



JAK SZYBKO WYKONYWAĆ ANALIZY I RAPORTY (NP. DLA APR) I UDOSTĘPNIAC JE WŁAŚCIWYM OSOBOM?

Tomasz Demski, StatSoft Polska Sp. z o.o.

Wymogi odnośnie analiz i raportów w przemyśle farmaceutycznym

Przedsiębiorstwa farmaceutyczne i inni producenci stosujący *Dobłą Praktykę Wytwarzania* (ang. *Good Manufacturing Practice*, GMP) zobowiązani są do regularnego wykonywania rozmaitych, często dosyć skomplikowanych analiz statystycznych. Raporty z tych analiz wykorzystywane są m.in. w:

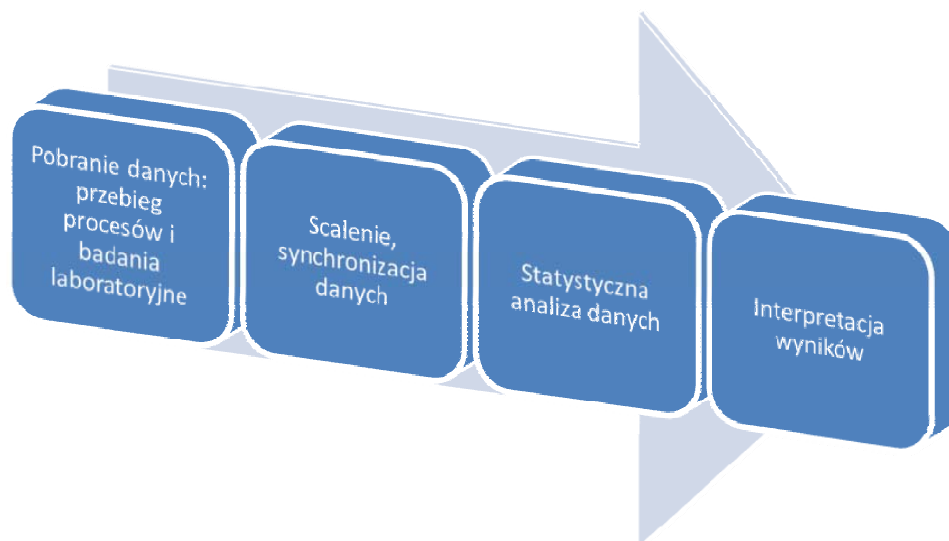
- ◆ okresowych lub rocznych przeglądach produktów (ang. *periodic/annual product review*),
- ◆ badaniu trendów (ang. *trending*),
- ◆ walidacji procesów,
- ◆ projektach PAT.

Na szczególną uwagę zasługuje zagadnienie okresowych przeglądów produktów. Zgodnie z wymogami FDA producenci są zobowiązani do przeglądania i analizy danych o procesie wytwarzania każdego produktu, co najmniej raz do roku. Wykonanie takiego raportu jest więc koniecznością, a ponieważ zazwyczaj wytwarzamy wiele produktów, musimy utworzyć wiele takich raportów.

Na schemacie poniżej widzimy typowy przebieg tworzenia raportu. Zaczynamy od pobrania danych dotyczących procesów i powstających w ich wyniku produktów. Zazwyczaj dane przechowywane są w różnych systemach i przed analizą konieczne jest ich przekształcenie oraz synchronizacja, co wymaga znajomości systemów i narzędzi bazodanowych, a także rzeczywistych relacji czasowych dotyczących zbierania danych, analiz laboratoryjnych itp.

Po uzgodnieniu danych badamy je metodami statystycznymi. Najczęściej korzystamy z takich metod jak: statystyki opisowe, karty kontrolne, wskaźniki zdolności procesu, testy statystyczne porównujące grupy (test t, ANOVA) i wykrywające trend (test Neumanna).

Na koniec potrzebna jest interpretacja uzyskanych wyników, np. stwierdzenie, czy proces jest uregulowany, czy spełnia wymogi jakościowe, czy trend niesie jakieś zagrożenie dla jakości produktu.



