

Curriculum Vitae

Informazioni personali

Nome / Cognome	Monica Scannapieco		
Indirizzo	OMISSIS		
Telefono	OMISSIS	Cellulare:	OMISSIS
E-mail	scannapi@istat.it		
Cittadinanza	Italiana		
Data di nascita	OMISSIS		
Sesso	femminile		

Istruzione e formazione

Studi

Date	23 Febbraio 2000
Titolo della qualifica rilasciata	Laurea in Ingegneria Informatica
Principali tematiche/competenze professionali possedute	La tesi di laurea si colloca nell'ambito dei sistemi informativi ed ha per titolo: <i>“Progettazione a Componenti Distribuiti. Proposta, Sperimentazione e Confronto di Paradigmi di Sviluppo in ambito Rete Unitaria della Pubblica Amministrazione (RUPA)”</i> . Ne è stato relatore il Prof. Carlo Batini. La tesi ha avuto carattere sperimentale, con oggetto la progettazione ed implementazione di soluzioni software basate su <i>middleware</i> a componenti distribuiti nel contesto di una pubblica amministrazione italiana (INPS). Votazione riportata: 110/110 e lode
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	Dipartimento di Informatica e Sistemistica, Università degli Studi di Roma La Sapienza
Date	1993
Titolo della qualifica rilasciata	Diploma di Maturità Scientifica
Principali tematiche/competenze professionali possedute	Diploma di maturità scientifica conseguito con votazione di 60/60

Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione Liceo Scientifico Statale "B. Mangino", Pagani (SA)

Istruzione e formazione	Studi Post-Laurea
Date	11 Febbraio 2004
Titolo della qualifica rilasciata	Dottorato di Ricerca in Ingegneria Informatica
Principali tematiche/competenze professionali possedute	L'attività di ricerca condotta nell'ambito del dottorato si colloca principalmente nell'area delle basi di dati, ed è, in particolare, incentrata sulla qualità dei dati. Il titolo della tesi di dottorato è: " <i>DaQuinCIS: Exchanging and Improving Data Quality in Cooperative Information Systems</i> ". Relatore della tesi: Prof. Carlo Batini. Il contributo fornito riguarda (i) modelli per la rappresentazione e l'interrogazione della qualità dei dati e (ii) tecniche per il miglioramento della qualità dei dati in sistemi informativi cooperativi. I risultati conseguiti sono tra i primi relativi alla gestione della qualità dei dati in sistemi informativi distribuiti. Ulteriori tematiche oggetto di ricerca nell'ambito del dottorato includono: modelli e linguaggi di interrogazione basati su XML, composizione di metriche di qualità dei dati, qualità dei dati nei <i>Web Information Systems</i> , interoperabilità nei sistemi informativi in rete e <i>trade-off</i> tra privacy e qualità dei dati.
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	Dipartimento di Informatica e Sistemistica, Università degli Studi di Roma La Sapienza

Istruzione e formazione	Post-Doctoral Fellowship
Date	2004 - 2006
Titolo della qualifica rilasciata	Assegnista di Ricerca
Principali tematiche/competenze professionali possedute	La ricerca in oggetto ha per titolo "Qualità dei Dati nei Sistemi Informativi Cooperativi", responsabile scientifico prof. Roberto Baldoni. Si è proseguita l'attività di ricerca intrapresa durante il dottorato, che ha riguardato la qualità dei dati nei sistemi informativi cooperativi.
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	Dipartimento di Informatica e Sistemistica, Università degli Studi di Roma La Sapienza
Date	Agosto 2005 - Dicembre 2005
Titolo della qualifica rilasciata	Temporary Research Assistant
Principali tematiche/competenze professionali possedute	La collaborazione ha avuto per oggetto lo studio di un framework per la valutazione del rischio di violazione della privacy e si è svolta presso l'università di Purdue.
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	Department of Computer Science e il CERIAS (Center for Education and Research in Information Assurance and Security) della University of Purdue (West Lafayette, Indiana, USA)

Istruzione e formazione	Abilitazione Professionale
--------------------------------	-----------------------------------

Date	Febbraio 2001
Titolo della qualifica rilasciata	Abilitazione alla Professione di Ingegnere
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	Università degli Studi di Roma La Sapienza

Istruzione e formazione	Corsi di Formazione
--------------------------------	----------------------------

Date	23-24-30-31/1/2012 e 1/2/2012
Nome Corso	IT Governance e Project Management
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	SAPIENZA – Università di Roma
Date	27/06/2011
Nome Corso	Evolution and merging of real-life ontologies
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	Tutorial al 19th Italian Symposium on Advanced Database Systems
Date	28/06/2011
Nome Corso	Querying and Reasoning about Massive Social Networks
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	Tutorial al 19th Italian Symposium on Advanced Database Systems
Date	Dal 4/6/2008 al 6/6/2008
Nome Corso	Statistical Machine learning
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	Università di Roma “La Sapienza”
Date	23/6/2008
Nome Corso	Secure Logistic Regression: shared computation across databases
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	Istat, Roma
Date	25/9/2007
Nome Corso	Adaptive Query Processing
Nome e tipo d'organizzazione erogatrice dell'istruzione e formazione	Tutorial alla 33rd VLDB Conference, Vienna, Austria
Date	Dal 22/06/2007 al 24/06/2007
Nome Corso	Algoritmi genetici e altre tecniche di calcolo evolutivo

Nome e tipo Corso di dottorato, Politecnico di Milano
 d'organizzazione erogatrice
 dell'istruzione e formazione
 Date 12/06/2007

Nome Corso Provenance in Databases
 Nome e tipo Tutorial alla 27th SIGMOD Conference, Pechino, Cina
 d'organizzazione erogatrice
 dell'istruzione e formazione
 Date 28/6/2006

Nome Corso Record Linkage: Similarity Measures and Algorithms
 Nome e tipo Tutorial alla 26th SIGMOD Conference Chicago, USA
 d'organizzazione erogatrice
 dell'istruzione e formazione
 Date Dal 3/5/2006 al 5/5/2006

Nome Corso Fondamenti di Linux
 Nome e tipo Istat, Roma
 d'organizzazione erogatrice
 dell'istruzione e formazione
 Date 17/06/2004

Nome Corso Tools for Design of Composite Web Services
 Nome e tipo Tutorial alla 24th SIGMOD Conference Parigi, Francia
 d'organizzazione erogatrice
 dell'istruzione e formazione
 Date 20/08/2002

Nome Corso Automation in Information Extraction and Data integration
 Nome e tipo Tutorial alla 28th VLDB Conference, Hong Kong, Cina
 d'organizzazione erogatrice
 dell'istruzione e formazione
 Date Dal 14/05/2001 al 18/05/2001

Nome Corso Scuola di Dottorato ERSADS – European Research Seminar on Advances
 In Distributed Systems
 Nome e tipo Bertinoro, Forlì, Italia
 d'organizzazione erogatrice
 dell'istruzione e formazione
 Date Dicembre 2000

Nome Corso Corso di Dottorato su “Data integration”
 Nome e tipo Università di Roma “La Sapienza”
 d'organizzazione erogatrice
 dell'istruzione e formazione

Esperienza professionale

Date	Dal 1 Marzo 2006 ad oggi
Lavoro o posizione ricoperti	Dal 01/03/2006 al 14/12/2016 Tecnologo Terzo Livello Professionale
Nome e indirizzo del datore di lavoro	Dal 15/12/2016 ad oggi Primo Tecnologo ¹ Istat, Via Cesare Balbo 16, 00184 Roma
Tipo di attività o settore	Ente Pubblico di Ricerca

Principali attività e responsabilità

Appartenenza a strutture Istat e incarichi di responsabilità di strutture Istat

Nel seguito si riportano le Strutture Istat (indicate nell'elenco seguente come SI) di afferenza e i relativi incarichi di responsabilità.

[SI8] Da Febbraio 2017 ad oggi, **Responsabile del Servizio “Architettura Integrata dei Dati e dei Processi” (MEC)** della Direzione Centrale per la Metodologia e il Disegno dei Processi Statistici (DCME), numero di **persone** attuali **25 (venticinque)**. (Rif. Delibera D16 15 DIRM 17 e successive D16 251 DIRM 2019, D16 714 DIRM 2019).

“Il Servizio provvede alla definizione e al monitoraggio dell’architettura di impresa dell’Istat, alla progettazione e alla cura dell’architettura informativa integrata dell’Istituto, comprensiva della relativa modellazione dei metadati strutturali. Si occupa altresì della predisposizione e dell’aggiornamento delle Linee guida per la realizzazione e gestione del sistema dei registri dell’Istituto, della progettazione e della cura dell’architettura di calcolo a supporto dei servizi statistici dei processi di produzione.” (Declaratoria Uffici Tecnici Non Generali)

[SI7] Da Aprile 2016 ad oggi, la struttura di afferenza è il Servizio “Architettura Integrata dei Dati e dei Processi” (MEC) della Direzione Centrale per la Metodologia e il Disegno dei Processi Statistici (DCME) (Rif. Delibera D28-6DIRM16)

[SI6] Da Marzo 2015 ad Aprile 2016, è stata assegnata al servizio DCIT/RST (Direzione centrale per le tecnologie dell'informazione e della comunicazione – Servizio Regolazione e sviluppo delle tecnologie ICT) ed ha rivestito il ruolo di **Dirigente Unità Operativa RST/G (Gestione di Big Data, Open Data e dati multidimensionali)** composta da **10 (dieci) persone**. (Rif. Delibera D16-6DIQR15)

[SI5] Da Novembre 2011 a Marzo 2015 è stata assegnata al servizio DCIT/RST (Direzione centrale per le tecnologie dell'informazione e della comunicazione – Servizio Regolazione e sviluppo delle tecnologie ICT) ed ha rivestito il ruolo di **Dirigente Unità Operativa RST/B (Sviluppo Sistemi informatici per la statistica)** composta da **9 (nove) persone**. (Rif. Delibera D16-24DIQR11)

¹ Vincitrice del concorso PT/2011/E, prima in graduatoria

[SI4] Da Giugno 2009 a Novembre 2011 è stata assegnata al servizio DCMT/SVS (Direzione Centrale per le tecnologie ed il supporto metodologico - Servizio Sviluppo Software) ed ha rivestito il ruolo di **Dirigente Unità Operativa SVS/C (Sviluppo Software per i processi di produzione statistica)** composta da **10 (dieci) persone**. (Rif. DELIBERA/PER 85/DGEN 15-09-2011 e DELIBERA/PER 107 16-06-2009, ODS 9009)

[SI3] Dal Novembre 2007 a Giugno 2009 ha avuto un **incarico per il coordinamento gestionale di 3 (tre) persone** (protocollo 2130 del 30-11-2007) della struttura di progetto DCMT/1 (Direzione Centrale per le tecnologie ed il supporto metodologico - Progetto Internalizzazione dello Sviluppo Software).

[SI2] Da Giugno 2006 a Giugno 2009 la struttura di afferenza è stata DCMT/1 (Direzione Centrale per le tecnologie ed il supporto metodologico - Progetto Internalizzazione dello Sviluppo Software).

[SI1] Da Marzo 2006 a Giugno 2006 la struttura di afferenza è stata DCMT/U (Direzione Centrale per le tecnologie ed il supporto metodologico - Ufficio del Direttore).

Principali attività e responsabilità

Gruppi di lavoro, commissioni e incarichi

Nell'ambito delle attività svolte in Istat, ho partecipato a Gruppi di Lavoro e Task Force, a carattere principalmente operativo, e a Commissioni e Comitati, a carattere prevalentemente strategico. Nel seguito ho distinto tra:

- **Gruppi di Lavoro e Task Force Istat (GDL),**
- **Commissioni e Comitati Istat (CI),**
- **Gruppi di Lavoro in ambito Europeo e Internazionale (GLEI),**
- **Commissioni e Gruppi Inter-istituzionali (CGI) e**
- **Incarichi (INC).**

Gruppi di Lavoro Istat

[GDL 31] **Comitato di indirizzo del Gruppo di Lavoro Interdipartimentale per la creazione del Register-based Analytics Framework dell'Istat (RAF) (D08 942 PRES 2019).**

Ruolo: Membro

[GDL 30] **Gruppo di lavoro interdipartimentale relativo alla ESSnet Big Data Pilots II (DOP626-2019 del 31-5-2019 e successive integrazioni D08 683 DGEN 2019 e D08 1096 DGEN 2019).**

Ruolo: Coordinatore (Risorse impegnate 46).

