

Audizione di docenti degli atenei di Roma esperti di ICT, con la finalità di acquisirne il punto di vista sul tema dell'indagine conoscitiva "Digitalizzazione e interoperabilità delle banche dati fiscali", anche con riferimento al progetto di migrazione delle banche dati della PA verso il Cloud nazionale.

Roma, 10 novembre 2021

Memoria del Prof. Riccardo Torlone (Università Roma Tre)

Gentile Presidente, onorevoli membri della Commissione,

vi ringrazio innanzitutto per l'invito a questa audizione. Sono professore ordinario presso l'Università Roma Tre dove dirigo un laboratorio di ricerca sulle basi di dati e sui Big Data situato nel Dipartimento di Ingegneria, del quale sono anche Direttore vicario. Mi occupo da molti anni di vari problemi relativi al trattamento informatico dei dati in tutte le fasi del loro ciclo di vita: dalla loro raccolta, alla trasformazione e integrazione, fino all'accesso, all'analisi e alla estrazione di conoscenza dei dati raccolti. In accordo con i miei colleghi qui presenti, farò un breve intervento riportando alcune mie considerazioni sul tema dell'indagine e sulle sue correlazioni con il progetto del Polo Strategico Nazionale¹ (PSN), relativo alla migrazione delle banche dati della Pubblica Amministrazione (PA) verso un servizio di Cloud nazionale, concentrandomi sul tema di mia maggiore competenza, ovvero sull'interoperabilità dei dati in questo contesto e sul problema connesso della loro fruibilità. I miei colleghi approfondiranno invece i temi del Cloud computing e della sicurezza dei dati in questo ambito. Premetto che potrò fare solo alcune considerazioni generali su questo argomento, non conoscendo nel dettaglio il dominio applicativo di interesse, ovvero la quantità, la qualità, la dislocazione geografica e soprattutto il livello di eterogeneità delle banche dati fiscali attualmente esistenti, sebbene sia al corrente che abbiamo davanti un panorama molto articolato con la presenza di più di 150 banche dati. Cercherò inoltre, da tecnico, di adottare un linguaggio scevro da tecnicismi troppo spinti, per rendere fruibile le mie considerazioni a un'ampia platea.

Vorrei innanzitutto chiarire che, tecnicamente parlando, la prospettata migrazione delle banche dati e dei servizi della Pubblica Amministrazione nel Cloud nazionale entro il 2025 non risolverà il problema della loro interoperabilità. Il Cloud, infatti, rappresenta solo una tecnologia che consente di offrire un'infrastruttura ovvero, per dirla in termini semplici, una "casa comune" sicura e protetta a dati e servizi della PA, garantendo sovranità nazionale e autonomia tecnologica. Viceversa, l'interoperabilità delle basi di dati è un problema di carattere metodologico, che investe cioè un processo teso a garantire l'interscambio e la condivisione di dati tipicamente eterogenei e memorizzati su sistemi diversi tra i quali una possibile scelta, ma certamente non l'unica, è il cloud. Si tratta in effetti di un problema con il quale, tipicamente, imprese e istituzioni hanno sempre avuto a che fare, da ben prima dell'avvento di risorse di cloud computing. Si pensi che il problema dell'integrazione di basi di dati è probabilmente uno dei più studiati sin dagli albori dell'avvento dell'informatica nella automatizzazione dei processi, sia in campo scientifico sia nel mondo produttivo, e per il quale una soluzione di carattere generale non è stata ancora trovata. In questo contesto, la disponibilità di una risorsa di cloud pubblico nel quale le varie banche dati della PA potranno alloggiare nel prossimo futuro rappresenta certamente un mezzo facilitatore, in quanto offre una infrastruttura comune ai dati ma, appunto, non dirimente alla soluzione del problema. Adottando un linguaggio comune, se due basi di dati "non si parlano" su infrastrutture tradizionali continueranno a non parlarsi sul cloud. A mio parere, quindi, la vera sfida della strategia Cloud per la Pubblica Amministrazione è rappresentata proprio dall'interoperabilità, ovvero dalla capacità di fare finalmente dialogare i dati della PA una volta che abiteranno in una casa comune.

