

Jakub Badowski
Instytut Nafty i Gazu, Kraków

Zunifikowany format danych jako kamień milowy w procesie gromadzenia informacji o awariach na gazociągach przesyłowych

Wprowadzenie

Wszelkie działania podejmowane przez człowieka wiążą się z pewnym ryzykiem. Tak samo jest w przypadku eksploatacji gazociągów, czyli transportu gazu ziemnego rurociągami. Nazywamy to ryzykiem eksploatacyjnym gazociągów. Gazociąg przesyłowy, jako obiekt strategiczny, aby mógł być uznany za bezpieczny powinien zostać objęty Systemem Zarządzania Ryzykiem (SZR). Integralną częścią SZR jest szacowanie ryzyka eksploatacyjnego (rysunek 1). Aby oszacować ryzyko eksploatacyjne gazociągu w sposób liczbowy (ilościowe szacowanie ryzyka) potrzebne są dane statystyczne dotyczące awarii, które wydarzyły się w przeszłości na podobnych gazociągach [2]. W Stanach Zjednoczonych takie dane gromadzi się od wczesnych lat siedemdziesiątych dwudziestego wieku. Służy do tego odpowiednio zaprojektowana i eksploatowana baza danych o zasięgu ogólnokrajowym [12]. Podobne doświadczenia w tym zakresie mają nasi sąsiedzi (głównie z Europy Zachodniej), zrzeszeni w organizacji pod nazwą EGIG [5].

Nasuwa się pytanie, jak radzą sobie z tym problemem krajowi operatorzy gazociągów. Czy gromadzą takie dane w sposób zgodny ze standardami europejskimi? A jeśli tak, to czy korzystają z tej wiedzy w kontekście Systemu Zarządzania Ryzykiem? Firmy zajmujące się przesyłem gazu w Polsce niechętnie dzielą się takimi informacjami, a literatura fachowa milczy na ten temat. Według wiedzy

autora, nawet jeśli operatorzy polskich gazociągów tego typu dane gromadzą, to tylko w obrębie własnego przedsiębiorstwa. Ponieważ do awarii na gazociągach dochodzi stosunkowo rzadko, bardzo trudno jest jednemu operatorowi zebrać wystarczającą ilość danych, aby przedstawiały one wartość z punktu widzenia statystyki.

Pojawia się więc zasadniczy problem, który można sformułować następująco: bez odpowiedniej ilości ustrukturyzowanych danych o awariach na gazociągach nie jest możliwe wdrożenie SZR. Wniosek nasuwa się sam: potrzebna jest ogólnopolska baza danych o awariach na gazociągach. Taka baza została zaproponowana przez autora w 2010 roku pod nazwą *System komputerowy przeznaczony do gromadzenia i analizy danych o awariach na gazociągach przesyłowych* [1]. Zostały też podjęte próby wdrożenia tego projektu, podczas których zaistniał problem techniczny zawarty w następującym pytaniu: jaki podmiot w kraju byłby odpowiedzialny za utrzymywanie tej bazy danych i gdzie fizycznie miałyby się ona znajdować? Aby rozwiązać tę kwestię, autor postanowił odstąpić na jakiś czas od założeń gromadzenia informacji w centralnej bazie danych, wprowadzając fazę przygotowawczą. Na tym

etapie operatorzy gazociągów mieliby gromadzić dane we własnym zakresie, ale w sposób ujednoczony i zgodny ze standardami. W tym celu autor publikacji opracował pewną strukturę danych, opartą na języku znaczników XML.



