

## Formaty danych – wektory

Dane w systemach informacji przestrzennej możemy podzielić na dwie grupy:

- Wektorowe,
- Rastrowe.

### Dane wektorowe określają:

- geometrię obiektu,
- jego współrzędne,
- jego atrybuty,
- metadane – np. układ odniesienia.

### Dane rastrowe określają:

- Obraz, jego wartości w pikselach,
- Metadane – np. układ odniesienia.

Dane wektorowe są szeroko wykorzystywane do przechowywania, przesyłania i przetwarzania danych w SIP. Istnieje wiele wykorzystywanych formatów danych wektorowych, z których każdy ma indywidualne cechy i zalety. Spośród najważniejszych wymienić można:

- Geography Markup Language (GML),
- GeoJSON,
- Shapefile (SHP),
- Keyhole Markup Language (KML),
- Well-known text/Well-known binary (WKT/WKB),
- Pliki tekstowe - CSV, TSV, SSV, PSV

Open Geospatial Consortium (OGC) jest organizacją, która opracowuje, implementuje i rozwija otwarte standardy danych i usług przestrzennych dla systemów informacji przestrzennej.

Standardy danych wektorowych zalecane przez OGC to GML oraz KML. Taki też standard został przyjęty przez Krajową Infrastrukturę Informacji Przestrzennej. W jej ramach udostępniane są bazy w różnych skalach i zróżnicowanej tematyce, jak na przykład Baza Danych Ogólnogeograficznych (BDO) czy Baza Danych Obiektów Topograficznych.

Standard danych GML jest oparty na języku XML - eXtensible Markup Language. To uniwersalny język znaczników, służący do definiowania struktur danych. Jest niezależny od platformy i pozwala na sprawniejszą wymianę i weryfikację poprawności strukturalnej danych. Jest standardem rekomendowanym przez organizację W3C - World Wide Web Consortium. XML jest między innymi podstawą HTML, przechowuje dane, nie dba o ich reprezentację.

## XML

XML ma strukturę drzewiastą. Jest to zespół zagnieżdżonych znaczników o ustalonej strukturze. Każdy dokument zaczyna się od nagłówka:

```
<?xml version="1.0" encoding="UTF-8"?>
```

Powyższe określa wersję XML - tutaj jest to 1.0. Możliwy jest wybór wersji 1.1. Są one ze sobą jednak niezgodne w obie strony.

Encoding oznacza sposób kodowania znaków, UTF-8 to jeden ze sposobów kodowania, który wyparł popularniejsze wcześniej ASCII

Następnie należy otworzyć główny znacznik, który obejmie cały dokument. Posiada on nazwę i definiowany jest za pomocą znaków "<>":

```
<nazwa>
```

Każdy otwarty znacznik, musi być zamknięty za pomocą znaku "/"

```
</nazwa>
```

W każdym z nich, mogą znajdować się wartości (typ prosty) lub zagnieżdżone znaczniki (typy złożone).

```
<nazwa>Wartość</nazwa>
```

Lub zagnieżdżając:

```
<nazwa><nazwa2>Wartość</nazwa2></nazwa>
```

Zwróć uwagę na sposób zagnieżdżenia, nie mogą być one przemieszane ze sobą. W celu ułatwienia czytelności, powyższe można zapisać:

```
<nazwa>
```

```
<nazwa2>Wartość</nazwa2>
```

```
</nazwa>
```

Zatem przykładowo, zapis pewnej notatki za pomocą języka XML może wyglądać tak:

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<notatka>
```

```
<do>Rygobert</do>
```

```
<fod>Saturnin</od>
```

```
<naglowek>Wiadomość</naglowek>
```

```
<tekst>Cześć!</tekst>
```

</notatka>

Unikamy polskich znaków w znacznikach!

Kontynuując powyższy przykład, jeśli chcemy przechowywać kilka notatek, należy stworzyć zawsze znacznik **główny** (tzw. root), który występuje tylko raz w całym dokumencie.

<?xml version="1.0" encoding="UTF-8"?>

<notatki>

<notatka>

<do>Rygobert</do>

<fod>Saturnin</od>

<naglowek>Wiadomość</naglowek>

<tekst>Cześć!</tekst>

</notatka>

<notatka>

<do>Rygobert</do>

<fod>Saturnin</od>

<naglowek>Wiadomość</naglowek>

<tekst >Cześć!</tekst>

</notatka>

</notatki>

## Zadania

Wszystkie zadania wysłane mają być do 2 tygodni po zajęciach na eSprawozdania -> Kamil Smolak > Studia stacjonarne -> SIPII

1. Napisz dokument XML, który będzie miał następującą strukturę:
  - a. Typ złożony linia,
    - i. która zawiera typ złożony punkt,
      1. który zawiera typy proste współrzędnych x oraz y
2. Napisz dokument XML, który będzie miał następującą strukturę:
  - a. Będzie przechowywał dane adresowe i osobowe kilku osób (minimum dwóch)

- b. będzie zawierać typy proste: imię, nazwisko, data\_urodzenia
  - c. będzie zawierać jeden typ złożony: miejsce zamieszkania z typami prostymi: typ domu, oraz kolejnym typem złożonym adres, ze typami prostymi: miasto, ulica, numer, lokal
3. Popraw schemat XML w pliku zadanie3, tak żeby był poprawny.