Rys. 1. Proces tworzenia raportu.

Tworzenie raportów (bez zastosowania odpowiedniego narzędzia) jest czasochłonne, pracochłonne i często wymaga zaangażowania wielu osób z przygotowaniem z różnych dziedzin. Dzieje się tak dlatego, że na poszczególnych etapach przygotowywania raportu wymagana jest różna wiedza i przygotowanie. Na pierwszych dwóch etapach potrzebujemy informatyka, na trzecim statystyka, a na czwartym statystyka i osoby rozumiejącej sam proces i doskonale orientującej się w wymogach odnośnie produktu.

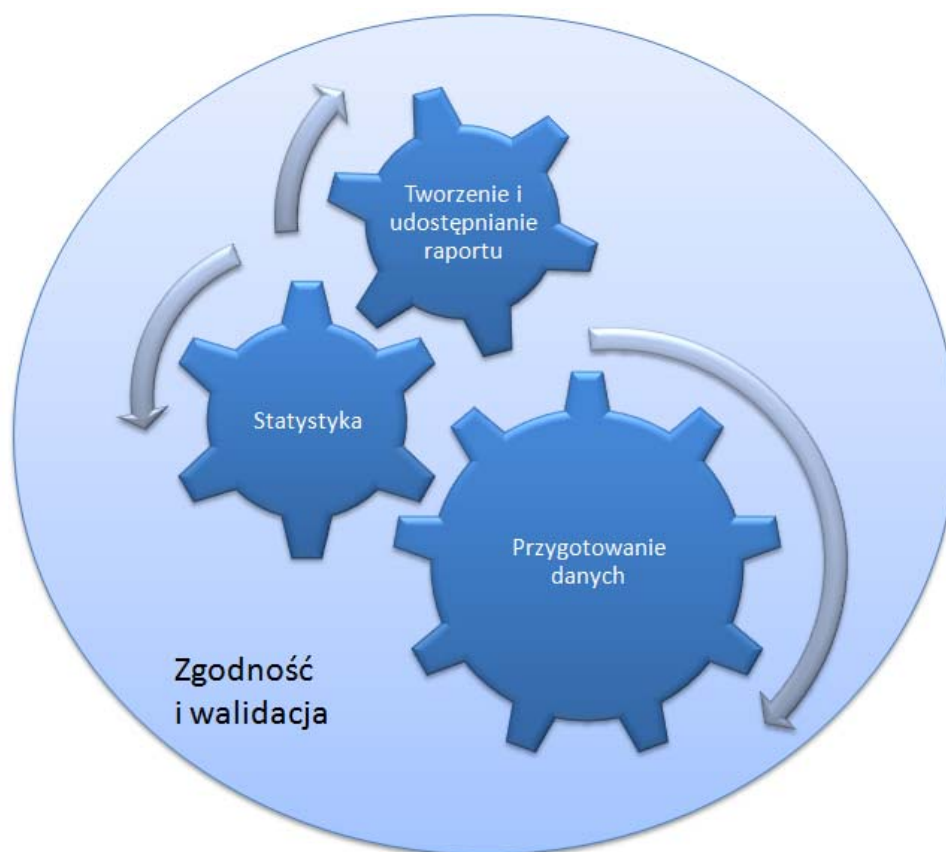
Zdarza się, że etapy 1 i 2 wykonywane są poprzez ręczne przepisywanie danych do arkusza, a potem agregowanie danych z użyciem ręcznie wpisywanych formuł.

Jak widać przygotowanie raportu może być uciążliwe i wymagać zaangażowania wielu osób o wysokich kwalifikacjach. Jednocześnie duża część pracochłonnych zadań sprowadza się do powtórzenia wykonanych wcześniej operacji: jeśli raz wykonamy operację scalenia danych, to wystarczy potem powtórzyć te operacje z niewielkimi zmianami (np. z innym warunkiem na identyfikator produktu lub okres, którego dotyczy analiza). Dodatkowo jeżeli operacje są wykonywane ręcznie przez człowieka, to stosunkowo łatwo jest o błąd. Z tego względu bardzo korzystne jest zbudowanie systemu wspomagającego tworzenie i upowszechnianie raportów.