Rys. 1. Ogólny schemat Systemu Zarządzania Ryzykiem [2]

Język XML i rodzina technologii z nim związanych

XML (ang. *Extensible Markup Language*) to standard przeznaczony do reprezentowania danych w sposób ustrukturyzowany. Jest także jednym z najczęściej stosowanych formatów wymiany informacji pomiędzy różnymi systemami, niezależnie od platformy [7]. O standardzie XML można powiedzieć, że jest językiem, gdyż służy do przekazywania pewnych treści. Z całą pewnością nie jest to natomiast język programowania. W rozwinięciu nazywa się go językiem znaczników, z racji tego, że w kodzie XML do oznaczania elementów służą znaczniki, czyli ciągi znaków (zwykle jakieś słowa) zawarte w nawiasach trójkątnych (<) i (>). Każdy element musi zaczynać się znacznikiem początku elementu, np. <autor>, oraz kończyć znacznikiem końca elementu, np. </autor>. Wyjątkiem od tej reguły są elementy puste, które wyglądają nieco inaczej, np. <element/>. Znaczniki mogą zawierać także atrybuty złożone z nazwy oraz wartości atrybutu. Są one definiowane w znaczniku początku elementu w następujący sposób: <tytul oprawa="twarda"></tytul>. Język XML jest bardzo elastyczny, ponieważ zbiór znaczników nie jest w żaden sposób ograniczony. Programista ma pełną swobodę w definiowaniu własnych znaczników, w zależności od swoich potrzeb. Przykładowy fragmentu kodu XML prezentuje rysunek 2.

```
<?xml version="1.0" encoding="UTF-8"?>
<ksiazki>
  <ksiazka>
    <tytul oprawa="twarda">Proces</tytul>
    <autor>Franz Kafka</autor>
  </ksiazka>
  <ksiazka>
    <tytul>Omerta</tytul>
    <autor>Mario Puzo</autor>
  </ksiazka>
</ksiazki>
```

Rys. 2. Przykładowy dokument XML

Konstrukcja języka XML pozwala na odizolowanie danych od sposobu ich prezentacji. Dokumenty XML nie zawierają żadnych informacji formatujących. Są zorientowane na zawartość, nie na prezentację. Aby zawarte w nich dane sformatować, należy skorzystać z arkuszy stylów. Realizuje się to w taki sposób, że w dokumencie XML umieszcza się referencję do pliku zawierającego informacje o sposobie wyświetlania danych.

Obecnie język XML ma bardzo wiele różnych zastosowań. Wykorzystuje się go do opisu zasobów w sieci, zapisu

informacji multimedialnej (np. grafiki wektorowej), elektronicznej wymiany danych (EDI) czy konfiguracji oprogramowania. Znalazł też szerokie zastosowanie w protokołach komunikacyjnych. Standard XML jako format wymiany danych cieszy się coraz większym zainteresowaniem, szczególnie za sprawą spopularyzowanych w ostatnich latach usług sieciowych (ang. *Web Services*) [10].

W niniejszej publikacji zostało opisane zastosowanie języka XML do przekazywania informacji w społeczności specjalistów z konkretnej dziedziny, mające charakter środowiskowego standardu. Jako bardziej znane przykłady takiego zastosowania języka XML można tu wymienić: MathML (*Mathematical Markup Language*) – do zapisu wzorów matematycznych, czy CML (*Chemical Markup Language*) – format zapisu informacji o związkach chemicznych.

Co ciekawe, XML jako język nie wnosi żadnych rewolucyjnych rozwiązań. Koncepty języków znakowania i hierarchicznych struktur danych znane są od bardzo dawna. To co jest w XML istotnie nowe, to szeroka akceptacja standardów oraz duża dostępność otwartych narzędzi. Ich zastosowanie jest niezbędne, aby w pełni wykorzystać możliwości języka XML. Standardy z tej rodziny opisują sposób formułowania pewnych typowych informacji, sposoby przetwarzania oraz przeszukiwania dokumentów XML itp. [11].