Con riferimento al tema dell'adozione di politiche tese a garantire la piena interoperabilità delle basi di dati della PA, vedo con pieno favore le iniziative attualmente in corso relative allo sviluppo di un *modello di*

¹ <https://innovazione.gov.it/dipartimento/focus/polo-strategico-nazionale/>

*interoperabilità unificato*² e di una *Piattaforma Digitale Nazionale Dati*³ (PDND) nelle quali sono coinvolti prioritariamente l'AgID e PagoPA.

Il modello di interoperabilità serve a definire, in termini tecnici precisi, un meccanismo di cooperazione standardizzato tra pubbliche amministrazioni e tra queste e soggetti terzi, per mezzo di soluzioni tecnologiche condivise che assicurano l'interazione e lo scambio di dati senza vincoli sulle implementazioni. Alla base c'è un principio secondo il quale ogni servizio software (per esempio il reperimento dei dati da una specifica banca dati) viene condiviso attraverso un'interfaccia digitale pubblica, chiamata Application Program Interface (API), che è basata su standard condivisi e che garantisce un accesso controllato al servizio in maniera indipendente dalla sua implementazione. Questa soluzione ha una serie di vantaggi: (1) favorisce lo sviluppo di nuove applicazioni a servizio dei cittadini e delle stesse PA costruita combinando vari servizi software; (2) assicura, nel rispetto del diritto alla privacy, l'accesso ai dati della Pubblica amministrazione anche a soggetti terzi; e (3) è coerente con i principi definiti in ambito europeo (in particolare con l'European Interoperability Framework⁴ della Commissione Europea). Viceversa, la Piattaforma Digitale Nazionale Dati per l'interoperabilità dei sistemi informativi e delle basi di dati (per la quale si è appena conclusa la consultazione pubblica di un primo documento descrittivo)⁵ è invece uno spazio virtuale atto ad agevolare le amministrazioni nell'adozione del succitato modello di interoperabilità e nel quale ciascun ente potrà rendere disponibili le proprie informazioni, secondo regole condivise basate sul meccanismo delle API. Più specificatamente, la PDND è costituita da: (1) un catalogo delle API prodotte dalle PA che la piattaforma rende disponibile ai soggetti accreditati, (2) da meccanismi per l'accreditamento, l'identificazione e la gestione dei livelli di autorizzazione dei soggetti accreditati e (3) da strumenti per la raccolta e la conservazione delle informazioni relative agli accessi e alle transazioni effettuati suo tramite. Va osservato che, in base al decreto legislativo del 2005 recante il "Codice dell'amministrazione digitale", in fase di prima applicazione la Piattaforma dovrebbe assicurare prioritariamente, con questo sistema, l'interoperabilità con le cosiddette "basi di dati di interesse nazionale" (che includono, per esempio, l'anagrafe nazionale della popolazione residente e il registro delle imprese) e con le banche dati dell'Agenzie delle Entrate individuate dal Direttore della Agenzia stessa. Ne consegue che gli enti che gestiscono le banche dati fiscali dovranno attrezzarsi in breve tempo per essere in grado di generare le API per l'accesso ai dati secondo le linee guida definite dall'AgID per l'adesione al modello di interoperabilità ed accreditarsi presso la PDND.

In questo quadro, ci sono alcune importanti considerazioni da fare.

La prima considerazione è che i progetti relativi al Cloud nazionale e alla Piattaforma Digitale Nazionale Dati sono pienamente coerenti. Infatti, il meccanismo di accesso alle informazioni basato sulle API adottato dal PDND è di tipo "black-box", ovvero ne nasconde i dettagli implementativi, inclusa la specifica tecnologia utilizzata per la memorizzazione dei dati. Ne consegue che è tecnicamente possibile sia modificare l'applicazione software che opera sui dati sia migrare i dati da una piattaforma ad un'altra senza che l'API venga modificata (purché ovviamente sia garantita la consistenza del suo comportamento esterno) e quindi senza che i suoi utilizzatori ne abbiano contezza. La migrazione nella PDND avrà quindi un impatto relativamente basso sul processo di adozione del modello di interoperabilità.