Wdrożenie systemu wspierającego tworzenie raportów nie tylko zmniejsza czasochłonność i pracochłonność, ale również umożliwia śledzenie całego procesu przygotowania raportu i łatwe odtworzenie go w przyszłości.

Jeśli system jest odpowiednio przygotowany (a tak jest w przypadku *STATISTICA Enterprise*), to walidacja automatycznego tworzenia raportów jest zdecydowanie łatwiejsza i mniej czasochłonna niż w przypadku ręcznego tworzenia raportów.

Ponadto w dużych organizacjach, mających wiele zakładów, bardzo pożyteczna jest standaryzacja i łatwość przeniesienia rozwiązania z jednego zakładu do drugiego. W takim środowisku przydatny bywa zdalny dostęp do systemu: przez interfejs WWW lub usługi terminalowe.



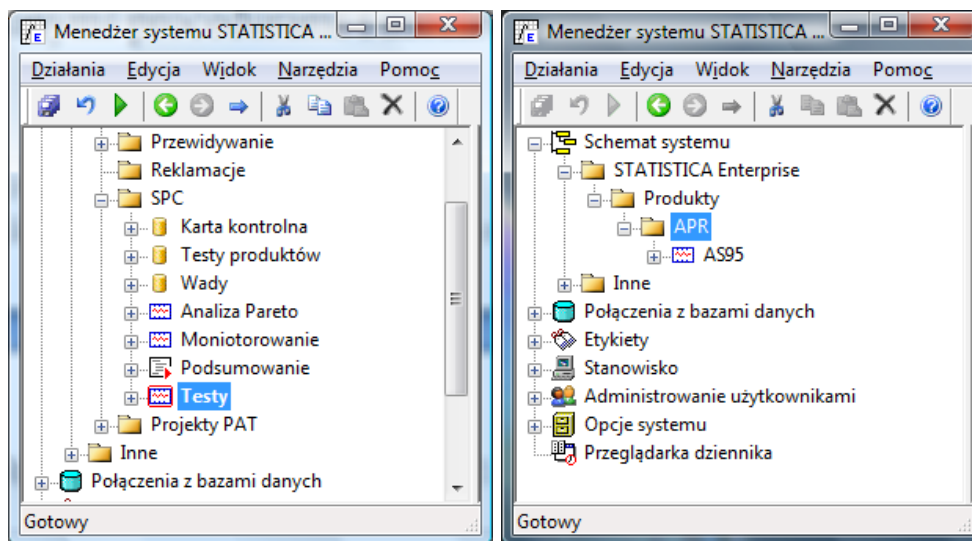
Rys. 2. Funkcje systemu wspomagającego tworzenie raportów.

System do wspierania tworzenia raportów powinien realizować trzy typy funkcji (rys. 2):

1. **Przygotowanie danych:** pobieranie danych z różnych źródeł, przekształcanie danych, agregacja i synchronizacja, sprawdzania i czyszczenie danych (etapy 1 i 2 na rys. 1).
2. **Statystyka** (etap 3 na rys. 1): system musi wykonywać wszystkie potrzebne procedury statystyczne, a dostęp do analiz powinien być łatwy i nie wymagać skomplikowanych operacji.
3. **Tworzenie i udostępnianie raportu:** wynikowy dokument powinien być drukowany lub zapisywany w standardowym formacie i zabezpieczony przed zmianami zgodnie z wymaganiami FDA CFR 21 part 11.

Cały system – zarówno jako oprogramowanie, jak i jego konkretne wdrożenie – powinien zapewniać zgodność z ustalonymi procedurami i możliwość walidacji.

