

LA COMMON STATISTICAL PRODUCTION ARCHITECTURE PER LA CONDIVISIONE E IL RIUSO

di Alessio Cardacino (alcardac@istat.it)

■ Negli ultimi anni, nonostante le restrizioni di bilancio per le organizzazioni che producono statistiche ufficiali, si è registrata una sempre maggiore richiesta di dati ad un livello di dettaglio sempre più disaggregato, da fornire in maniera tempestiva e senza perdere in qualità. Tali ragioni hanno spinto molte organizzazioni ad attivare una industrializzazione e standardizzazione dei processi di produzione statistica, basate su soluzioni IT efficienti, affidabili, economiche e riutilizzabili. All'inizio di quest'anno l'High-Level Group for the Modernisation of Statistical Production and Services (HLG), gruppo di lavoro di alto profilo a cui partecipano presidenti di istituti nazionali di statistica e Chief Statisticians delle organizzazioni internazionali, ha dato vita a un progetto per la creazione di una Common Statistical Production Architecture (CSPA, o "Architettura Plug and Play"), ovvero di un'architettura di tipo industriale per la produzione statistica orientata alla condivisione e al riuso di processi, dati e componenti software.

LO SPRINT DI OTTAWA

Il progetto è stato organizzato in attività intensive da effettuare in un arco limitato di tempo (sprint). Il primo sprint si è tenuto dall'8 al 12 aprile 2013 a Ottawa presso l'Istituto Nazionale di Statistica del Canada, ed ha avuto come oggetto la definizione della Common Statistical Production Architecture in termini di obiettivi, principi, requisiti e modelli architetturali di riferimento. Allo sprint hanno partecipato esperti di architetture d'impresa e di servizi IT per la produzione statistica

appartenenti alle seguenti organizzazioni: Statistics Canada, ABS (Australia), Statistics New Zealand, ONS (Regno Unito), Istat, Statistics Netherlands (CBS), Statistics Sweden (SCB), INEGI (Messico). L'attività è stata basata su riunioni di brainstorming. L'architettura è stata strutturata in livelli (architettura di business, informativa, applicativa e tecnologica) definendo per ciascuno di essi un'adeguata descrizione. L'idea alla base dell'architettura consiste nel supportare i processi di produzione statistica definiti in maniera conforme allo standard GSBPM (Generic Statistical Business Process Model), combinando un insieme di componenti generalizzati con la medesima facilità con cui si collega una periferica hardware di tipo plug and play (ad esempio una memoria USB) ad un PC.

L'informazione statistica scambiata dai componenti (in termini di dati e metadati) deve essere strutturata in modo standard (possibilmente utilizzando protocolli internazionali come SDMX o DDI) e deve basarsi su un information model compatibile con lo standard GSIM (Generic Statistical Information Model).

I processi, i componenti e i metadati devono essere resi disponibili all'interno di cataloghi localizzati nelle singole organizzazioni, oppure condivisi da più organizzazioni con l'obiettivo di costituire un catalogo globale di artefatti da rendere disponibile a tutte le istituzioni statistiche del mondo.

Da un punto di vista applicativo l'architettura si configura come una Service Oriented Architecture (SOA), in cui i componenti sono implementati come servizi dotati

di interfacce standard. Tali servizi possono essere sviluppati ex-novo, oppure costruiti tramite un incapsulamento (wrapping) di componenti software preesistenti (tool, API, ecc.).

RISULTATI OTTENUTI E PROSSIMI PASSI

Al termine dei cinque giorni di attività, sulla base dei materiali raccolti durante lo sprint, è stata redatta una prima versione del documento sulla Common Statistical Production Architecture. Come prossimo passo è prevista l'implementazione, entro la fine del 2013, di un *proof of concept* dell'architettura che ne dimostri la fattibilità di realizzazione e di utilizzo da parte degli istituti statistici. Ed è al disegno di tale *proof of concept* che è stato dedicato il secondo sprint, svolto a Roma dal 10 al 14 giugno 2013.