La seconda considerazione è che il modello di interoperabilità unificato e la PDND non costituiscono ancora la soluzione definitiva al problema dell'interscambio di informazioni tra le banche dati della PA. Definiscono infatti uno strumento tecnologico fondamentale basato su standard condivisi ma sarà necessario sviluppare nuove applicazioni che, utilizzando questo strumento, siano in grado di comporre le funzionalità offerte dalle API disponibili nella PDND al fine di offrire una soluzione in grado di combinare dati memorizzati in basi di dati di Amministrazioni diverse o in basi di dati diverse della stessa Amministrazione. Per ottenere questo, è indispensabile un altro importante ingrediente: far uso, nella produzione delle API, di un vocabolario comune

² <https://www.agid.gov.it/it/infrastrutture/sistema-pubblico-connettivita/il-nuovo-modello-interoperabilita>

³ <https://www.pagopa.it/it/prodotti-e-servizi/piattaforma-digitale-nazionale-dati>

⁴ https://ec.europa.eu/isa2/eif_en

⁵ <https://docs.italia.it/AgID/documenti-in-consultazione/lg-pdnd-docs/it/bozza/>

o, per dirla in termini più tecnici, di una ontologia condivisa o, quantomeno, di una collezione di ontologie specializzate per i vari domini applicativi. In termini semplici, un'ontologia serve a capire che due termini diversi utilizzati in due banche dati rappresentano in realtà lo stesso concetto e a descrivere l'esistenza di una correlazione semantica tra termini che rappresentano concetti diversi. Questo aspetto è citato nelle Linee Guida che definiscono il modello di interoperabilità, nelle quali si invitano le amministrazioni ad aderire, nella produzione delle API, a una collezione di ontologie di dominio che l'AgID sta raccogliendo, ma è di fondamentale importanza per il raggiungimento di una piena interoperabilità tra le banche dati della PA. A tale riguardo, segnalo che la prima linea di indirizzo strategico del PSN, propedeutica alla realizzazione del Polo, consiste in una classificazione completa dei dati e dei servizi in funzione del danno che una loro compromissione in termini di confidenzialità, integrità e disponibilità, provocherebbe al sistema Paese. Si potrebbe sfruttare questa occasione per una classificazione più completa che integri queste informazioni con una descrizione più esaustiva, anche in termini ontologici, dei dati e dei servizi della PA in maniera da favorire il successivo processo di adesione al modello di interoperabilità dell'AgID e al PDND.

L'ultima considerazione che vorrei fare è legata al problema della fruizione dei dati da parte dei cittadini in questo contesto. Deve essere chiaro che il meccanismo delle API a supporto dell'interoperabilità non è pensato per essere utilizzato direttamente dall'utente finale (il cittadino per intenderci) ma favorisce semplicemente il dialogo tra componenti software e quindi nasce per offrire un supporto ai programmatori nello sviluppo di applicazioni. Una volta quindi che queste operazioni andranno in porto sarà necessario un ulteriore sforzo da parte delle istituzioni e, perché no, da soggetti terzi atto a garantire lo sviluppo di una serie di applicazioni al servizio del cittadino che, facendo leva sul modello e sull'infrastruttura tecnologica che è stata realizzata, sia in grado di semplificare l'accesso ai servizi della Pubblica Amministrazione attraverso un meccanismo con accesso unificato e di facile uso.

Chiudo il mio intervento osservando che un contributo importante a tutte le attività di cui ho parlato può essere fornito dai centri di ricerca pubblici e in particolare dalle Università, lungo due direzioni. La prima è quella della formazione: uno degli ostacoli maggiori al processo che ho delineato è la scarsità di competenze tecniche ed in questo le Università giocano un ruolo fondamentale, essendo loro compito istituzionale quello di formare professionisti capaci di affrontare le sfide che abbiamo davanti. L'altra, riguarda il contributo che docenti e ricercatori delle Università possono offrire in termini dei metodi e degli strumenti innovativi che sono il risultato delle loro attività di ricerca, attraverso quel tanto auspicato trasferimento di conoscenze che le Università possono fornire alle istituzioni e alle imprese.

Vi ringrazio per l'attenzione.

Riccardo Torlone

Università Roma Tre
Dipartimento di Ingegneria
Via della Vasca Navale, 79
00146 Roma
Web: <http://torlone.dia.uniroma3.it/>
Email: riccardo.torlone@uniroma3.it
Tel. 3290552319