Jak powinna wyglądać praca z systemem z punktu widzenia użytkownika? Zaczynamy od uwierzytelnienia w systemie. Na podstawie uprawnień użytkownika system udostępnia mu odpowiednią listę obiektów: raportów, szablonów analiz i pobierania danych. Użytkownik wybiera raport, w razie potrzeby podaje dodatkowe kryteria (np. wybiera konkretny produkt, serię, okres itp.) i dostaje wyniki analizy, które może interpretować.



Rys. 3. Przykładowe listy raportów dla różnych użytkowników.

Nieważne, jak skomplikowane operacje wykonywane są w tle, aby uzyskać finalny raport: użytkownik otrzymuje go po kilku kliknięciach i, nie marnując czasu na żmudne, powtarzalne operacje, może przystąpić do wyciągania wniosków na podstawie wyników analizy.

Wdrażanie systemu wspierającego wykonywanie analiz i raportowanie

StatSoft ma wieloletnie doświadczenie w tworzeniu systemów wspierających raportowanie dla przemysłu farmaceutycznego. Z naszych doświadczeń wynika, że właściwy jest przedstawiony poniżej sposób realizowania takich projektów. Oczywiście wszystkie poniższe etapy są dokumentowane i realizowane zgodnie z zasadami walidacji systemów komputerowych.

I. Zbadanie, jak raporty są tworzone i wykorzystywane obecnie

Na tym etapie decydujemy, czy po prostu zautomatyzujemy istniejącą procedurę, czy też w jakiś sposób ją uporządkujemy i udoskonalimy. Warto zauważyć, iż proste przeniesienie istniejących procedur niekoniecznie jest najlepszym rozwiązaniem: warto wykorzystać wprowadzanie systemu jako okazję do przejrzenia istniejących procedur i udoskonalenia ich. W szczególności ponieważ system bez angażowania zasobów może wykonywać nawet bardzo skomplikowane analizy, można zrezygnować z pewnych uproszczeń i ograniczeń, które były konieczne przy ręcznym tworzeniu raportów.

Na tym etapie określamy również stopień automatyzacji tworzenia raportów w zależności od rodzaju zagadnienia i potrzeb przedsiębiorstwa. W szczególności raport może być tworzony całkowicie automatycznie i zapisywany w pliku PDF albo automatycznie wykonywane jest tylko przygotowanie danych i wstępna analiza, a człowiek decyduje, co dalej, np. wykonuje dalsze analizy i przygotowuje raport ze swoimi wnioskami.



Czasami w jednej organizacji stosuje się więcej niż jedną procedurę; w takim wypadku należy wybrać jedną z nich. Bardzo często pewne decyzje czy stwierdzenia w raportach były umieszczane arbitralnie na podstawie opinii osoby wykonującej analizę. Przykładowo analityk mógł stwierdzać, że trend jest „mały” albo „duży”, określać, że jakaś obserwacja jest odstająca. W takim wypadku należy opracować odpowiednie reguły postępowania, np. zaproponować właściwy test statystyczny sprawdzający, czy obserwacja jest odstająca.

Na tym etapie należy wziąć pod uwagę, jak ewentualne zmiany w procedurach raportowania wpływają na działanie przedsiębiorstwa, procesy decyzyjne itp.

II. Źródła danych

Ten etap to głównie zbadanie istniejących źródeł danych, z których będziemy korzystać przy tworzeniu raportów. Określamy tutaj, które dane przechowywane są w poszczególnych źródłach, w jaki sposób można je odczytać oraz jakie reguły obowiązują dla danych i jak przekładają się one na rzeczywistość.

Na tym etapie warto przeprowadzić audyt danych, sprawdzić ich poprawność i zachowanie pewnych zasad, reguł kodowania itp. w całym okresie, dla którego będą wykonywane analizy (np. czy wyniki tego samego testu nie były kodowane w różny sposób w różnych latach).

Musimy tu także określić sposób scalania, synchronizacji i agregacji danych pochodzących ze wszystkich źródeł danych.