[GDL 29] **Gruppo di lavoro interdirezionale relativo alla ESSnet Shared Services (DOP/381/2019 del 27-03-2019).**

Ruolo: Esperto

[GDL 28] **Task Force con il compito di progettare il programma della XIII Conferenza nazionale di statistica (DOP/101/2018).**

Ruolo: Membro. Ha curato i contenuti scientifici di tutte le sessioni del

percorso “Integrazione” per un totale di 15 interventi invitati.

[GDL 27] **Task Force** per la identificazione delle aree, eventi e/o fasi del processo a rischio corruttivo e per la programmazione delle opportune misure di **prevenzione della corruzione** (428/PER del 13/06/2017).

Ruolo: Responsabile

[GDL 26] **Gruppo di lavoro interdipartimentale sui Big Data** (costituito con Delibera DGEN D08-19 del 9 febbraio 2016).

Ruolo: Responsabile Linea di Attività su Memorizzazione e analisi di fonti Big: Web, Mobile, Social Media

[GDL 25] **Task Force** con il compito di supportare la transizione delle indagini verso la **raccolta dati** centralizzata (113/DGEN del 17/11/2015).

Ruolo: Membro

[GDL 24] **Analisi e reingegnerizzazione** del processo di compilazione delle statistiche di **finanza pubblica** all'interno del sistema dei Conti Nazionali (costituito con Delibera DGEN D08-93 in data 9 Ottobre 2015).

Ruolo: Referente Linea di Attività di modellazione dei processi, dei dati e dei servizi.

[GDL 23] Progettazione e realizzazione del **sistema informativo ANVIS (ANagrafe Virtuale Statistica)** con gruppo di lavoro definito in del. 45/DGEN 16/4/2015.

Ruolo: Coordinatore (Risorse impegnate: 38).

[GDL 22] **Gruppo di corso** per la realizzazione dell'iniziativa formativa su "**Data Scientist**" (PRES Protocollo 387 del 18 Aprile 2014).

Ruolo: Membro

[GDL 21] **ProbLink (PrivacyOpen Big Data Linkage)**: Il gruppo di lavoro (del. 39/DIQR 15/9/2014) ha i compiti di: (i) definire metodologie e strumenti per misurare la qualità statistica dei processi di record linkage probabilistici realizzati con il tool RELAIS (REcord Linkage At IStat); (ii) investigare soluzioni di record linkage in ambito privacy-preserving, Open data e Big Data.

Ruolo: Coordinatore

[GDL 20] Progettazione del **nuovo sistema informativo del servizio commercio estero** e attività internazionali delle imprese (del. 96/DGEN del 29/10/2014).

Ruolo: Coordinatore (Risorse Impegnate 23).

[GDL 19] Progettazione e realizzazione di un **sistema di controllo e correzione dei dati** acquisiti dalle **indagini campionarie sulle famiglie** (del. 159/DGEN 17/10/2013 e del.107/DGEN del 16/05/2015).

Ruolo: Vice-Coordinamento tecnico e organizzativo delle risorse impegnate nella progettazione e nello sviluppo del sistema - Team tecnico (Risorse impegnate: 5).

[GDL 18] Progetto definito da un **Gruppo di Lavoro** (del. 34/DGEN 27/2/2013) **per la diffusione dei dati censuari**. Workpackage 4: Progettazione e Realizzazione del Data Warehouse Primario – Diffusione Italiana ed Europea.

Ruolo: Responsabile del Workpackage 4. Coordinamento tecnico e organizzativo delle risorse impegnate nel WP4 (Risorse impegnate: 4).

[GDL 17] Gruppo di Lavoro interno al DIQR denominato **“Gruppo di Lavoro sulle attività dello Statistical Network”**. Riferimento: D08-56 DIQR del 25 Novembre 2013.

Ruolo: Componente del Team tecnico-scientifico su Common and Unified Linkage Tools.

[GDL 16] **Gruppo di Lavoro Interdipartimentale su Big Data**, incaricato di individuare i temi di innovazione tecnologica e metodologica relativi all’uso dei Big Data per la Statistica Ufficiale. Riferimento: D08 - 18 DIQR del 26 Marzo 2013.

Ruolo: Membro

[GDL 15] **Gruppo di lavoro** con il compito di coordinare le iniziative atte allo sviluppo, uso ed implementazione dello **standard SDMX** (Del. 26/DIQR/23-07-2012).

Ruolo: Membro

[GDL 14] **Gruppo di Lavoro Interdipartimentale sugli Open Data**, con il compito di supportare le strategie Istat nell’ambito degli Open Data. Riferimento: 15 DIQR dell’11 Aprile 2012.

Ruolo: Responsabile per la Direzione Informatica (DCIT).

[GDL 13] **Gruppo di lavoro** che gestisce la raccolta e la diffusione dei **software generalizzati**, le relative informazioni e gli aspetti informatici connessi alla applicazione web alla base dell'OTS. Riferimento: 100/DPTS – 13/05/2011, Durata: maggio 2011 – Novembre 2012

Ruolo: Membro

[GDL 12] **Progetto per lo sviluppo e la gestione dei dati censuari – DIDAC**, riferimento: D08-5-DIQR-28/10/2011, Durata: Novembre 2011- Dicembre 2014

Ruolo: Vice responsabile (Progetto DIDAC) (Risorse impegnate: 60)

[GDL 11] **Gruppo di Lavoro per lo “Studio ed applicazione di tecniche di Record Linkage per la costruzione di archivi longitudinali sulla popolazione straniera”**, riferimento: D08-54-DPTS 11, Durata: marzo 2011- giugno 2012

Ruolo: Membro

[GDL 10] **Gruppo di Lavoro** avente il compito di definire le procedure per la **gestione operativa**, presso tutte le strutture organizzative coinvolte, delle attività inerenti l'applicazione delle sanzioni: D08-92-DPTS-maggio 2011

Ruolo: Membro

[GDL 9] **Gruppo di lavoro per il Progetto Interdirezionale “Sistema Informativo Statistico delle Pubbliche Amministrazioni – Portale Statistico della PA”**, D08 172 DPTS 10, Durata: Ottobre 2010 – Ottobre 2013

Ruolo: Membro

[GDL 8] **Gruppo di Lavoro per la “Realizzazione degli idonei sviluppi metodologici e l'estensione degli ambiti applicativi del sistema SMART”**, riferimento D08 64 DPTS 10, Durata: Aprile 2010 – Ottobre 2011

Ruolo: Membro

[GDL 7] **Gruppo di lavoro “RELAIS: Metodi e Strumenti per il Record Linkage”**, riferimento: D08-205 – DPTS 10, Durata: dicembre 2010 - novembre 2013,

Ruolo: Membro

[GDL 6] **Gruppo di lavoro per la gestione degli interventi di sviluppo applicativo e/o di manutenzione evolutiva**. Riferimento: D08 165 DPTS 09. Durata: Novembre 2009 –,

Ruolo: Membro

[GDL 5] **Circolo di Qualità “Metodologie e Strumenti Generalizzati”**, D01-UDC-08, 2008

Ruolo: Membro

[GDL 4] **Gruppo di Lavoro “Metodologie e software per l'abbinamento di dati da fonti diverse”** Riferimento: D08 – 122- DPTS- Ottobre 2008, Durata: Ottobre 2008 – Marzo 2010,

Ruolo: Coordinatore

[GDL 3] **Circolo di Qualità “Metodologie e Strumenti Generalizzati”**, D01-UDC-SCS-07, 2007

Ruolo: Membro

[GDL 2] **Gruppo di Lavoro ISTAT-CISIS** avente il compito di definire la fattibilità della produzione delle **stime sul mercato del lavoro**. Riferimento: D08-14-DPTS, Durata: Inizio: marzo 2006 - Fine: settembre 2008

Ruolo: Membro

[GDL 1] **Gruppo di Lavoro** incaricato di procedere alla valutazione di metodi di **tutela della riservatezza** di tipo perturbativo per la pubblicazione di aggregati delle statistiche strutturali sulle imprese. Riferimento: 83/06/DPTS. Durata: Maggio 2006 –,

Ruolo: Membro

Commissioni e Comitati Istat (CI)

[CI 7] **Commissione Big Data** (costituita con Delibera PRES D08-4-2016 del 26/1/2016) e Segreteria Tecnico Scientifico (costituita con Delibera PRES D08-12-2017 del 16/6/2017)

Ruoli: Membro della Commissione e Referente dell'Area “Big Data Reference Architecture” nell'ambito della segreteria tecnica

[CI 6] Commissione Big Data (costituita con Delibera PRES D08-20-2013 del 14/2/2013)

Ruolo: Responsabile segreteria tecnica

[CI 5] **Commissione esaminatrice** del concorso pubblico nazionale per titoli ed esami a 25 posti per il profilo di **collaboratore tecnico enti di ricerca** VI livello professionale dell'Istat, area informatica (Del. 313/PER dell'11/04/2011).

Ruolo: Membro

[CI 4] **Commissione Tecnica** incaricata di aggiudicare le procedure di gara per la **fornitura di beni e servizi informatici** (84/PER del 1/02/2011).

Ruolo: Membro supplente

[CI 3] **Comitato metodologie** con il compito di provvedere all'avvio e al corretto funzionamento della rete per l'innovazione metodologica della produzione statistica dell'Istituto (Del. 110/DPTS/06-07-2010).

Ruolo: Membro

[CI 2] **Commissione esaminatrice** del concorso pubblico nazionale per titoli per il profilo di **dirigente di ricerca** dell'Istat nell'area statistico informatica (3684/06/PER del 10/10/2006).

Ruolo: Segretario

[CI 1] **Commissione esaminatrice** del concorso pubblico nazionale per titoli ed esami per il profilo di **primo ricercatore** di secondo livello professionale dell'Istat nell'area statistico informatica (3685/06/PER del 10/10/2006).

Ruolo: Segretario

Commissioni e Gruppi Inter-Istituzionali (CGI)

[CGI 6] **Gruppo di lavoro inter-istituzionale** (DOP/225/2018) per l'effettuazione di una sperimentazione volta a verificare le potenzialità dell'architettura applicativa definita nell'ambito del progetto **DAF (Data Analytics Framework)** del Team per la Trasformazione Digitale (<https://teamdigitale.governo.it/>).

Ruolo: Coordinatore

[CGI 5] **Accordo Istat - Banca di Italia** sul tema **Privacy-Preserving Record Linkage (PPRL)** per l'integrazione e l'arricchimento informativo tra fonti. (rif. Approvato Comitato di Presidenza Istat-Seduta Aprile 2017, Attestazione attività 18-9-2018)

Ruolo: Coordinamento tecnico e organizzativo per Istat

[CGI 4] **Commissione Valutatrice** per la copertura mediante costituzione di rapporto di lavoro a tempo determinato della posizione di **dirigente responsabile dell'area gestione ed integrazione dati** nell'ambito della Direzione Sistemi Informativi e Agenda Digitale del **Comune di Milano** (rif. Determinazione Dirigenziale 835/2017 del Direttore dell'Area Amministrazione Risorse Umane del Comune di Milano).

Ruolo: Componente Esperto

[CGI 3] **Gruppo di Valutazione** per l'attribuzione di due posizioni dirigenziali di **Data Scientist profilo statistico** e **Data Scientist profilo informatico** nell'ambito della Direzione Centrale Tecnologie e Innovazione dell'**Agenzia delle Entrate** (rif. Gruppo di valutazione, nominato con atto n. 47028 dell'8 marzo 2017).

Ruolo: Componente

[CGI 2] **Gruppo di lavoro su Competenze professionali specialistiche per professionisti e futuri professionisti ICT**, nell'ambito della realizzazione di un "Piano nazionale per la cultura, la formazione e le competenze digitali" promosso da **AgID** (rif. SP/637.2013).

Ruolo: Componente

[CGI 1] **Gruppo di lavoro** finalizzato alla definizione delle **linee guida per i dati di tipo aperto** promosso da **AgID** (rif. SP/225.2013).

Ruolo: Membro rappresentante Istat

Gruppi di Lavoro in ambito Europeo e Internazionale (GLEI)

[GLEI 9] **Eurostat Steering Group on Big Data**. Gruppo ristretto di istituti di statistica europei selezionati sulla base dell'esperienza sul tema. Il gruppo ha carattere strategico e orienta i lavori della Task Force su Big Data (vedi a seguire [GLEI 4]) che ha un carattere prevalentemente operativo (rif. Attestazione attività del 16-10-2019, num. prot. 2766466/19).

