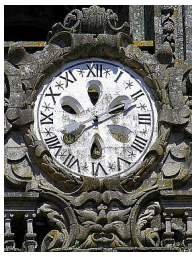

Zegar w katerdrze Santiago de Compostela. Autor: Miguel Bellido Sánchez (Arbuga), *CC BY-SA 3.0 [via Wikimedia Commons]*

Nic tak dobrze nie ilustruje trudności, jakie napotyka tłumaczenie języka ludzkiego na język komputerowy, jak zapis jednostek czasu. Kalendarz którym się posługujemy pochodzi z roku 46 BC (wprowadzony przez Juliusza Cezara) z późniejszymi poprawkami z roku z 1582 (papieża Grzegorza VIII). Jest to najpopularniejszy, ale nie jedyny kalendarz, są inne w szerokim zastosowaniu jak Hijri, Hindu, kalendarz chiński czy hebrajski. W kalendarzu gregoriańskim podajemy dzień, miesiąc i rok, i wydaje nam się, że komputery powinny te ‘proste dane’ zawsze zrozumieć. Niestety, prostota ta jest bardzo pozorna.

Pierwszym problemem jest zapis daty. Po polsku zapisujemy 2 VI 1945 albo 11 września 1950 albo 3.4.23. Po angielsku zapisujemy 3/4/23 co może znaczyć 3 kwietnia 1923 roku albo 4 marca 2023 roku, zależnie od kraju i domyślnego wieku. Do niedawna komputer nie umiał zrobić nawet tak prostej konwersji, dziś próbuje odgadnąć język, którym się posługujemy, i czasem udaje się zgadnąć (a czasem nie) właściwą datę z zapisu.

Drugim problemem jest zapis czasu. W USA doba jest podzielona na godziny ‘ranne’ oznaczane AM i ‘popołudniowe’ - PM. Ale nikt nie chce się umawiać w poniedziałek o 12 AM ani we wtorek o 12 PM, gdyż te zapisy nie są jednoznaczne. Zegar 24 godzinny usuwa przynajmniej tę jednoznaczność (godzina 24:00 w poniedziałek i 0:00 we wtorek to ten sam moment w czasie, ale należący do innego dnia).

Kolejny problem to zapis z okresloną dokładnością. Ktokolwiek pracował z arkuszem rozliczeniowym, nawet tym najnowszym, wie, że można zapisać dzień (bez określenia czasu) albo czas co do minuty. Możemy więc zapisać wydarzenie które ma miejsce danego dnia, albo w danej minucie, ale nie można zapisać roku (bez określenia miesiąca), roku i miesiąca (bez określenia tygodnia), roku, miesiąca i tygodnia (bez określenia dnia), ani też tylko godziny, godziny i minuty, itp. (inaczej niż przez opisanie tego słownie, co całkowicie zbija z tropu

komputer).... i tak dalej.

Wiele programistów borykając się z takimi trudnościami, stwarzało własne standardy zapisu dat. Jest ich wiele, co stwarza kolejny problem kompatybilności. Dopiero wprowadzenie standardu ISO, oraz przyjęcie go przez konsorcjum W3C pozwoliło na jednolity zapis podstawowych danych dotyczący czasu. Poniżej jego najważniejsze elementy, oparte o standard Schema z W3C, który ogranicza nieco różnorodność zapisu z ISO 8601.

Data

- Datę zapisujemy jako YYYY-MM-DD, np. **1867-12-05**
- Możemy opuścić dzień np. **1867-12**, albo dzień i miesiąc, np. **1867**.
- Możemy zamiast miesiąca i dnia podać tydzień w roku: **1945W06**

Pora dnia

- Porę dnia zapisujemy jednolicie jako hh:mm:ss, np. **13:25:17**
- Możemy opuścić sekundy np. **13:25** albo minuty i sekundy, np. **13**
- Możemy dodać ułamek sekundy z dowolna dokładnością np **13:25:17.50255**
- Jeśli podajemy dzień i czas, rozdzielamy je znakiem T, np. **1867-12-05T18:45**
- Możemy też oznaczyć strefę czasową (bez niej czas jest lokalny albo bez zdefiniowanej strefy). Czas np **13:25:17Z** oznacza czas UTC (dawniej Greenwich), a **13:25:17-05:00** lub
2013-06-05-05:00
oznaczają czas lub dzień w strefie czasowej Nowego Jorku.