XML Schema

Spośród technologii, które znalazły zastosowanie w niniejszej pracy, na szczególną uwagę zasługuje XML Schema, który określa sposób deklarowania składni dokumentów XML, a w konsekwencji służy do ich walidacji. Za pomocą tego schematu określa się reguły budowy dokumentu. Można w nich ustalać nazwy, następstwo i sposób zawierania się elementów, atrybuty oraz typy danych, jakie są w nich zawarte. Istnieje możliwość tworzenia własnych typów bądź skorzystania ze słownika typów elementarnych. Za pomocą XML Schema można określać warunki poprawności danych, między innymi przez definiowanie przedziałów dopuszczalnych wartości bądź dopasowywanie wartości do wyrażenia regularnego. Dzięki schematom możliwe jest tworzenie nowych „dialektów” XML, czyli sformalizowanych struktur danych. Wykorzystanie schematów bardzo upraszcza przetwarzanie dokumentów XML w programach, gdyż proces ich analizy można ograniczyć jedynie do dokumentów poprawnych. Kontrolę poprawności przejmuje biblioteczna procedura walidująca, która korzysta ze schematu.

XML Schema nie jest jedyną technologią tego typu, choć uważa się ją za najbardziej popularną. Jako alternatywne rozwiązania można w tym miejscu wymienić jeszcze standard DTD (ang. *Document Type Definition*) oraz RelaxNG [4].

XSLT

Można śmiało powiedzieć, że XSLT (ang. *Extensible Stylesheet Language Transformations*) jest jednym z najważniejszych składników rodziny technologii XML. To język ogólny, który może być wykorzystywany przez różne aplikacje w różnych systemach informatycznych [8]. Jest przeznaczony do przekształcania struktury dokumentu XML na dowolny inny format zgodny ze składnią XML, a także na czysty tekst, format HTML,

PDF i mnóstwo innych rodzajów dokumentów. Świetnie nadaje się do prezentowania danych pochodzących z dokumentów XML w przeglądarkach internetowych. Może być również użyty do konwersji pomiędzy różnymi formatami danych. Stosowanie rozwiązań XSLT pozwala na przekształcenie kodu XML do postaci wygodnej do oglądania. Dzięki oddzieleniu treści od warstwy prezentacji istnieje możliwość utworzenia oddzielnych arkuszy stylów dla przeglądarek internetowych, urządzeń typu PDA czy telefonów komórkowych, przy czym – w żadnym z wymienionych przypadków nie zachodzi konieczność wprowadzania zmian w samej strukturze danych. Autor publikacji wykorzystuje język XSLT w procesie tworzenia raportów oraz do prezentacji wybranych informacji w Internecie.

Opis formatu danych

Po teoretycznym wprowadzeniu przystąpiono do opisu właściwego formatu danych o awariach na gazociągach, opartego na języku znaczników XML. Ponieważ do analizy statystycznej zdarzeń awaryjnych potrzebne są dwa odrębne zbiory danych: inwentarz sieci oraz rejestr zdarzeń [1] – przygotowano odpowiednio dwa formaty:

- format opisujący jednostkę gazociągu (odcinek). Założono, że sieć składa się z odcinków zdefiniowanych przez operatora. Każdy z takich odcinków jest obiektem posiadającym swój unikatowy numer identyfikacyjny (ID). Zdefiniowany odcinek powinien być w miarę możliwości jednolity pod względem parametrów technicznych, aby było możliwe opisanie go pewnym określonym zestawem informacji.
- format opisujący zdarzenie (awarię) – dotyczy zdarzeń awaryjnych zaistniałych na sieci gazowej, które spełniają określone kryteria.

Jednostki gazociągów oraz zdarzenia awaryjne zostały sklasyfikowane według ustalonych zasad, aby było możliwe utrzymanie określonej struktury dokumentu. Format opisujący jednostkę gazociągu (odcinek) składa się z następujących elementów:

- id – identyfikator odcinka (unikatowy klucz). Praca ta nie narzuca operatorom sposobu identyfikowania swoich gazociągów. Niemniej jednak każda jednostka gazociągu musi mieć w danym systemie swój unikatowy (niepowtarzalny) identyfikator;
- opis (ang. *description*) – dodatkowe informacje pomocne przy identyfikacji odcinka. Opis nie jest obowiązkowy;
- długość odcinka (ang. *length of pipeline*) – wartość podawana w metrach;