Ruolo: Componente Esperto rappresentante Istat

Date: Dal 2018 ad oggi

[GLEI 8] **Eurostat Task Force on the Use of Privately Held Data.** Ha l'obiettivo di definire una roadmap di azioni legali, tecniche e organizzative per facilitare l'accesso a dati forniti da soggetti privati (ad esempio dati di telefonia) da parte degli istituti di statistica europei (rif. <https://webgate.ec.europa.eu/fpfis/wikis/display/EstatBigData/Use+of+privately+held+data+for+official+statistics>) (rif. Attestazione attività del 16-10-2019, num. prot. 2766466/19).

Ruolo: Componente Esperto rappresentante Istat

[GLEI 7] **Eurostat Enterprise Architecture (EA) Board.** Ha l'obiettivo di organizzare e valutare le revisioni EA dei progetti ESS.VIP (European Statistical System Vision Infrastructure Project) assicurando che siano allineati con la ESS EA e ESS Vision 2020 e che gli standard esistenti siano utilizzati correttamente. (rif. https://ec.europa.eu/eurostat/cros/content/ess-ea-board_en e Ref. Ares(2018)4594671 - 07/09/2018).

Ruolo: Componente Esperto rappresentante Istat (solo 7 paesi sono entrati nel board con selezione avvenuta per esperienza sulla tematica)
Date: Dal 2016 ad oggi

[GLEI 6] **UNECE HLG-MOS (High Level Group for Modernization of Official Statistics) Executive Board,** gruppo che coordina i progetti e i gruppi operativi dell'HLG-MOS (Attestazione attività 18-9-2018 e Attestazione attività del 16-10-2019, num. prot. 2766466/19).

Ruolo: co-rappresentante per Istat (insieme al Dott. Piero Falorsi) Date: dal 2016 ad oggi

[GLEI 5] **UNECE Blue Sky Thinking Network** (rif. <https://statswiki.unece.org/display/hlgbas/Modernisation+Groups>). Il Blue Sky Thinking Network è la “fabbrica di idee” per la comunità internazionale di modernizzazione della statistica. I report Blue Skies forniscono visione su argomenti selezionati. Questi documenti rispondono a quattro domande chiave: qual è l'argomento, qual è il valore per le statistiche ufficiali, cosa è già stato fatto, quali sono le opportunità di collaborazione. (Attestazione attività 18-9-2018 e Attestazione attività del 16-10-2019, num. prot. 2766466/19, rif. <https://statswiki.unece.org/display/hlgbas/Modernisation+Groups>)

Ruolo: Componente Esperto e rappresentante Istat nel core group. Ho contribuito al primo Blue Sky Report su "Machine Learning in Official Statistics". Sto contribuendo al Blue Sky Report su “Secure Multi-party Computation and Official Statistics”

Date: dal 2016 ad oggi

[GLEI 4] **Eurostat Task Force su Big Data.** La Task Force ha l'obiettivo di supportare l'adozione del Big Data Action Plan di Eurostat. (Eurostat TF Certification Ref. Ares(2018)4417214 - 28/08/2018).

Ruolo: Membro rappresentante per Istat. Date: dal 2015 al 2018

[GLEI 3] **UNECE Modernization Committee on Products and Sources**, con l'obiettivo di sviluppare ed esplorare le fonti di dati per supportare la modernizzazione e i nuovi prodotti da realizzare per soddisfare le crescenti esigenze degli utenti.

Ruolo: co-rappresentante per l'Istat (insieme al Dott. Stefano De Francisci) dal 2014 al 2016

[GLEI 2] **UNECE CSPA Architecture Working Group**. Obiettivo del gruppo è il supporto al disegno e alla realizzazione di servizi statistici secondo lo standard architetturale Common Statistical Production Architecture (rif. <https://statswiki.unece.org/display/CSPA/Common+Statistical+Production+Architecture>). Ho contribuito attivamente alle attività del gruppo e alla stesura dello standard.

Ruolo: Membro rappresentante per Istat. Date: Dal 2013 al 2014

[GLEI 1] **Eurostat Working Group SISAI (Statistical information Systems Architecture and Integration)**, con l'obiettivo di affrontare temi specifici di architetture informative per la statistica e di riportare i risultati al gruppo dei direttori IT dell'ESS (ITDG).

Ruolo: Membro rappresentante per l'Istat. Date: Dal 2013 al 2014

Incarichi (INC)

[INC 10] Incarico del Direttore della Direzione Centrale per la Metodologia e il Disegno dei Processi Statistici (DCME), Dott. Piero Demetrio Falorsi, di partecipare alla **ESSnet Big Data Pilots I (Fase 2)** (rif. DCME Protocollo 650774/17).

Ruoli: 1) National contact point per Istat, 2) responsabile del progetto per Istat, 3) coordinamento aspetti organizzativi per Istat, nell'ambito del WP0 (Coordination) **4) coordinamento degli aspetti tecnici per tutto il consorzio** del WP2 (Web Scraping Enterprise Characteristics) (10 persone ca.) con coinvolgimento dei paesi Bulgaria, Polonia, Olanda, Svezia e Inghilterra. **5) partecipazione** al WP5 su (Mobile Phone data).
Date: Agosto 2017-Maggio 2018

[INC 9] Incarico di **Responsabile dell'iniziativa “ME Big Data: Progetti di implementazione in ambito europeo”** per gli anni 2018 e 2019 (rif. Ordine di Servizio-ODS 13618 e Ordine di Servizio-ODS 15519).

[INC 8] Incarico di **Responsabile dell'iniziativa “ME Big Data: Progetti pilota nell'ambito di iniziative europee”** per il 2018 Ordine di Servizio-ODS 6718.

[INC 7] Incarico di **Responsabile dell'iniziativa “Linee guida per il Sistema integrato dei registri”** per il 2017 Ordine di Servizio-ODS 8917

[INC 6] Incarico del Direttore della Direzione Centrale per la Metodologia e il Disegno dei Processi Statistici (DCME), Dott. Piero Demetrio Falorsi, di partecipare alla **ESSnet Big Data Pilots I (Fase 1)**. (rif. DCME Protocollo 650773/17)

Ruoli: 1) National contact point per Istat, 2) responsabile del progetto

per Istat, 3) coordinamento aspetti organizzativi per Istat, nell'ambito del WP0 (Coordination) **4) coordinamento degli aspetti tecnici per tutto il consorzio** del WP2 (Web Scraping Enterprise Characteristics) (10 persone ca.) con coinvolgimento dei paesi Bulgaria, Polonia, Olanda, Svezia e Inghilterra. **5) partecipazione** al WP5 su (Mobile Phone data).
Date: Febbraio 2016-Luglio 2017

[INC 5] Incarico del Presidente Prof. Giorgio Alleva, conferito in data 09/10/2014 (SP/683.2014), come **“referente”** per il progetto di **modernizzazione dell'Istat**. I referenti del progetto hanno il ruolo di analizzare le fonti amministrative e le indagini, i progetti strategici, le sinergie e le interrelazioni critiche fra le diverse aree organizzative, la struttura organizzativa ed i flussi decisionali e autorizzativi nell'ambito delle strutture.

[INC 4] Incarico del Presidente Prof. Giorgio Alleva, conferito in data 23/12/2014 (SP/871.2014), per lavorare al **progetto di modernizzazione** dell'Istat con un impegno pari al 40% del proprio tempo.

[INC 3] **Attestazione di incarichi** ricevuti ed attività svolte da parte del Direttore Dott. Francesco Castanò della Direzione Centrale per le tecnologie dell'informazione e della comunicazione (DCIT). Rif. Protocollo 1376 del 6 Ottobre 2015 Periodo: Novembre 2013 – Agosto 2015

[INC 2] **Nota di Encomio** da parte del Direttore Dott. Domenico Donvito della Direzione Centrale per le tecnologie dell'informazione e della comunicazione (DCIT). Rif. Protocollo 1720 del 31 Ottobre 2013

[INC 1] Incarico del Direttore della Direzione Centrale per le tecnologie ed il supporto metodologico (DCMT), Dott. Piero Demetrio Falorsi, di **coordinare** per la DCMT la fasi iniziali di impianto del web data warehouse dell'Istat (i.Stat: dati.istat.it). L'incarico ha avuto carattere aggiuntivo rispetto all'attività ordinaria (rif. DCMT Protocollo 187 del 04/02/2011).

Principali attività e responsabilità

Progetti Europei e Internazionali (PEI)²

[PEI 15] **ESSnet Big Data Pilots-II**. Il progetto è in corso, come proseguimento del progetto ESSnet Big Data Pilots-I (vedi [PEI 10]). Analogamente al progetto ESSnet Big Data Pilots-I è coordinato da Statistics Netherlands. Prevede il coinvolgimento di 22 partner europei (principalmente Istituti di Statistica) per un totale di 12 Workpackage. Istat, nella persona della sottoscritta, coordina il workpackage “Process and Architecture” finalizzato alla definizione di una Big Data Reference Architecture per gli istituti di statistica europei e per il sistema statistico europeo.

Ruoli: Country contact per Istat, coordinatore del workpackage “Process and Architecture”, partecipazione come membro ai workpackage: “Enterprise Characteristics”, “Earth Observation” e

² Alcuni dei progetti elencati hanno previsto dei gruppi di lavoro e degli incarichi e sono dunque citati anche negli altri elenchi. L'obiettivo di questo elenco è di fornire una breve descrizione degli obiettivi dei progetti e della dimensione.

“Smart Statistics”. Date: Novembre 2018 - Dicembre 2020

[PEI 14] **ESSnet Implementing shared services**. Il progetto è coordinato dall'istituto di statistica francese (INSEE) e prevede oltre all'Istat, gli istituti di statistica di Portogallo, Svezia e Norvegia. Obiettivo del progetto è lo sviluppo e la condivisione di servizi software tra istituti nazionali di statistica e altre autorità dell'ESS.

Ruolo: Componente esperto e Referente organizzativo per Istat. Date: 2019-2020

[PEI 13] **ESSnet Linked Open Statistics**. Progetto coordinato dall'istituto di statistica bulgaro e che ha coinvolto oltre l'Istat, gli istituti di statistica di Francia e Irlanda. Ha l'obiettivo di dimostrare il valore del LOD per l'ESS (European Statistical System) e per gli utenti e le parti interessate, di fornire soluzioni sostenibili che possano essere riutilizzate da altri istituti e di costruire una “community of practice” sui LOD all'interno dell'ESS e della comunità internazionale degli uffici statistici ufficiali. (Attestazione attività 18-9-2018)

Ruolo: Membro. Ho contribuito alla scrittura della proposta e ha coordinato gli aspetti di progetto delle ontologie e dei metadati. Date: 2017-2019

[PEI 12] **Eurostat DIGICOM**. Digital communication, User analytics and Innovative products (DIGICOM) è un progetto VIP (Vision Infrastructure Project) finanziato da Eurostat. Ha l'obiettivo generale di migliorare il “dialogo” degli istituti di statistica con i propri utenti, agendo su coinvolgimento proattivo, soluzioni innovative e velocità di reazione. E' organizzato in 4 work package. (Attestazione attività 18-9-2018)

Ruolo: Referente Istat per il Workpackage “Open Data Dissemination”. Date: 2016-2019

[PEI 11] **UNECE Data Architecture**. Progetto che ha l'obiettivo di definire una architettura dati di riferimento (reference data architecture) per gli istituti nazionali di statistica. (Attestazione attività 18-9-2018)

Ruolo: Membro. Ho contribuito attivamente alla stesura della prima versione dell'architettura di riferimento nota come CSDA (Common Statistical Data Architecture) (rif. <https://statswiki.unece.org/download/attachments/129177312/HLG-MOS%20Reference%20Data%20Architecture%20v1.0.pdf>), anche partecipando a sprint sessions con gli altri membri del progetto. Date: Gennaio 2017 -Dicembre 2017"

[PEI 10] **ESSnet Big Data Pilots-I**. ESSnet Big DataPilots-I ha l'obiettivo generale del progetto è di preparare il sistema statistico europeo all'integrazione delle fonti di big data nella statistica ufficiale. L'idea è che l'attività si concentri sull'avvio di progetti pilota che analizzino le potenzialità delle fonti di big data per la produzione o il contributo alla produzione statistica. Il progetto, che vede il coinvolgimento di un consorzio di 22 partner, di cui 20 sono Istituti Nazionali di Statistica e 2 Authorities Statistiche, è coordinato dall'Istituto di statistica dei Paesi Bassi. Il progetto si articola in 10 work package, Istat, nella persona della sottoscritta, coordina il work package

“web scraping enterprise characteristics”.

