

Exakt eller på ett ungefär?

Brist på information om precision leder till felaktig användning



2010-04-14: Sven-Håkan Olsson

PRECISION KRÄVS Är det här en exakt siffra, eller bara ungefärlig? När vi definierar tjänstekontrakt eller hela informationsmodeller glömmar vi ofta bort att beskriva fältens exakthet.

En maskiningenjör beskriver alltid mått uttryckta med precisionsnivå. Det står inte bara 1,5 cm utan det står 15 mm +/- 0,1 mm. En maskindel kan få vara 15,09 mm men tillåts inte vara 15,11 mm. I industrialismens barndom förstod man att fullständigt exakta mått helt enkelt inte existerar i sinnevärlden. Ska man klara av massproduktion, underleverantörer och utbytbara reservdelar så måste man ange med vilken precision en maskindel ska tillverkas för att sedan passa i den färdiga konstruktionen.

Inom datavärlden däremot finns det faktiskt ”exakt” information. Jag bor i Stockholms län, varken mer eller mindre. Jag heter Sven-Håkan Olsson. Oops, nej det gör jag ju inte, i folkbokföringen står det Sven Håkan Olsson! Bindestrecket stoppades in av lågstadieläraren som hade svårt att hålla isär tre olika Håkan i klassen. Sedan har bindestrecket hängt kvar och finns i en mängd register.

Så även om det ibland existerar exakt information så är detta inte alltid fallet. Det här måste beskrivas när informationen definieras. Gör en begreppsmodell eller semantikbeskrivning, se till att denna refereras i maskingränssnittsdefinitioner och XML-scheman – precisionsinformationen ska ingå i informationskontraktet.

Exempel: ofullständiga personnummer

Ett genomgående exempel i Sverige är personnummer. Många befintliga databaser innehåller inte sekulinformation för personnumren. I samband med manuell handläggning är det oftast självklart ifall en person är 101 år eller ett, åldring eller bäbis. Men storskalig databearbetning kan inte klara den bedömningen exakt. I enstaka applikationer görs idag faktiskt hårresande automatiska gissningar.

Varje gång man tänker: ”Nu ska vi göra ett ryck, nu inför vi verkligen 12 siffrors personnummer” så faller det ofta på att man har äldre information lagrad som ska konverteras in i det nya systemet och där sekulinformation saknas. Eller man behöver en integration med något gammalt system. Det här blir som en slags smitta, finns det bara något system utan sekeldata i applikationsfloran i organisationen så smittar det ner de andra med sin lägre precision.

Nyss funderade jag på en meddelandedefinition där flera system ska kunna skicka information till ett mottagande system. Personnummer ingår i informationsbehovet. Det mottagande systemet nykonstrueras och tänks kunna lagra sekulinformation men flera av de sändande har inga sekelsiffror.

Som jag ser det finns det bara två lösningar på detta problem. Det ena är att införa sekelfo i alla sändande system och göra en helt manuell genomgång av alla småbarn och äldrepersoner. I vårt fall är denna lösning fullständigt ogörlig av kostnads- och tidsskäl.

Den andra lösningen är att sända personnummren SAMT info om deras precision!

Det blir därmed inte sämre än idag, utan ofta bättre, men varje gång en manuell handläggning görs i mottagande system så har användaren en chans att veta om personnumret är exakt eller om en åldersbedömning måste göras. Användargränssnittet ska förstås avspegla precisionen på ett pedagogiskt sätt så att användaren har en chans att förstå aktuell exakthet – ifall vi vet sekel eller inte – och om en ålderskoll bör utföras.

Ungefär så här är mitt förslag i detta fall. Två separata fält skickas för varje nummer:

- SamordnPersNummer
 - Samordningsnummer/personnummer för en person. Format ÅÅMMDD-NNNN.
- SamordnSekel
 - I de fall som sekelsiffror för personen är tydligt utredd och känd i sändande system så fylls sekelfo i i detta fält. Format alltid ÅÅ, till exempel ”19”. Tanken är att slippa suddig gissningslogik och hellre överföra det som sändande system faktiskt vet. Annars utelämnas fältet, vilket blir en indikation på lägre precision.

Helt nya personnummer?

Sedan är det en helt annan sak att personnumren enligt vissa regelförslag ska göras om i framtiden så att inte födelsedatum går att vaska fram ur numren som idag. Att bygga om alla Sveriges applikationer ifall det blir så kommer nog att ge högre precision på köpet MEN kostnaden för en sådan regeländring kommer att bli mycket hög. Kanske inte lika hög som att åtgärda år-2000-problemet, men förmodligen i miljardklassen för hela Sverige!

Minns var ni läste det först (om jag nu är först) – och köp aktier i IT-konsultbolagen eller så...



Sven-Håkan Olsson är en fristående konsult som särskilt arbetar med att kombinera verksamhetsnytta med teknikhöjd. Han har en lång karriär sedan 70-talet som it-konsult (it-arkitektur, systemdesign, programmering, reviewer, utredningar, kursledning). Sven-Håkan är också medgrundare av Know IT och var dess teknikchef 1990-2003. Han utsågs till en av "Sveriges topputvecklare" av Computer Sweden 2008.

Sven-Håkan håller regelbundet kurser åt Dataföreningen Kompetens, till exempel "Cloud Computing integration och migration". Läs gärna mer på hans blogg www.definitivus.se.