Powtarzający się okres czasu

Schema W3C definiują powtarzające się okresy czasu: dzień, dzień-miesiąc i miesiąc.

- Dzień powtarza się w każdym miesiącu (i trwa jeden dzień) np. **---15** to piętnasty każdego miesiąca.
- Dzień miesiąc powtarza się tylko raz w roku, np. **--12-24** to dzień 24 grudnia.
- Miesiąc powtarza się raz w każdym roku (i trwa cały miesiąc), np. **--08** to sierpień każdego roku.

Czas trwania

Możemy też zapisać czas trwania jakiegoś wydarzenia. Czas ten może być określony mniej lub bardziej precyzyjnie, ale w każdym przypadku posiadamy narzędzie do jego zapisu. Zaczynamy zapis od **P** (period) a następnie dodajemy trwanie w latach (**Y**), miesiącach (**M**), dniach (**D**) godzinach (
H
), minutach(
M
) i sekundach (
S
). Przykłady:

P12Y5M (12 lat i 5 miesięcy)
P36M (36 miesięcy)
P20DT15H30M (20 dni, 15 godzin i 30 minut (potrzebny symbol T aby oddzielić datę od czasu)
P1356S (1345 sekund)
P99.486S (99.486 sekund) itp,

Używanie zapisu ze standardu ISO 8601 jest wskazane, szczególnie tam, gdzie ważna jest precyzyjność danych (a więc m.in. przy wpisywaniu danych w archiwistyce, wypełnianiu tabel itp.). Jak sie jednak można domyśleć, ten standard to dopiero poczatek. W genealogii od dawna znany jest standard GEDCOM o szerszym zakresie zapisu dat, a archiwistyce stosowane są w postaci opisowej różne inne przypadki. Mamy więc zapis wydarzenia które wystąpiło **między** dwoma punktami w czasie, np. miedzy 1923 a 1925, albo wydarzenia które trwało
od-do
np. od maja 1923 do grudnia 1925. Pierwsze mogło trwać jeden dzień, drugie ok 32 miesięcy. Potrzebujemy zapisać datę, której precyzja jest nieznana, np. około 1923 albo około maja 1950, albo w przybliżeniu między majem a sierpniem 2013. Albo wydarzenie które nastąpiło w drugiej

dekadzie 20 wieku, czy też okres który zaczął się w znanym punkcie w czasie, a skończył w nieznanym (albo wcale nie skończył).

Biblioteka Kongresu USA pracuje nad nowym standardem który ma w zamiarze pozwolić na wyrażenie tego typu danych. Standard EDTF (Extended Date/Time format) jest jeszcze w trakcie budowy, ale wstępna wersja jest juz dostępna w sieci, i można z niej skorzystać (a także dokonać walidacji swoich danych).

Problemów z datami jest oczywiście więcej. Na przykład Rzymianie nie znali zera - patrz liczby rzymskie - więc nie było też roku zerowego. Zero jest znane od 976 roku - wymyślone przez perskiego matematyka Muhammada ibn Ahmad al-Khwarizmi, i po tysiącu lat zaczynamy powoli używać zera przy liczeniu lat. Jak widać nowinki matematyczne zostają wprowadzane do powszechnego użycia dość powoli. Ponieważ jednak komputeryzacja zmienia świat dużo szybciej, używanie opisanej powyżej, standardowej notacji jest bardzo polecane, gdyż zwiększa szanse, że tak cierpliwie zapisywane przez nas dane dotyczące czasu zostaną bezbłednie zinterpretowane przez ciągle jeszcze niezbyt mądre komputery.

_____

1) “Pytanie: Czy czas *południe* i *północ* zapisujemy jako 12 a.m. i 12 p.m. Odpowiedź: To jest podchwytliwe pytanie ponieważ 12 a.m. i 12 p.m. są niejednoznaczne i nie powinny byc używane “ - ze strony *US National Institute of Standards and Technology (NIST)* .

Więcej o zapisie dat

- Standard ISO 8601 - artykuł w Wikipedii
- Wszystko o kalendarzach
- O zegarze 12 godzinnym (AM i PM)
- Standard EDTF

Marek Zieliński

Artykuł ukazał się 1 czerwca 2013 w *Blogu archiwistów i bibliotekarzy Instytutu Piłsudskiego*

Może Cię też zainteresować:

- Wstęp do standardów metadanych
- Digital Humanities
- Digitalizacja a cyfryzacja