Ruolo: Responsabile del progetto per Istat, Responsabile del workpackage "Web scraping enterprise characteristics", country-contact Istat, (rif. DCME Protocollo 650773/17 e DCME Protocollo 650774/17). Date: Febbraio 2016 - Luglio 2018

[PEI 9] **UNECE Linked Open Metadata**. Progetto UNECE per la realizzazione di un insieme di classificazioni e modelli in formato LOD (Attestazione attività 18-9-2018).

Ruolo: Rappresentante per Istat e Coordinamento tecnico e organizzativo del gruppo europeo (20 persone.)

Prodotti realizzati: Linee guida per la pubblicazione di classificazioni e modelli in format LOD. Artefacts LOD disponibili su Github e su database <https://unece-stardog.ichec.ie>. Date: Gennaio 2016- Dicembre 2016

[PEI 8] **Eurostat “Facilitation of Methodological Cooperation in the ESS”** (ESS/B/2013/002). Obiettivo del progetto è attuare azioni a supporto di scambi tra esperti di temi specifici di interesse metodologico per la statistica ufficiale. Il mio coinvolgimento ha riguardato il tema "Big Data and Official Statistics".

Ruolo: Esperto e facilitatore per il tema “ Methodology, quality issues and accreditation of Big Data sets”. Partecipazione alla stesura del report:

<https://ec.europa.eu/eurostat/cros/system/files/Concept%20paper%20ESS%20Big%20Data%20Event.pdf> Date: 2014

[PEI 7] **UNECE CSPA (Common Statistical Production Architecture)**. Ho partecipato a due progetti promossi dall'UNECE, CSPA 2013 e CSPA 2014. CSPA 2013 è stato finalizzato alla prima stesura della specifica CSPA (CSPA 1.0) e alla realizzazione prototipale di servizi statistici conformi a tale specifica (CSPA service EditRules Wrapper, CSPA service Autocoding 2). CSPA 2014 ha previsto la revisione della specifica CSPA (CSPA 1.1) e la realizzazione di un ulteriore servizio (CSPA service di Error Correction) (DCMT/RST protocollo 1376 del 6/10/2015).

Ruolo: Membro rappresentante Istat e Coordinamento tecnico e organizzativo delle risorse impegnate nella progettazione e nello sviluppo (Risorse impegnate: 3) Date: 2013-2014

[PEI 6] **Progetto Internazionale Statistical Network CULT**. Obiettivo del progetto è la definizione di un framework comune per il record linkage in Istituti di Statistica (<http://www1.unece.org/stat/platform/display/statnet/CULT+Project>). (DCMT/RST protocollo 1376 del 6/10/2015)

Ruolo: Membro esperto per Istat sul tema privacy-preserving record linkage. Prodotti realizzati: Stato dell'arte su nuove metodologie e software per il record linkage. Stato dell'arte su sistemi software per il record linkage.

Date: 2013-2015

[PEI 5] **Progetto Eurostat Implementing the SDMX Istat Strategy.** Progetto finanziato con un grant bilaterale Istat-Eurostat, finalizzato alla realizzazione della SDMX Reference Infrastructure a supporto della diffusione dei dati Istat. (DCMT/RST protocollo 1376 del 6/10/2015 e 314 del 10/3/2016)

Ruolo: Responsabile del Workpackage “Statistical Open Data: from SDMX to RDF data cube format”. Coordinamento tecnico e organizzativo delle risorse impegnate nel WP (Risorse impegnate: 2). Prodotti realizzati: Realizzazione del web service di diffusione, integrato con la SDMX Reference Infrastructure, per la traduzione on-the-fly da SDMX a RDF Data Cube Vocabulary. Date: 2012-2014

[PEI 4] **CORE (COmmon Reference Environment).** ESSnet finanziata da Eurostat nell’ambito dello Statistical Programme 2010. L’Istat ha coordinato il progetto, con partner gli istituti di statistica di Olanda, Norvegia, Francia, Portogallo e Svezia.

L’obiettivo generale del progetto ha riguardato la progettazione e l’implementazione di un’architettura a supporto della definizione ed esecuzione automatica dei processi statistici.

Il progetto CORE è ritenuto strategico in ambito europeo nell’ambito delle iniziative di industrializzazione e standardizzazione dei processi statistici promosse dall’HLG-BAS (High-level Group for Strategic Developments in Business Architecture in Statistics). (DCMT/RST protocollo 1376 del 6/10/2015, 314 del 10/3/2016)

Ruoli: Technical Manager dell'intero progetto e responsabile dei Work package WP1-Project Management e WP3-Generic interface design for interconnecting GSBPM sub-processes (architettura del sistema). Date: 2010-2012

[PEI 3] **CORA (COmmon Reference Architecture).** ESSnet finanziata da Eurostat nell’ambito dello Statistical Programme 2009. L’Istat ha coordinato il progetto, con partner gli Istituti di Statistica di Olanda, Norvegia, Svizzera, Lettonia e Svezia.

L’ obiettivo generale del progetto ha riguardato un’analisi della fattibilità di una architettura comune ai vari Istituti di Statistica Europei, secondo le dimensioni di business, technology e organization. I principali risultati di CORE sono relativi a: (i) un information model, che è stato un input importante per il successivo progetto CORE, (ii) un repository di software open source utilizzati da Istituti di Statistica Europei, attualmente mantenuto dall’UNECE

(<http://www1.unece.org/stat/platform/display/msis/Software+Inventory>).

Il progetto CORA si è concluso con successo, e l’esito positivo ha determinato il finanziamento per il successivo progetto CORE.

Ruoli: Technical Manager di tutto il progetto e responsabile dei Work package WP1-Project Management e WP5-Project Dissemination. Date: 2009-2010

[PEI 2] **ESSnet Data Integration.** Il progetto è stato condotto dall’Istat,

con partecipazione degli istituti di statistica di Olanda, Norvegia, Polonia, Spagna e Svizzera. Si è occupato della diffusione e del trasferimento di conoscenza di metodi statistici per l'integrazione di fonti dati esistenti nella produzione della statistica ufficiale.

Ruolo: Membro. Ho partecipato alle attività relative al Work package 3 "Methodological Developments", Task 3.1 Record Linkage. Date: 2009-2011.

[PEI 1] **CENEX Statistical Methodology Project - Area Integration of survey and administrative data.** Nel 2005 Eurostat ha promosso dei centri di eccellenza nel campo della statistica, denominati European Centres and Networks of Excellence (CENEX), con l'obiettivo di rafforzare la cooperazione tra gli istituti nazionali di statistica (questi centri si sono poi "evoluti" nelle ESSnet). Questo CENEX coordinato dall'Istat riguarda l'area "Integration of Survey and Administrative Data" e ha previsto come istituti nazionali di statistica partecipanti, oltre all'Istat, l'istituto spagnolo (INE), quello olandese (CBS), quello austriaco (AT) e quello ceco (CSO). (protocollo 2130 del 30-11-2007)

Ruolo: Membro e coordinamento Work package "Software Tools for Integration Methodologies", relativo all'analisi dei sistemi software esistenti per effettuare record linkage probabilistico. Date: 2006-2008

Principali attività e responsabilità

Docenze di corsi Istat ed Europei.

Nel seguito si riporta il dettaglio dell'attività di docenza erogata dalla sottoscritta per conto di Istat. L'attività è di due tipi: (i) docenza interna (indicata con DI) e (ii) docenza nell'ambito di iniziative internazionali (indicata con DE e con dettaglio dell'iniziativa stessa).

[D37] Titolo Corso: **Enterprise Architecture**

Periodo: 22-25 Ottobre 2019

Tipologia: DE, ESTP (European Statistical Training Programme)

[D36] Titolo Corso: **Information Standards and Technologies for Describing, Exchanging and Disseminating Data and Metadata**

Periodo: 10-13 Settembre 2019

Tipologia: DE, ESTP (European Statistical Training Programme)

[D35] Titolo Corso: **Introduzione ai Big Data**

Periodo: 25-26 Giugno 2019

Tipologia: DI

[D34] Titolo Corso: **Enterprise Architecture**

Periodo: 22-25 Ottobre 2018

Tipologia: DE, ESTP (European Statistical Training Programme)

[D33] Titolo Corso: **Introduction to big data and its tools**

Periodo: 12-14 febbraio 2018

Tipologia: DE, ESTP (European Statistical Training Programme)"

[D32] Titolo Corso: **Enterprise Architecture**

Periodo: 16-19 Ottobre 2017

Tipologia: DE, ESTP (European Statistical Training Programme)

[D31] Titolo Corso: **Information Standards**

Periodo: 12-15 Settembre 2017

Tipologia: DE, ESTP (European Statistical Training Programme)

[D30] Titolo Corso: **Information Standards and Technologies for Describing, Exchanging and Disseminating Data and Metadata**

Periodo: 30-31 Marzo 2017

Tipologia: DE, ESTP (European Statistical Training Programme)

[D29] Titolo Corso: **Introduction to big data and its tools**

Periodo: 24-26 Gennaio 2017

Tipologia: DE, ESTP (European Statistical Training Programme)

[D28] Titolo Corso: **Enterprise Architecture and the different EA layers, application to the ESS context**

Periodo: 11-14 Ottobre 2016

Tipologia: DE, ESTP (European Statistical Training Programme)

[D27] Titolo Corso: **Information Standards and Technologies for Describing, Exchanging and Disseminating Data and Metadata**

Periodo: 19-22 Aprile 2016

Tipologia: DE, ESTP (European Statistical Training Programme)

[D26] Titolo Corso: **Introduzione ad un linguaggio comune per saper leggere e interpretare i dati statistici: metadati, standard internazionali e statistica**

Periodo: 21-23 Marzo 2016

Tipologia: DI

[D25] Titolo Corso: **Introduction to big data and its tools**

Periodo: 29 Febbraio - 3 Marzo 2016

Tipologia: DE, ESTP (European Statistical Training Programme)

[D24] Titolo Corso: **Methods and Tools for Record Linkage and Statistical Matching**, SILC Survey Periodo: 08-09 Giugno 2015,

Tipologia: DE, Cooperazione internazionale

[D23] Titolo Corso: **Big Data for Official Statistics**

Periodo: 18-22 Maggio 2015

Tipologia: DE, ESTP (European Statistical Training Programme)

[D22] Titolo Corso: **Introduzione alle Basi di Dati Relazionali e SQL Base**

Periodo: 26-29 gennaio 2015

Tipologia: DI

[D21] Titolo Corso: **Record Linkage and Statistical Matching**

Periodo 10-13 Novembre 2014

Tipologia: DE-ESTP

[D20] Titolo Corso: **Introduzione alle Basi di Dati Relazionali e SQL Base**

Periodo: 10-13 Novembre 2014

Tipologia: DI

[D19] Titolo Corso: **Record Linkage**

Periodo: 16-19/9/2014

Tipologia: DI

[D18] Titolo Corso: **CSPA (Common Statistical Production Architecture)**

Periodo: 24-26 Giugno 2014

Tipologia: DE, ESTP (European Statistical Training Programme)

[D17] Titolo Corso: **Big Data: effective processing and analysis of very large unstructured data for official statistics**

Periodo: 5-9 Maggio 2014

Tipologia: DE, ESTP (European Statistical Training Programme)

[D16] Titolo Corso: **Record Linkage**

Periodo: 19-21 Marzo 2013

Tipologia: DE, ESTP (European Statistical Training Programme)

[D15] Titolo Corso: **CORA (Common Reference Architecture)**

Periodo: 22-23 Gennaio 2013

Tipologia: DE, ESTP (European Statistical Training Programme)

[D14] Titolo Corso: **Record Linkage e Statistical Matching**

Periodo: 28-29 Settembre 2011

Tipologia: DE, Essnet Data Integration

[D13] Titolo Corso: **Training on the Job RELAIS**

Periodo: 5-8 Luglio 2011

Tipologia: DE, Essnet Data Integration

[D12] Titolo Corso: **Record Linkage**

Periodo: 16-19 Marzo 2009

Tipologia: DI

[D11] Titolo Corso: **Advanced Course on R Programming**

Periodo: 15-18 Dicembre 2009

Tipologia: DE

[D10] Titolo Corso: **Introduzione all'integrazione: principi, metodi e strumenti**

Periodo: 17 Febbraio 2009

Tipologia: DI

[D9] Titolo Corso: **Developpement du Systeme d' Informations Statistiques sur les Entreprises (SISE) à L'Institut National de la Statistique Tunisia Relais- un instrument pour le lien probabilistique**