Niezwykle ważne jest staranne i dokładne przeprowadzenie tego etapu oraz udokumentowanie go. Absolutnie nie należy oszczędzać czasu poświęconego na ten etap!

III. Ustalenie zawartości i formatu raportu

Tu określamy analizy statystyczne, które będą wykonywane oraz to, jakie konkretne tabele i wykresy trafią do raportu. Oczywiście należy tu ustalić również formatowanie raportu, dodatkowe składniki umieszczane w raporcie (np. logo, schemat procesu).

IV. Wdrożenie

Poprzednie etapy dostarczają specyfikacji określających zawartość i sposób tworzenia raportu. Na etapie IV realizujemy to w postaci systemu informatycznego.

V. Testowanie i akceptacja przez użytkowników

Po stworzeniu systemu należy sprawdzić jego działanie, aby upewnić się, że nie tylko działa poprawnie, ale również spełnia wymogi i oczekiwania użytkowników.



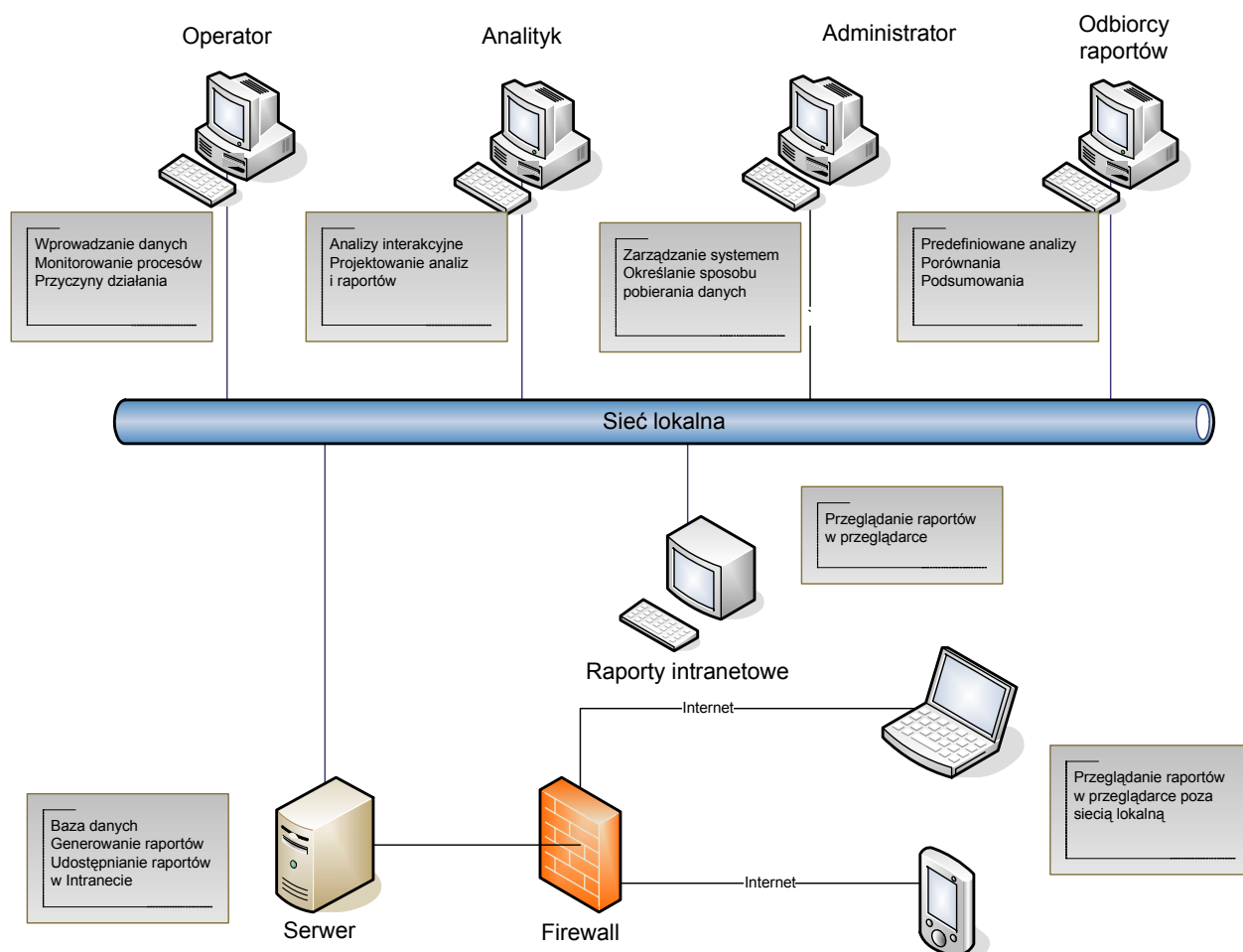
Rozwiązanie: **STATISTICA Enterprise**

STATISTICA Enterprise to specjalna wersja *STATISTICA* przeznaczona do wykorzystania przez zespoły ludzi, pracujących wspólnie i potrzebujących wygodnego, wydajnego i bezpiecznego środowiska do analizy danych i raportowania. Ponieważ *STATISTICA Enterprise* łączy w sobie narzędzia do przeprowadzenia całego procesu przekształcania surowych danych w użyteczną wiedzę nazywamy go zintegrowanym systemem analizy danych.

W skład systemu wchodzi:

- ◆ centralna baza danych,
- ◆ narzędzia analityczne *STATISTICA* (statystyka, wizualizacja i *data mining*),
- ◆ generator raportów,
- ◆ system bezpiecznego przechowywania dokumentów zgodny z wymaganiami FDA CFR 21 part 11 (ta część jest opcjonalna).

Poglądowy schemat działania systemu i główne role użytkowników przedstawia rys. 4.



Rys. 4. Schemat *STATISTICA Enterprise*.



Rdzeniem zintegrowanego systemu jest centralna baza danych. Możemy w niej przechowywać:

- ◆ wyniki pomiarów właściwości procesu i towarzyszące im dodatkowe informacje (etykiety), takie jak np. identyfikator osoby wprowadzającej konkretne obserwacje,
- ◆ specyfikacje sposobu pobierania danych z zewnętrznych źródeł danych (baz danych, plików, sterowników automatyki przemysłowej i innych),
- ◆ scenariusze wprowadzania danych (mierzone właściwości, dopuszczalne wartości, urządzenia pomiarowe itd.),
- ◆ scenariusze analizy danych i szablony raportów,
- ◆ uprawnienia dostępu użytkowników do poszczególnych funkcji i danych systemu.

Warto zwrócić uwagę na kluczową zaletę przechowywania powyżej wymienionych danych w centralnej bazie danych. Dzięki temu mamy zapewnioną dostępność i zgodność danych, a potrzebne ustawienia i zmiany wystarczy wykonać raz. Jeśli przykładowo zmieniamy szablon raportu, np. dodając do niego nowy wykres, to wystarczy wykonać to raz i zapisać nowe ustawienia w bazie danych. Zmiana ta jest potem automatycznie przenoszona na wszystkie stanowiska, nie trzeba obchodzić każdego komputera i zmieniać na nim ustawień. Ponadto dane wprowadzone na jednym stanowisku są dostępne dla użytkowników na innych stanowiskach praktycznie od razu.

W skład *STATISTICA Enterprise* wchodzi narzędzia analizy danych z rodziny *STATISTICA*. Zakres metod statystycznych, wizualizacji i *data mining* zawartych w systemie zależy od potrzeb danego użytkownika. Z narzędzi analitycznych korzystamy na kilka sposobów (przy czym wiele osób nie tylko nie musi umieć obsługiwać programu *STATISTICA*, ale nawet wiedzieć, że jest źródłem uzyskiwanych informacji), są to:

- ◆ interakcyjne wykonywanie analiz (tak jak w zwykłej *STATISTICA*), projektowanie szablonów analiz i raportów dla innych użytkowników,
- ◆ korzystanie z przygotowanych scenariuszy analizy w środowisku *STATISTICA*,
- ◆ przeglądanie raportów (np. w przeglądarce internetowej lub PDF),
- ◆ odbiór automatycznie generowanych przez system powiadomień (alarmów) poprzez e-mail, sms itp.