- rodzaj sieci (ang. *type of network*) – może przyjmować dwie wartości:
 - sieć przesyłowa (ang. *transit gas pipeline*),
 - sieć dystrybucyjna (ang. *distribution gas pipeline*);
- materiał (ang. *material*) – niesie ze sobą informacje, z jakiego materiału zbudowany jest odcinek gazociągu. Może przyjmować następujące wartości:
 - stal (ang. *steel*),
 - PE (Polietylen) (ang. *Polyethylene*),
 - żeliwo (ang. *Cast iron*),
 - inny (ang. *other*);
- średnica nominalna gazociągu (DN) (ang. *diameter class*) – wartość podawana w [mm];
- grubość ścianki (ang. *nominal wall thickness*) – wartość podawana w [mm];
- rok budowy (ang. *year of construction*);
- ciśnienie M.O.P. (maksymalne ciśnienie robocze) (ang. *maximum operating pressure*) – wartość podawana w barach;
- powłoka ochronna (ang. *coating*) – mówi o tym, czy na danym odcinku znajduje się powłoka ochronna. Może przyjmować dwie wartości:
 - tak (ang. *yes*),
 - nie (ang. *no*);
- ochrona katodowa (ang. *Cathodic protection*) – mówi o tym, czy dany odcinek ma zapewnioną ochronę katodową. Może przyjmować dwie wartości:
 - tak,
 - nie;
- grubość przykrycia gazociągu (ang. *cover class*) – podawana w [cm].

```

<?xml version="1.0" encoding="utf-8" ?>
<pipes>
  <pipe id="0001">
    <description></description>
    <pipe-length unit="m">4190</pipe-length>
    <type-of-network>transit gas pipeline</type-of-network>
    <material>steel</material>
    <diameter-class unit="mm">150</diameter-class>
    <nom-wall-thickness unit="mm">7</nom-wall-thickness>
    <year-of-construction>1985</year-of-construction>
    <max-operating-pressure unit="bar">15</max-operating-pressure>
    <coating>yes</coating>
    <cathodic-protection>yes</cathodic-protection>
    <cover-class unit="cm">80</cover-class>
  </pipe>
</pipes>

```

Rys. 3. Struktura dokumentu XML opisująca jednostkę gazociągu, z przykładowymi wartościami

Dane charakteryzujące zdarzenie awaryjne składają się z trzech bloków:

a) dane identyfikacyjne (ang. *identification data*):

- id – identyfikacja zdarzenia (unikatowy klucz),
- data zdarzenia (ang. *date of incident*),
- czas zdarzenia (ang. *time of incident*),
- id odcinka gazociągu (ang. *segment id*) – identyfikator odcinka, na którym doszło do awarii;

b) określenie zdarzenia (ang. *incident attributes*):

- wielkość wycieku (ang. *size of leak*):
 - nieszczelność (ang. *pinhole*),
 - pęknięcie (ang. *hole*),
 - rozerwanie (ang. *rupture*),
 - inne (ang. *other*);
- przyczyna zdarzenia (ang. *cause*):

– czynniki zewnętrzne (ang. *external interference*),

– korozja (ang. *corrosion*),

– wady konstrukcyjne/materiałowe (ang. *construction defects/material failures*),

– ruchy gruntu (ang. *ground movement*),

– niewłaściwe przyłączenie (ang. *hot tap made by error*),

– inne;

• ciśnienie chwilowe w momencie zdarzenia (ang. *transient pressure*) – wartość podawana w barach;

c) skutki zdarzenia (ang. *effects of incident*):

• określenie rozmiaru uszkodzenia (ang. *extent of damage*):

– punktowe,

```

<?xml version="1.0" encoding="utf-8" ?>
<incident id="21">
  <identification-data>
    <date>2012.05.03</date>
    <time>20:00</time>
    <segment-id>0001</segment-id>
  </identification-data>
  <incident-attributes>
    <size-of-leak>pinhole</size-of-leak>
    <cause>corrosion</cause>
    <transient-pressure unit="bar">25</transient-pressure>
  </incident-attributes>
  <effects>
    <extent-of-damage>other</extent-of-damage>
    <ignition>no</ignition>
    <explosion>no</explosion>
    <fatalities>0</fatalities>
    <material-damage symbol="PLN">50000</material-damage>
  </effects>
</incident>