Periodo: 19-21 Gennaio 2009

Tipologia: DE – Cooperazione Internazionale

[D8] Titolo Corso: **Programmazione avanzata con R**

Periodo: 28-30 Maggio 2008

Tipologia: DI

[D7] Titolo Corso: **Programmazione avanzata con R**

Periodo: 19-21 Marzo 2008

Tipologia: DI

[D6] Titolo Corso: **Web services: Progetto e Programmazione in Java**

Periodo: 28-30 Marzo 2007

Tipologia: DI

[D5] Titolo Corso: **L'ambiente R elementi di programmazione e librerie per la produzione statistica**

Periodo: 21-29 Maggio 2007

Tipologia: DI

[D4] Titolo Corso: **The R software for statistical computing**

Periodo: 19-23 Marzo 2007

Tipologia: DE – Cooperazione Internazionale

[D3] Titolo Corso: **L'ambiente R elementi di programmazione e librerie per la produzione statistica**

Periodo: 6-15 Febbraio 2007

Tipologia: DI

[D2] Titolo Corso: **L'ambiente R elementi di programmazione e librerie per la produzione statistica.**

Periodo: 4 Dicembre 2006 -12 Gennaio 2007

Tipologia: DI

[D1] Titolo Corso: **L'ambiente R elementi di programmazione e librerie per la produzione statistica.**

Periodo: 6-10 Novembre 2006

Tipologia: DI

Principali attività e responsabilità

Principali progetti e risultati conseguiti per tema

Nel seguito, descrivo le principali attività e risultati conseguiti in Istat, raggruppandoli per aree. Le quattro principali aree in cui ho lavorato sono:

- (i) Area Big Data,*
- (ii) Area Integrazione, Registri Statistici e Qualità dei dati,*
- (iii) Area Open data e*
- (iv) Area Enterprise Architecture, Standardizzazione di Processi e Sistemi Informativi Statistici.*

I riferimenti a lavori scientifici ed a note e relazioni tecniche rimandano all'elenco delle pubblicazioni presente alla fine di questo documento. I riferimenti ai Gruppi di lavoro, Commissioni, Incarichi e Progetti sono riferiti agli elenchi precedenti

Area Big Data

Data Inizio: Settembre 2013

Data Fine: In corso

Ho lavorato nell'area Big Data a partire dalle prime azioni intraprese per introdurre i Big Data in Istat. A Settembre 2013, sono stata incaricata di presentare la strategia dell'Istat per l'uso dei Big Data alla riunione DGINS (riunione annuale dei presidenti degli Istituti di Statistica Europei), nell'ambito della quale si è prodotto lo "Scheveningen Memorandum Big Data and Official Statistics" (primo documento di visione sull'uso dei Big Data a supporto della Statistica Ufficiale).

Nell'Ottobre 2013, il Presidente Enrico Giovannini ha costituito una commissione di esperti del mondo accademico, industriale e dei media, per la quale ho svolto un ruolo di coordinamento della segreteria tecnica [CI6]. Insieme al Dott. Emanuele Baldacci, ho curato la stesura della "Istat's Roadmap for the Adoption of Big Data for Official Statistics" che ha costituito il principale output della commissione [RT26].

Di recente (Aprile 2019) è stato approvato dal Comitato di Presidenza dell'Istat il documento "Big Data: Proposta di Roadmap per il triennio 2019-2021" [RT40], redatto dalla sottoscritta e dal Dott. Piero Demetrio Falorsi, che delinea obiettivi strategici e operativi in continuità con la precedente roadmap.

Dal 2013 al 2016, nell'ambito dei ruoli di Dirigente di Unità Operativa RST/B (Sviluppo di Sistemi Informatici per la Statistica) [SI5] e RST/G (Gestione di Big data, Open Data e Dati Multidimensionali) [SI6], e di partecipazione al Gruppo di Lavoro Interdipartimentale sui Big Data [GDL16], ho lavorato ai seguenti progetti sperimentali, con ruolo di indirizzo tecnico e coordinamento organizzativo:

- **Progetto Internet-as-a-Data Source.** Progetto finalizzato alla (i) sperimentazione di tecnologie di *scraping* ed (ii) alla realizzazione di un sistema di *web scraping e text mining* per la stima di indicatori calcolati dall'indagine "Uso dell'ICT da parte di Imprese".
- **Progetto Analisi dei Social Media.** Sperimentazione di un sistema di acquisizione ed analisi di dati da Social Media per la stima di indicatori di *sentiment*.

Entrambi i progetti hanno prodotto importanti risultati sperimentali (e.g. [J22], [J18], [C65] e [C54]) che sono stati la base per il consolidamento e la messa in esercizio dei relativi processi di produzione.

A partire dal 2016, il Presidente Giorgio Alleva ha costituito una seconda commissione Big Data [CI7], alla quale partecipo come membro e come referente dell'area "Big Data Reference Architecture". Nell'ambito di tale area, ho coordinato e scritto un capitolo del Rapporto Annuale della Commissione Big Data [RT36].

Dal 2016 ad oggi, nel servizio di cui sono attualmente responsabile, e quindi sotto la mia diretta supervisione (Servizio MEC "Architettura Integrata dei Dati e dei Processi") [SI8], sono presenti **6 iniziative connesse ai Big Data** [INC8][INC9]:

- 253 (conclusa) "ME metodi di calcolo e trattamento statistico dei Big Data 2016". Responsabile: Dott. Diego Zardetto
- 1153 (attiva) "ME metodi di calcolo e trattamento statistico dei Big Data". Responsabile: Dott. Diego Zardetto
- 1301 (attiva) "ME Metodi di Accesso e Data Preparation". Responsabile: Dott. Diego Zardetto
- 1302 (conclusa) "Big data Progetti Pilota nell'ambito di iniziative europee". Responsabile: Monica Scannapieco
- 1406 (attiva) "Big data: Progetti di implementazione in ambito europeo". Responsabile: Monica Scannapieco
- 1145 (attiva) "Laboratorio per l'innovazione". Responsabile: Dott. Enrico Orsini

Le iniziative 253, 1153 e 1301 hanno portato al perfezionamento, consolidamento e messa in produzione come **statistiche sperimentali** dei progetti Internet as a data source e Analisi dei social media. Il progetto Internet as a Data source è stato seguito congiuntamente dai servizi MEC e MEA (responsabile Dott. Giulio Barcaroli). Gli indicatori pubblicati sono consultabili sul sito dell'Istat nella sezione dedicata alle statistiche sperimentali, disponibile a: <https://www.istat.it/it/statistiche-sperimentali/sperimentazioni-su-big-data>

In aggiunta, le attività correnti delle iniziative 1301 e 1153 riguardano: sperimentazione di tecniche di *word embeddings* [C67] per il trattamento e l'analisi di dati testuali e il *deep learning* come approccio automatico alla classificazione di immagini e testi [IT13][C72].

A Marzo 2019, il team Istat, costituito da tre persone del servizio che dirigo, ha conquistato il **secondo posto allo European Big Data Hackaton** https://ec.europa.eu/eurostat/cros/EU-BD-Hackathon_en.

L'iniziativa 1145 ha portato all'inaugurazione a marzo 2018 del primo **Laboratorio per l'Innovazione dell'Istat**, un'infrastruttura per la ricerca che consente di testare le proprie idee in uno spazio fisico attrezzato [RT37].

A partire dal 2016, ho partecipato e coordinato diverse iniziative europee e internazionali. Ho avuto un importante ruolo di coordinamento per Istat della **ESSnet Big data Pilots I** [PEI 10], progetto che ha coinvolto la quasi totalità degli istituti di statistica europei. Per lo stesso progetto, ho coordinato alcuni paesi europei [INC 6] per il Workpackage "Enterprise Characteristics" che ha prodotto un insieme di indicatori sperimentali pubblicati sul sito del progetto all'indirizzo: https://webgate.ec.europa.eu/fpfs/mwikis/essnetbigdata/index.php/WP2_Experimental_statistics

Il progetto ha avuto un secondo finanziamento che è attualmente in corso e terminerà alla fine del 2020: **ESSnet Big Data Pilots II** [PEI15].

Per tale progetto, sono il responsabile per Istat, coordino un gruppo interno all'Istituto di più di 40 persone e coordino il workpackage "Process and Architecture" a cui partecipano 8 paesi (Bulgaria, Danimarca, Estonia, Francia, Germania, Italia, Olanda, Portogallo) [GDL30] [INC9].

Inoltre, per conto di Istat, ho svolto attività di docenza sui Big Data come docente di 8 edizioni di corsi su Big Data dello European Statistical Training Programme (ESTP), ho recentemente fatto una docenza interna su "Introduzione ai Big Data" [D35] e partecipato alla pianificazione della formazione interna su Data Science con il Gruppo di Corso per la realizzazione dell'iniziativa formativa su "Data Scientist" [GDL22].

Infine, ho una attività di produzione scientifica sul tema Big Data, ed in particolare, Big Data quality, pubblicata su riviste e proceedings di conferenze internazionali (e.g. [J25], [J24], [J23] [J21], [J19], [C73], [C72],[C66]).

In ambito Big Data ho partecipato e partecipo a gruppi Eurostat (Big Data Steering Committee [GLEI9], Big Data Task Force [GLEI4], Task Force on the use of Privately held Data [GLEI 8]) e UNECE (Blue Sky Thinking Network [GLEI5] e HLG-MOS Modernization Committee on Products and Sources [GLEI3]).

Per il 2019 sono stata membro del Comitato Programma del Convegno Intermedio della Società Italiana di Statistica (SIS) denominato "Smart Statistical Methods for Smart Applications".

Area Integrazione, Registri Statistici e Qualità dei Dati

Data Inizio: Marzo 2006

Data Fine: In corso

Fin dalla mia assunzione in Istat mi sono occupata di metodi e strumenti per l'integrazione dei dati. All'ultima Conferenza Nazionale di Statistica tenutasi a Luglio 2018, ho curato l'organizzazione di tutte le sessioni del

percorso “Integrazione” (per un totale di 15 lavori invitati sul tema, cfr. <https://www.istat.it/it/tredicesima-conferenza/programma-e-relazioni>).

Da Marzo 2006 ho seguito il progetto **RELAIS (REcord Linkage At IStat)** finalizzato alla progettazione e realizzazione dei metodi di un toolkit per realizzare un processo di record linkage tra due (o più) fonti. RELAIS nasce da un’idea della sottoscritta che ha seguito il progetto di progettazione dei metodi e di realizzazione del sistema fin dall’inizio. La versione 1.0 del software RELAIS è stata rilasciata sul sito Istat a Febbraio 2008, ed è stato il primo sistema software rilasciato dall’Istat con la EUPL (European Union Public License), aderendo alle iniziative indirizzate all’adozione ed alla promozione di software a sorgente aperto. Il sistema, attualmente alla versione 3.0, è disponibile per il download dal sito Istat (sezione Strumenti di elaborazione/Integrazione): <https://www.istat.it/it/metodi-e-strumenti/metodi-e-strumenti-it/elaborazione/strumenti-di-elaborazione/relais>

RELAIS è stato utilizzato, ed è tuttora in utilizzo, in diversi processi di produzione dell’Istat (e.g. i censimenti generali della popolazione e delle abitazioni) e da Istituti di Statistica Europei ed Internazionali (e.g. Australian Bureau of Statistics).

Per il progetto RELAIS ho curato la progettazione dei metodi e dell’intero sistema e il coordinamento delle risorse operative, prima con il coordinamento del gruppo di lavoro [GDL4] da Ottobre 2008 a Marzo 2010 e successivamente con coordinamento di un ulteriore gruppo di lavoro [GDL21]. Inoltre, per conto di Istat, ho svolto attività di docenza su metodi per il record linkage e RELAIS di 8 edizioni di corsi, sia interni che europei.

Dal 2009 al 2010 ho lavorato al **progetto I.stat** che ha avuto come obiettivo principale l’introduzione di un corporate data warehouse in Istat. Infatti, al tempo, la pubblicazione dei dati da parte dell’Istat avveniva in modo non coordinato e si avvertiva la necessità di una centralizzazione dell’informazione diffusa dall’Istituto. Pertanto, nel 2009, sono partite diverse attività che hanno poi portato, alla fine del 2010 alla piena pubblicazione su un sistema di web warehousing, I.stat appunto, di gran parte dei dati Istat.

Ho avuto il ruolo di **coordinatore del progetto per la Direzione Centrale per le tecnologie ed il supporto metodologico (DCMT)** come da [INC1], in cui si evidenzia il carattere di incarico aggiuntivo rispetto all’attività ordinaria.

