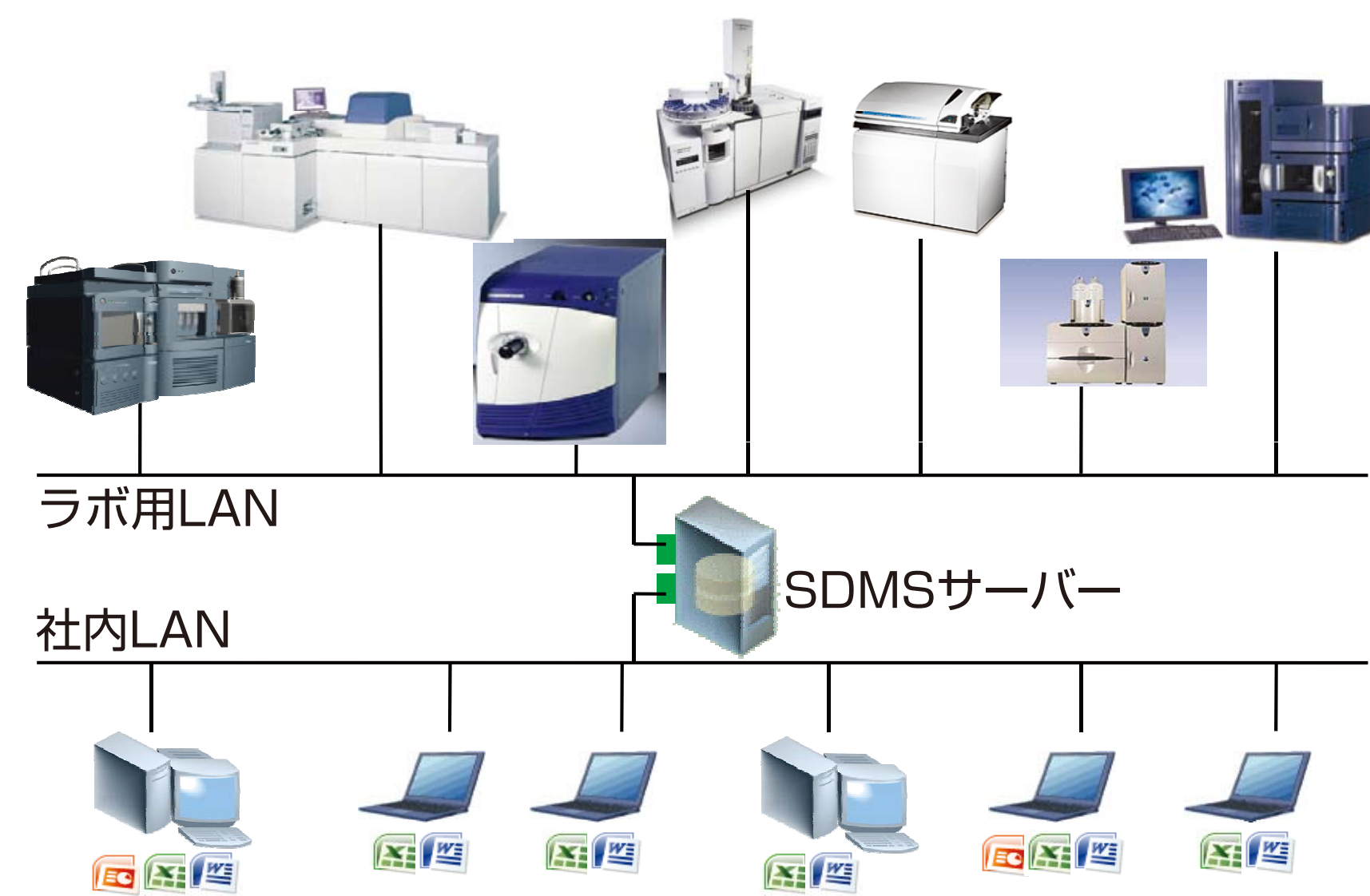


Fig.6 システム構築事例

結論

SDMSを使用することで、バラバラに単体で保存されている各種分析機器のデータを自動的に、電子的に一元管理することが可能となる。そのため、データを紙で保存していた際の保管スペースを大幅に削減することができ、スペースとコストと、紙の管理に必要であった人的コストも削減できる。また、SDMSからデータを二次利用することで、手作業の多いワークフローの効率化が実現できる。