```

Rys. 4. Struktura dokumentu XML opisująca pojedyncze zdarzenie

- szczelina,
- wzdłużne,
- przekrojowe,
- inne;
- czy nastąpił zapłon (ang. *ignition*):
 - tak,
 - nie;
- czy nastąpił wybuch (ang. *explosion*):
 - tak,
 - nie;
- ofiary w ludziach (ang. *fatalities*) – liczba ofiar śmiertelnych;
- straty materialne (ang. *material damage*) – liczone w [zł].

Podsumowanie

Gromadzenie danych o awariach na gazociągach jest procesem ciągłym. Jeśli się go praktykuje dostatecznie długo, dokładnie i systematycznie, owocuje bezcenną wiedzą, która użyta w odpowiedni sposób przekłada się na wymierne korzyści, a przede wszystkim wpływa na poprawę bezpieczeństwa. Świadczą o tym najlepiej – wspomniane już na wstępie – doświadczenia krajów Europy Zachodniej oraz Stanów Zjednoczonych. Ważne jest, aby gromadząc dane o awariach, trzymać się określonych standardów.

Zdaniem autora, świetnie nadaje się do tego zadania format danych zaprezentowany w niniejszej pracy, oparty na języku znaczników XML oraz kilku narzędziach z jego rodziny, rekomendowanych przez światową organizację W3C [6]. Prezentowane rozwiązanie jest w zamyśle autora etapem przygotowawczym do wdrożenia ogólnopolskiej bazy danych o awariach na gazociągach. Jego zastosowanie znacznie ułatwi proces pozyskania wiedzy statystycznej wykorzystywanej w Systemach Zarządzania Ryzykiem.

Literatura

- [1] Badowski J.: System komputerowy przeznaczony do gromadzenia i analizy danych o awariach na gazociągach przesyłowych. „Nafta-Gaz” 2010, nr 12, s. 1132-1137.
- [2] Dietrich A.: Problem gromadzenia i analizy danych o awariach na gazociągach wysokiego ciśnienia a System Zarządzania Ryzykiem. „Nafta-Gaz” 2001, nr 5, s. 268-281.
- [3] EGIG: 8th Report of the European Gas Pipeline Incident Data Group. December 2011.
- [4] <http://relaxng.org/> [dostęp 10 września 2012 r.]
- [5] <http://www.egig.eu/> [dostęp 10 września 2012 r.]
- [6] <http://www.w3.org/> [dostęp 10 września 2012 r.]
- [7] <http://www.w3.org/standards/xml/> [dostęp 10 września 2012]
- [8] Otegem M.: *XSLT dla każdego. Uniwersalny język przekształcania dokumentów XML*. Wydawnictwo Helion, Gliwice 2003.
- [9] Tadeusiewicz R.: *Mobilny samouczek projektowania systemów informacyjnych*. Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie.
- [10] Trachtenberg A., Skalar D.: *PHP Receptury*. Wydawnictwo Helion, Gliwice 2007.
- [11] Traczyk T.: *XML – stan obecny i trendy rozwojowe*. Politechnika Warszawska, Październik 2003.
- [12] Vieth P. H., Kiefner J. F. et al.: *Final Report on Analysis of DOT Reportable Incidents for Gas Transmission and Gathering Pipelines – January 1, 1985 through December 31, 1994*. Kiefner & Associates, Inc. 1996.



Mgr inż. Jakub BADOWSKI – asystent w Zakładzie Informatyki Instytutu Nafty i Gazu w Krakowie, absolwent Informatyki Stosowanej oraz studiów podyplomowych z zakresu gazownictwa na Akademii Górniczo-Hutniczej w Krakowie. Zajmuje się tworzeniem oprogramowania komputerowego wykorzystującego technologie oparte o Internet. Entuzjasta wolnego oprogramowania.