Da Novembre 2011 a Dicembre 2014, ho lavorato al **progetto DIDAC (Diffusione Dati Censuari)**, relativo allo sviluppo e alla gestione del sistema di diffusione dei dati dei censimenti generali. Il progetto ha avuto una notevole complessità che si manifesta nell’elevato numero di workpackage costitutivi pari a 22 e nell’elevato numero di risorse coinvolte (circa 60).

Il progetto ha portato alla progettazione e realizzazione del Data Warehouse del XV Censimento generale della Popolazione e delle Abitazioni dell’Istat a supporto della diffusione nazionale dei dati censuari. Il progetto ha anche prodotto il Data Warehouse a supporto della diffusione europea dei dati censuari [IS6].

Ho contribuito al progetto come membro dello staff di **coordinamento**

tecnico con ruolo di vice-responsabile dell'intero progetto ([GDL12]). In aggiunta nel ruolo di **responsabile di team tecnico** del Workpackage Data Warehouse Primario–Infrastruttura tecnologica ed Impianto [GDL18], ho curato l'impianto del data warehouse che ha previsto l'utilizzo di soluzioni progettuali e tecnologiche innovative, che includono l'utilizzo di sistemi di Business Intelligence e OLAP. Ho inoltre coordinato e realizzato e la diffusione dei dati censuari a livello di sezione di censimento come Linked Open Data (LOD) [IS6].

Dal 2016 ad oggi, lavoro alla realizzazione del progetto Sistema Integrato dei Registri (SIR), che si colloca nell'ambito del programma strategico dell'Istat PG1-“Costruzione del Sistema Integrato dei Registri”, che risulta correntemente essere il programma strategico che impegna il maggior numero di risorse dell'Istituto.

Dal 2016 ad oggi, nel servizio di cui sono attualmente responsabile (Servizio MEC “Architettura Integrata dei Dati e dei Processi”), sono presenti 2 linee di attività connesse al SIR:

- 245 Linee Guida per il Sistema Integrato dei Registri (Responsabile attuale: Ing. Roberta Radini, Responsabile precedente: Monica Scannapieco) [INC7] e (protocollo 2130 del 30-11-2007)
- 1199 Architettura Informativa. Responsabile: Ing. Roberta Radini

Il mio ruolo è stato di impostazione tecnica e coordinamento organizzativo di tutte le attività di tipo architeturale relative al Sistema Integrato dei Registri, ed in particolare:

- Modellazione concettuale delle ontologie dei domini di produzione dell'Istat per il Sistema Integrato dei Registri e per la diffusione esterna all'Istituto (sito Istat: <https://www.istat.it/it/metodi-e-strumenti/ontologie>)
- Progettazione e sviluppo dei prototipi e delle prime versioni consolidate del:
 - Registro Base degli individui (RBI), prima versione rilasciata del registro (Gennaio 2018).
 - Registro del Lavoro (RTL) Prima versione rilasciata del registro RTL (Luglio 2018).
 - Registro Statistico di Base dei Luoghi- componente indirizzi (RSBL/IND). Prima versione rilasciata del registro RTL (Giugno 2018).

Il documento [RT38] “Il Sistema Integrato dei Registri: architettura Informativa e Metodologica”, contiene lo stato di avanzamento e le linee guida di realizzazione di tutto il sistema (rif. [INC 7]).

Rilevo il carattere innovativo dell'impostazione architeturale del sistema, basata su un approccio di integrazione semantica dei dati e condiviso in contesti nazionali ed internazionali (e.g. [C69], [C68], [C64], [C63], [C61] e [IT12]).

Ho partecipato a diversi progetti internazionali sull'integrazione come membro (rif. [PEI1], [PEI2], [PEI5], [PEI6] e [PEI13]), con ruolo di referente per Istat (rif. [PEI 12]) e con ruolo di coordinamento (rif. [PEI9] e [PEI2]).

Sul piano nazionale, sono stata nominata coordinatore (insieme al Dott.

Natale Renato Fazio e alla Dott.ssa Marina Venturi) di un **gruppo inter-istituzionale** (rif. [CGI6]) **che ha coinvolto Istat e il Team per la Trasformazione Digitale per la realizzazione del Data Analytics Framework (DAF)**. In quest'ambito ho coordinando le attività per le modellazione delle ontologie, alcune realizzate congiuntamente da Istat e dal Team per la Trasformazione Digitale.

Mi sono anche occupata per Istat di un tema di frontiera relativo all'integrazione privata dei dati (privacy-preserving data sharing). Nell'aprile 2017, **Istat e Banca d'Italia hanno stipulato un accordo di collaborazione** [rif. CGI5], nell'ambito del quale ho coordinato il gruppo tecnico che ha condotto una sperimentazione su *private set intersection*, ovvero linkage esatto in presenza di requisiti di privacy [RT39].

Sono attualmente membro per Istat, insieme al Dott. Natale Renato Fazio, della Task Force Eurostat on the Use of Privately Held Data [GLEI8].

Ho una produzione scientifica dedicata al tema dell'integrazione dati, ed in particolare: record linkage (e.g. [CH5], [J16], [J15], [J13], [J10]), qualità dei dati e integrazione multi-fonte (e.g., [CH3], [J25], [J12], [J11], [J8], [C69],[C58]). Ho scritto due libri pubblicati dalla Springer su qualità dei dati e dell'informazione ([B3], [B1]). Sono stata *program chair* della diciassettesima edizione della *International Conference on Information Quality* (ICIQ) e sono *Associate Editor* dell'*ACM Journal on Data and Information Quality* dal 2012.

Area Open Data

Data Inizio: Novembre 2013

Data Fine: In corso

A partire dal 2013, sono stata tra i primi in Istat ad introdurre l'uso dei modelli, dei linguaggi e delle tecnologie del Web semantico per la diffusione dei dati dell'Istat.

Dal 2013 al 2015, ho lavorato alla progettazione e realizzazione del **portale Linked Open Data dell'Istat** (progetti IS-LOD e CENSUS LOD, rif. [INC2] [INC3]). Il portale, disponibile al sito datiopen.istat.it, è un importante esempio di innovazione su aspetti di integrazione di dati, ha un numero di accessi che lo posizionano tra i siti più visitati tra quelli di diffusione dell'Istituto ed è stato motivo di lustro per l'Istituto in vari contesti nazionali e internazionali.

Il mio ruolo è stato di ideazione dell'intero progetto e di coordinamento tecnico dello stesso [INC2][INC3].

Nell'ambito del [CGI 1], **Gruppo di Lavoro finalizzato alla definizione delle linee guida per i dati di tipo aperto**, promosso dall'**Agenzia per l'Italia Digitale (AgID)**, ho lavorato attivamente alla produzione delle "Linee Guida per la Valorizzazione del Patrimonio Informativo Pubblico". Insieme ad AgID, ho lavorato alla pubblicazione di classificazioni in formati RDF (Resource Description Framework) (e.g. <https://www.istat.it/it/archivio/17888>)[C52].

In collaborazione con AgID e DAF [(rif. [CGI6])], ho coordinato per Istat le attività organizzative e tecniche per la pubblicazione di ontologie OWL (Web Ontology Language) (e.g. <https://www.istat.it/it/archivio/217525>).

Ho partecipato e partecipo correntemente a diversi progetti europei ed internazionali sul tema open data:

- Nell'ambito del progetto Eurostat "Implementing the SDMX Istat Strategy", ho coordinato le attività tecniche di realizzazione di un traduttore da SDMX a RDF [C55].
- Progetto DIGICOM [PEI12] con ruolo di referente Istat per il workpackage su Open Data Dissemination.
- UNECE Linked Open Metadata, in cui ho lavorato con un ruolo di coordinamento tecnico ed organizzativo insieme al Dott. Franck Cotton dell'INSEE di un gruppo internazionale che ha prodotto "artifact" (ontologie, classificazioni e sistemi di gestione) in formato LOD. Cito tra gli artifact, il sistema Classification Explorer per il quale abbiamo ricevuto un **award** da **Oracle** come **vincitori della Challenge Semstats 2016** [C59].
- Progetto ESSnet Linked Open Statistics [PEI 13], del quale ho curato la stesura della proposta e che ho seguito indirizzando le attività relative al progetto delle ontologie.

Ho una produzione scientifica dedicata sull'area open data ed in particolare Linked Open Data (e.g. [J17], [C62], [C60], [C59], [C48]). Sono stata per tre anni (2014-2016) Program Chair del workshop "Semantics Statistics", tenutosi nell'ambito della *International Semantic Web Conference (ISWC)*.

Area Enterprise Architecture, Standardizzazione di Processi e Sistemi Informativi Statistici

Data Inizio: 2009

Data Fine: In corso

Dal 2009 al 2012, ho seguito due progetti europei (ESSnet) sul tema Enterprise Architecture per Istituti Nazionali di Statistica. I due progetti sono **CORA (COMmon Reference Architecture)** [PEI3] e **CORE (COMmon Reference Environment)** [PEI4].

Il mio ruolo in entrambi i progetti è stato di **Technical Manager** per l'intero progetto e di responsabile di Workpackage organizzativi e tecnici (Rif. [PEI3] e [PEI4]).

La ricaduta in Istat derivante dall'esperienza acquisita in questi progetti è avvenuta principalmente nell'ambito delle attività dell'Unità Operativa "Sviluppo di sistemi informatici per la statistica" [SI5]. Come evidenziato in [INC2], mi si riconosce un contributo "nel coordinamento dello sviluppo di sistemi informatici innovativi a supporto della produzione statistica con particolare riferimento ai progetti Sanzioni per la gestione dell'iter sanzionatorio a carico della Direzione Generale, ReGe (Registro Generale) per l'acquisizione trimestrale dei dati giudiziari delle procure italiane, Web Ranker per il calcolo online di indicatori sintetici e Nuova Procedura Famiglie. Le attività hanno costituito una sostanziale evoluzione verso l'industrializzazione e l'armonizzazione delle soluzioni informatiche per la produzione statistica."

Dal 2013 ad oggi, ho continuato a seguire le attività internazionali ed europee sul fronte Enterprise Architecture mediante:

- Partecipazione all'Eurostat Working Group SISAI (Statistical Information Systems Architecture and Integration) [GLEI 1].
- Partecipazione all'UNECE CSPA Architecture Working Group [GLEI 2].
- **Eurostat Enterprise Architecture (EA) Board [GLEI 7]**, incarico particolarmente prestigioso data la selezione effettuata dei soli Paesi esperti sul tema.
- UNECE Data Architecture [PEI 11].
- ESSnet Implementing Shared Services [PEI 14].

In ambito Istat, ho portato avanti alcuni significativi progetti di ottimizzazione dei processi ed impianto di Enterprise Architecture. In particolare:

- **Progetto COE**, finalizzato alla reingegnerizzazione del sistema informativo del servizio commercio estero e attività internazionali delle imprese, con **ruolo di coordinatore** [GDL20]
- **Progetto Nuova Raccolta Dati**, finalizzato al disegno dei processi e delle funzionalità della nuova direzione di Raccolta Dati prevista dal Progetto di Modernizzazione dell'Istat, con **ruolo di referente dell'attività** [INC2][INC3] e **membro della Task Force di migrazione** [GDL25].
- **Progetto Riorganizzazione Finanza Pubblica**, finalizzato all'analisi e reingegnerizzazione del processo di compilazione delle statistiche di Finanza Pubblica all'interno dei Conti Nazionali, con ruolo di Referente della modellazione dei processi, dei dati e dei servizi [GDL24].
- **Progetto Anticorruzione**, finalizzato a modellare i processi per identificare rischi di corruzione e programmare misure di prevenzione, con **ruolo di responsabile** [GDL27].

Dal 2016 ad oggi, sono **responsabile** di un servizio, **Servizio MEC "Architettura Integrata dei Dati e dei Processi"**, che ha nella sua declaratoria il compito di provvedere "alla definizione e al monitoraggio dell'architettura di impresa dell'Istat" [SI8]. Nell'ambito del servizio è presente una **iniziativa dedicata all'Enterprise Architecture**:

- 1119 - ME Enterprise Architecture: architettura dei Servizi e dei Processi nel contesto del Processo Unico Responsabile: Dott. Mauro Bruno.

L'iniziativa ha prodotto il documento "Linee Guida dell'Enterprise Architecture in Istat" [RT31], che contestualizza e specializza l'Enterprise Architecture Reference Framework di Eurostat al contesto Istat.