Z praktycznego punktu widzenia warto zwrócić uwagę na aspekt łatwości korzystania z wyników analiz w *STATISTICA Enterprise*. Otóż w praktyce najczęściej większość osób nie wykonuje analiz, a jest odbiorcą ich wyników. Takim osobom dużo łatwiej jest korzystać z szablonów analizy przygotowanych przez specjalistów od analizy danych i wdrożonych w *STATISTICA Enterprise*, niż własnoręcznie określać i uruchamiać analizę w zwykłym (tj. niewbudowanym w system) programie do analizy danych.

Generator raportów służy do uruchamiania całej analizy w zadanych odstępach czasu i przekształcania jej wyników na powszechnie zrozumiały format (np. PDF) – tak aby mógł z nich korzystać praktycznie każdy. Układ i formatowanie raportu określane jest podczas



jego tworzenia przez użytkownika. Oprócz wyników analizy (tabel, wykresów i informacji tekstowych) raport może być wzbogacony:

- ◆ dodatkowymi danymi: datą i czasem utworzenia, nazwą użytkownika, który uruchomił raport, wersją raportu itp.,
- ◆ tekstami (np. tytułem) i obiektami graficznymi (np. logo, schemat produkowanego elementu, którego dotyczy analiza).

Generator raportów jest w stanie tworzyć raport spełniający nawet bardzo wygórowane wymagania dotyczące jego zawartości, ale również bezpieczeństwa danych i monitorowania wszelkich zmian. Przykładowy raport przedstawia rys. 5.



Rys. 5. Przykładowy raport *STATISTICA Enterprise*.

Generator raportów może w praktyce okazać się niezwykle pożyteczny. Standaryzacja i automatyzacja tworzenia zestawień pozwala spożytkować czas analityków na twórcze



wykorzystanie wiedzy o procesie i jego doskonalenie, zamiast na wykonywania żmudnych i powtarzalnych czynności. Zastosowanie takiego systemu oszczędza tysiące godzin pracy rocznie (przykładowo przygotowanie zestawienia podsumowującego w jednej z firm farmaceutycznych przed wdrożeniem systemu bazującego na *STATISTICA Enterprise* trwało 40 godzin, a po wdrożeniu 5 min!). Oprócz bezpośredniej oszczędności czasu i środków, zastosowanie systemu zmniejsza ryzyko popełnienia błędu i daje bardziej dokładny wgląd w proces, poprzez dostępność bardziej złożonych metod analitycznych.

Samo przechowywanie wszystkich danych w centralnej bazie danych oraz system uprawnień użytkowników zapewnia bezpieczeństwo danych, zwłaszcza że *STATISTICA Enterprise* dysponuje funkcją zapisu działań użytkowników systemu. Jednak w pewnych zastosowaniach jest to za mało – potrzebujemy jeszcze:

- ◆ bezpiecznego przechowywania dokumentów tworzonych przez system,
- ◆ zatwierdzania dokumentów,
- ◆ kontroli i zapisu zmian (innymi słowy przechowywania wszystkich wersji dokumentów, szablonów analiz i raportów).

W takim wypadku potrzebna jest specjalna wersja *STATISTICA Enterprise* zintegrowana ze *STATISTICA Document Management System* (opis tego systemu znajduje się na stronie <http://www.statsoft.pl/dms.html>). Ta wersja *STATISTICA Enterprise* jest zgodna z wymaganiami FDA CFR 21 part 11 i nosi nazwę *Compliance Edition*.