Di particolare rilevanza, il **Progetto SINTESI** Sistema INTEgrato per le Statistiche sulle Imprese - che ha come obiettivo la realizzazione di un sistema generalizzato, che permetta di unificare metodi e strumenti per la gestione delle indagini economiche sulle imprese. Il progetto è coordinato dal Dott. Mauro Bruno, afferente al servizio MEC di cui sono responsabile e costituisce il più recente investimento dell'Istat sul tema della industrializzazione dei processi statistici [C70].

Sono inoltre stata docente per Istat di corsi specifici su Enterprise Architecture e standard di dati e processi (4 corsi di EA, 5 corsi su standard dati e 2 corsi su standard di processo sia in ambito europeo che interno).

Ho inoltre una produzione scientifica dedicata sul tema Enterprise Architecture e ottimizzazione di processi (e.g., [Ch5], [J22], [J6], [C70], [C57], [C56], [C41]).

Date	Maggio 2017
Lavoro o posizione ricoperti	Consulenza per attività scientifica
Nome e indirizzo del datore di lavoro	Sapienza Università di Roma
Principali attività e responsabilità	Consulenza come Quality manager del Progetto ETER (European Tertiary Educational Register) finalizzato a costruire un registro contenente dati su istituti di istruzione superiore (Higher Education Institution) europei raccolti mediante un processo dedicato [C75]
Tipo di attività o settore	Ricerca
Date	Novembre 2015 – Dicembre 2015
Lavoro o posizione ricoperti	Consulenza per attività scientifica
Nome e indirizzo del datore di lavoro	Statistical Office of the Republic of Serbia (SORS)
Principali attività e responsabilità	La collaborazione ha avuto per oggetto l'assistenza tecnica all'implementazione del progetto "Data Collection Project Income living conditions statistics (SILC)" mediante la progettazione ed implementazione di un processo generale di integrazione dei dati in grado di includere nel repertorio ILC le fonti amministrative disponibili.
Tipo di attività o settore	Ricerca
Date	Luglio 2014 –Dicembre 2014
Lavoro o posizione ricoperti	Consulenza per attività scientifica
Nome e indirizzo del datore di lavoro	Sapienza Università di Roma
Principali attività e responsabilità	Consulenza come Quality manager del Progetto ETER (European Tertiary Educational Register) [IS7]
Tipo di attività o settore	Ricerca

Date	Giugno 2014 –Luglio 2014
Lavoro o posizione ricoperti	Consulenza per attività scientifica
Nome e indirizzo del datore di lavoro	Food and Agriculture Organization of the United Nations (FAO)
Principali attività e responsabilità	Progettazione e realizzazione di un repository semantico di risorse Web.
Tipo di attività o settore	Agenzia Specializzata Organizzazione Intergovernativa (ONU)
Date	Aprile 2011-Marzo 2012
Lavoro o posizione ricoperti	Contratto di collaborazione coordinata e continuativa per attività di ricerca
Nome e indirizzo del datore di lavoro	Dipartimento di Informatica e Sistemistica Università di Roma “La Sapienza”
Principali attività e responsabilità	La collaborazione ha per oggetto lo svolgimento di attività di studio e sperimentazione relativa all'utilizzo di tecniche di record linkage per il mining di processi a partire da dati semi-strutturati (e.g. email).
Tipo di attività o settore	Ricerca
Date	Luglio 2009 – Giugno 2010
Lavoro o posizione ricoperti	Contratto di collaborazione coordinata e continuativa per attività di ricerca
Nome e indirizzo del datore di lavoro	Dipartimento di Informatica e Sistemistica Università di Roma “La Sapienza”
Principali attività e responsabilità	La collaborazione ha avuto per oggetto lo svolgimento di attività di studio e sperimentazione relativa alla tematica dell'Object Matching applicata al contesto Web (Web Object Identification).
Tipo di attività o settore	Ricerca
Date	Aprile 2008 – Dicembre 2008
Lavoro o posizione ricoperti	Contratto di collaborazione coordinata e continuativa per attività di ricerca
Nome e indirizzo del datore di lavoro	Dipartimento di Informatica e Sistemistica Università di Roma “La Sapienza”
Principali attività e responsabilità	La collaborazione ha avuto per oggetto lo svolgimento di attività di studio e sperimentazione relativa al problema del record linkage affrontato con un approccio basato su algoritmi evolutivi.
Tipo di attività o settore	Ricerca
Date	Giugno 2007 – Dicembre 2007
Lavoro o posizione ricoperti	Contratto di collaborazione coordinata e continuativa per attività di ricerca
Nome e indirizzo del datore di lavoro	Dipartimento di Informatica e Sistemistica Università di Roma “La Sapienza”
Principali attività e responsabilità	La collaborazione ha avuto per oggetto lo svolgimento di attività di studio di meccanismi per l'integrazione dei dati in modalità <i>Privacy Preserving</i> e studio di algoritmi per l'elaborazione delle interrogazioni in sistemi Peer to Peer con attenzione alla qualità dei dati e alla privacy.
Tipo di attività o settore	Ricerca
Date	Giugno 2006 – Maggio 2007

Lavoro o posizione ricoperti	Contratto di collaborazione coordinata e continuativa per attività di ricerca
Nome e indirizzo del datore di lavoro	Dipartimento di Informatica e Sistemistica Università di Roma “La Sapienza”
Principali attività e responsabilità	La collaborazione ha avuto per oggetto lo svolgimento di attività di su “Record Linkage Dinamico mediante Combinazione di Tecniche Empiriche e Probabilistiche”.
Tipo di attività o settore	Ricerca
Date	Maggio 2004 – Giugno 2004
Lavoro o posizione ricoperti	Contratto di collaborazione coordinata e continuativa per attività di ricerca
Nome e indirizzo del datore di lavoro	Dipartimento di Informatica e Sistemistica Università di Roma “La Sapienza”
Principali attività e responsabilità	La collaborazione ha avuto per oggetto lo svolgimento di attività di ricerca su problemi di qualità dei dati nei sistemi adattivi multicanale.
Tipo di attività o settore	Ricerca
Date	Aprile 2004
Lavoro o posizione ricoperti	Contratto di collaborazione coordinata e continuativa per attività di ricerca
Nome e indirizzo del datore di lavoro	Dipartimento di Informatica e Sistemistica Università di Roma “La Sapienza”
Principali attività e responsabilità	La collaborazione ha avuto per oggetto l’Usabilità ed Accessibilità di Contenuti di e-Learning.
Tipo di attività o settore	Ricerca
Date	Novembre 2003 – Marzo 2004
Lavoro o posizione ricoperti	Contratto di collaborazione coordinata e continuativa per attività di ricerca
Nome e indirizzo del datore di lavoro	Dipartimento di Informatica e Sistemistica Università di Roma “La Sapienza”
Principali attività e responsabilità	La collaborazione ha avuto per oggetto Classificazione e Confronto di Dimensioni della Qualità dei Dati nel Contesto della Pubblica Amministrazione Italiana.
Tipo di attività o settore	Ricerca
Date	Gennaio 2004 – Marzo 2004
Lavoro o posizione ricoperti	Collaborazione coordinata e continuativa
Nome e indirizzo del datore di lavoro	Istat , Via Balbo 16, Roma
Principali attività e responsabilità	La collaborazione ha avuto per oggetto lo studio della problematica della qualità dei dati toponomastici in grandi basi di dati amministrative.
Tipo di attività o settore	Ente Pubblico di Ricerca
Date	Giugno 2003 – Novembre 2003
Lavoro o posizione ricoperti	Collaborazione coordinata e continuativa
Nome e indirizzo del datore di lavoro	Istat , Via Balbo 16, Roma
Principali attività e responsabilità	La collaborazione ha avuto per oggetto lo studio della problematica della qualità dei dati toponomastici in grandi basi di dati amministrative.
Tipo di attività o settore	Ente Pubblico di Ricerca

Date	Ottobre 2000 – Dicembre 2000
Lavoro o posizione ricoperti	Incarico Professionale
Nome e indirizzo del datore di lavoro	AT&T Laboratories , Indirizzo: Florham Park, New Jersey, USA

Principali attività e responsabilità La collaborazione ha avuto per oggetto lo studio di nuovi modelli per la rappresentazione di dati tramite XML e si è svolta presso gli AT&T Research Lab.

Tipo di attività o settore Ricerca

Date	Dicembre 2000 – Maggio 2001
Lavoro o posizione ricoperti	Incarico professionale
Nome e indirizzo del datore di lavoro	Scuola Superiore della Pubblica Amministrazione , Presidenza del consiglio dei Ministri

Principali attività e responsabilità Organizzazione di corsi sui metodi e le tecnologie della società dell'informazione.

Tipo di attività o settore Ente Pubblico

Date	Aprile 2000 – Ottobre 2000
Lavoro o posizione ricoperti	Contratto a tempo indeterminato
Nome e indirizzo del datore di lavoro	Centro di Ricerca in Matematica Pura ed Applicata (CRMPA) , Via Ponte Don Melillo, Salerno

Principali attività e responsabilità

- Docenza in corsi di formazione sulle piattaforme a componenti e sulle tecnologie middleware.
- Collaborazione su tematiche riguardanti la reingegnerizzazione (BPR) dei processi aziendali e l'applicabilità di WfMS (Workflow Management System).
- Ideazione e stesura della proposta di progetto MURST (Ministero Università e Ricerca Scientifica e Tecnologica) "GAP-Grid Application Provision".

Tipo di attività o settore Ricerca

Date	Dal 2002 al 2011
Lavoro o posizione ricoperti	Docenza a Contratto
Nome e indirizzo del datore di lavoro	Università degli Studi di Roma "La Sapienza"

Principali attività e responsabilità Nell'anno accademico 2010/2011 è stata incaricata della docenza a contratto del corso di Basi di Dati (3 crediti), Corso di Laurea in Ingegneria Gestionale, dell'Università degli Studi di Roma "La Sapienza".

Nell'anno accademico 2008/2009 è stata incaricata della docenza a contratto del corso di Basi di Dati (6 crediti), Corso di Laurea in Ingegneria Gestionale, dell'Università degli Studi di Roma "La Sapienza".

Nell'anno accademico 2007/2008 è stata incaricata della docenza a contratto del corso di Progettazione del Software (5 crediti), Corso di Laurea in Ingegneria Gestionale, dell'Università degli Studi di Roma "La Sapienza".

Nell'anno accademico 2006/2007 è stata docente a contratto del corso di

Progettazione del Software (5 crediti), Corso di Laurea in Ingegneria Gestionale, dell'Università degli Studi di Roma "La Sapienza".

Nell'anno accademico 2005/2006 è stata docente a contratto del corso di Progettazione del Software (5 crediti), Corso di Laurea in Ingegneria Gestionale, dell'Università degli Studi di Roma "La Sapienza".

Nell'anno accademico 2004/2005 è stata docente a contratto del corso di Progettazione del Software (5 crediti), Corso di Laurea in Ingegneria Gestionale, dell'Università degli Studi di Roma "La Sapienza".

Nell'anno accademico 2003/2004 è stata docente a contratto del corso di Progettazione del Software (5 crediti), Corso di Laurea in Ingegneria Gestionale, dell'Università degli Studi di Roma "La Sapienza".

Nell'anno accademico 2002/2003 è stata tutor del corso di Progettazione del Software, Corso di Laurea in Ingegneria Gestionale, dell'Università degli Studi di Roma "La Sapienza", docente prof. Giuseppe De Giacomo.

Tipo di attività o settore Formazione e ricerca

Altri Titoli Scientifici e Professionali

Partecipazioni ad altri progetti di ricerca

In aggiunta alle attività svolte presso l'Istat, come precedentemente dettagliato, nell'ambito del Dottorato di Ricerca in Ingegneria Informatica, e successivamente come assegnista di ricerca e collaboratore scientifico, da Novembre 2000 ho collaborato con il Dipartimento di Informatica e Sistemistica dell'Università di Roma "La Sapienza" ai seguenti progetti di ricerca nazionali ed internazionali, in cui spesso ho rivestito il ruolo di coordinatore di attività e risorse. In particolare:

- 2013-2017: European Tertiary Education Register (ETER), rilevazione dati europea finanziata da Eurostat, responsabile del progetto Prof. Cinzia Daraio, Dipartimento di Ingegneria Informatica, Automatica e Gestionale, A. Ruberti, "Sapienza" Università di Roma. Ruolo: "Quality Manager".
- 2011-2012: MailOfMine project, ricerca cofinanziata dall'Unione Europea all'interno del Settimo Programma Quadro 2007-2013 attraverso il progetto "Smart Vortex – management and analysis of massive data streams to support large-scale collaborative engineering projects", responsabile del progetto prof. Tiziana Catarci, Dipartimento di Informatica e Sistemistica A. Ruberti, "Sapienza" Università di Roma.
- 2009-2010: Progetto di Ateneo Federato dal titolo "Web Object Matching", C26F09BY73, ricerca finanziata per 24 mesi, ricerca finanziata per 24 mesi, responsabile della ricerca prof. Tiziana Catarci, Dipartimento di Informatica e Sistemistica A. Ruberti, "Sapienza" Università di Roma.
- 2007-2008: Progetto di Ricerca di Università (ex Progetto di Ateneo) C26A074R53, "Metodi per migliorare l'accuratezza e la consistenza

- dei dati basati su algoritmi evolutivi”, ricerca finanziata per 24 mesi, responsabile della ricerca prof. Tiziana Catarci, Dipartimento di Informatica e Sistemistica A. Ruberti, “Sapienza” Università di Roma.
- 2006-2007: progetto ESTEEM (Emergent Semantics and cooperATIOn in multi-knowledgE EnvironMents) (PRIN 2005) (durata biennale da Gennaio 2006 a Dicembre 2007), responsabile scientifico prof. Tiziana Catarci. Le attività svolte riguardano il coordinamento della linea di ricerca connessa al processamento trust- and quality-aware di dati in contesti P2P. Il mio ruolo è stato di indirizzo scientifico e di coordinamento organizzativo del gruppo di lavoro del Dipartimento di Informatica e Sistemistica (2 persone), nonché di supporto al coordinatore del progetto.
 - 2005-2008: progetto eG4M (eGovernment for Mediterranean Countries) (FIRB 2003) (durata triennale da Novembre 2005 a Ottobre 2008), responsabile scientifico prof. Tiziana Catarci. Le attività svolte riguardano il coordinamento della linea di ricerca connessa alla qualità dei documenti elettronici ed in particolare alle tecniche di misurazione, controllo e miglioramento di tale qualità. Il mio ruolo è stato di indirizzo scientifico e di coordinamento organizzativo del sottogruppo di lavoro sulla qualità del Dipartimento di Informatica e Sistemistica (2 persone).
 - 2003-2006: rete di eccellenza europea “INTEROP” (durata triennale da Novembre 2003 a Ottobre 2006), finanziato dalla Commissione Europea nell'ambito del programma quadro IST FP 6, responsabile scientifico prof. Tiziana Catarci; nell'ambito della rete di eccellenza la tematica di interesse è quella di modellazione dei dati tramite ontologie. Il mio ruolo è stato di indirizzo scientifico e di coordinamento organizzativo del gruppo di lavoro del Dipartimento di Informatica e Sistemistica (composto da 2 persone)
 - 2003-2005: Progetto nazionale “Le nuove tecniche di governo e l'innovazione tecnologica nella formazione e attuazione delle politiche pubbliche - Metodologia di analisi della problematica di accesso ai dati pubblici" (Progetti Attivi di Amministrazione dello Stato – Legge FISR) (durata biennale da Settembre 2003 a Settembre 2005), responsabile scientifico prof. Tiziana Catarci. In particolare, in tale contesto, ho avuto il compito di studiare le problematiche di accesso ai dati pubblici e di caratterizzarli rispetto al relativo livello di qualità. Il mio ruolo è stato di indirizzo scientifico e di coordinamento organizzativo del gruppo di lavoro del Dipartimento di Informatica e Sistemistica (composto da 2 persone).
 - 2002-2005: progetto MAIS (Multichannel Adaptive Information Systems) (FIRB 2001) (durata triennale da Novembre 2002 a Ottobre 2005, poi esteso a Giugno 2006), responsabile scientifico prof. Tiziana Catarci). L'ambito di competenza ha riguardato la progettazione di un repository di qualità del servizio. Il mio ruolo è stato di membro del gruppo di lavoro del Dipartimento di Informatica e Sistemistica (composto da 5 persone).
 - 2002-2003: progetto nazionale “DaQuinCIS - Data Quality in Cooperative Information Systems" (durata biennale dal Gennaio 2002 a Dicembre 2003), finanziato dal MIUR (COFIN 2001), responsabile scientifico prof. Tiziana Catarci. Le competenze scientifiche hanno

riguardato: la definizione di un modello di cooperazione specificatamente orientato alla qualità dei dati, e la progettazione di un'architettura cooperativa per il supporto alla gestione ed al miglioramento della qualità dei dati. Inoltre, ho avuto il ruolo di (i) di indirizzo scientifico e di coordinamento organizzativo del gruppo di lavoro del Dipartimento di Informatica e Sistemistica (composto da 3 persone), e (ii) di assistente del responsabile scientifico prof. Tiziana Catarci su tutte le tematiche organizzative ed amministrative del progetto, di cui il Dipartimento di Informatica e Sistemistica è stato coordinatore nazionale.

Partecipazione a Comitati di Programma e Comitati Editoriali di Riviste

- **Membro** del Comitato Scientifico della International Scientific Conference on New Techniques and Technologies for Statistics (NTTS) nel 2021.
- **Membro** del Comitato Programma del Convegno Intermedio SIS 2019 –“Smart Statistical Methods for Smart Applications”, 2019.
- **Panel Chair** ACM womENCourage – Women in Computing, 2019
- **Program Chair** International Workshop on Semantics Statistics, in conjunction with International Conference on Semantic Web per gli anni 2014, 2015, 2016
- **Associate Editor** dell'ACM International Journal on Data and Information Quality dal 2012.
- **Membro dell'Editorial Board** e **Associate Editor** dell'International Journal on Information Quality (ISSN (Online): 1751-0465 & ISSN (Print): 1751-0457).
- **Program Chair** della 17th International Conference on Information Quality (ICIQ 2012).
- **Program Chair** del 1st INTEROP Dissemination Workshop tenutosi in cooperazione con OnTheMove (OTM 2004) Federated Conferences a Cipro, Grecia, il 27 Ottobre 2004.
- **Program Chair** dell'International Workshop on Information Quality in Information Systems, tenutosi in cooperazione con 23 rd ACM SIGMOD International Conference on Management of Data a Parigi, Francia, il 18 Giugno 2004.
- **Guest editor** dell'IEEE Special Issue Data Quality in the Internet Area, 2011 [Ed3].
- **Membro** del gruppo IFIP 2.6 su “Data Semantics” dal 2004 e ne è stata segretario dal 2004 al 2007.
- **Guest editor** dello Special Issue di Information Systems (Elsevier) “Interoperability in Open Information Systems through Data and Service Integration” (Data di pubblicazione: Ottobre 2008).
- **Program Chair** dell'International Workshop on Semantic Data and Service Integration (SDSI 2007) tenutosi in cooperazione con VLDB 2007 a Vienna, Austria, il 23 Settembre 2007.
- **Program Chair** del INTEROP-MIOS Workshop tenutosi in cooperazione con On-TheMove (OTM 2005) Federated Conferences a Cipro, Grecia, il 31 Ottobre 2005.
- **Local Chair** della 4th International Conference on Web Information

Systems Engineering (WISE 2003), tenutasi a Roma il 10-13 Dicembre 2003.

- **Local Chair** dell' International Workshop on Data Quality in Cooperative Information Systems, tenutosi a Siena il 10-11 Gennaio 2003.

Seminari Invitati e Tutorial

- Seminario invitato “Big Data @ Istat: Current and Future Challenges”, Università di Bologna, Ottobre 2019.
- Seminario invitato “Use of machine learning for estimating enterprises characteristics from web data” al convegno Methods for Big Data in Official Statistics, Herleen, Olanda, Ottobre 2018.
- Seminario invitato “Big Data Linkage: how to deal with unstructured data and private data”, Università di Bologna, Ottobre 2018.
- Seminario invitato “New sources for Official Statistics: IT Challenges”, Expert Week, Corso di Laurea in Data Science, Università di Milano Bicocca, Febbraio 2018
- Seminario invitato “Linking Data: From Record Linkage to Object Matching””, Università di Bologna, Prof. Daniela Cocchi, Settembre 2017.
- Tavola Rotonda su “Amministrazione digitale e valutazione, quale sinergie” convegno Innovazione e Società. Metodi Statistici per la Valutazione (IES2017), Napoli, Settembre 2017.
- Seminario invitato “Object matching: la nuova frontiera del record linkage per i Big Data”, Università di Bologna, Prof. Daniela Cocchi, Settembre 2016.
- Seminario invitato “The Use of Internet as Data Source for Official Statistics: the steps so far and the way forward”, Riunione Scientifica SIEDS (Società Italiana di Economia, Demografia e Statistica), Roma, Maggio 2016.
- Seminario invitato, “On the use of Internet as Data Source for Official Statistics: the steps so far and the way forward”, SIEDS (Società Italiana di Economia, Demografia e Statistica), 53esima Riunione Scientifica, Roma, Maggio 2016.
- Seminario invitato “Record Linkage Solutions based on Evolutionary Computation”, Statistics Netherlands, Dr. Barteld Braaksma, Novembre 2015
- Seminario invitato “Big Data Linkage Nuove sfide per il record linkage”, Università di Bologna, Prof. Daniela Cocchi, Ottobre 2015
- Università di Bologna, Prof. Claudio Sartori, (Maggio 2012).
- Seminario invitato “CORE and GSIM” in Web Conference nell’ambito dello GSIM Sprint, Slovenia, (Febbraio 2012).
- Seminario invitato dal titolo “Record Linkage and Data Integration: Can You Have One Without the Other?”, Workshop sulla Qualità dei Dati nella Pubblica Amministrazione, SAS Italy (Aprile 2007).
- Keynote Speech dal titolo “The Object Identification Problem: New Trends in Record Linkage”, 2nd French Workshop on Data and Knowledge Quality, ENIC, Villeneuve d'Ascq, France (Gennaio 2006).

- Seminario invitato dal titolo “Dealing with Data Quality in Distributed and Heterogeneous Information Systems”, CS Colloquium, Purdue University, West Lafayette, Indiana, USA (Novembre 2005).
- Seminario invitato dal titolo “Data Integration and Trust in Data Quality Frameworks”, Università di Milano Bicocca, Italia (Giugno 2005).
- Seminario invitato dal titolo “Dealing with Data Inconsistencies in Cooperative Information Systems” all'IRISA, Rennes, Francia (Giugno 2004).
- Seminario invitato dal titolo “eGovernment e Governo Economici Locali”, Università di Bergamo, Italia (Maggio 2004).
- International Conference and Research Center for Computer Science Dagstuhl Schloss - Seminar on “Data Quality on the Web”, Dagstuhl, Germania (Settembre 2003).
- Panel su “Data Quality & Data Integrity Research Issues” alla 6th IFIP TC-11 WG 11.5 Working Conference on Integrity and Internal Control in Information Systems, EPFL, Losanna, Svizzera (Novembre 2003).
- E' autore, insieme a Tiziana Catarci e Carlo Batini, del seguente tutorial: “Data Quality in Cooperative Information Systems”, presentato alla International Conference on Entity Relationships (ER 2004), Shanghai, Cina, 2004.

Capacità e competenze personali

Madrelingua(e) **Italiano**

Altra(e) lingua(e)

Autovalutazione

Livello europeo

Lingua Inglese

Comprensione

Ascolto

Lettura

C2

C2

Parlato

Interazione
orale

Produzione
orale

C2

C2

Scritto

C2

Competenze Informatiche

Pacchetti di Office Automation: Microsoft Office, Open Office

Sistemi Operativi: Microsoft Windows, Linux

Tecnologie e linguaggi per il Web: OWL, RDF, XML, HTML

Strumenti di editing di testi: LaTeX

Linguaggi di Programmazione: Java e framework correlati, R, C, Prolog, LISP.

Strumenti e linguaggi di Progettazione: ArchiMate, Entity-Relationship (ER), Microsoft Visio, Visual Paradigm, ERWin

Gestione dei dati: SQL, MySQL, SQLite, ORACLE, SQL Server, ORACLE Spatial and Graph

Middleware, application framework e Web server: MS Business Intelligence, Web service (SOAP, REST), ActiveMQ, Tomcat

Altre Capacità Ho una spiccata **attitudine all'innovazione**, testimoniata da:

- Più di 100 pubblicazioni in ambito nazionale ed internazionale con un **h-index** pari a **27** e più di **3600 citazioni** (fonte Google Scholar, Ottobre 2019).
- Partecipazione a **25 progetti di ricerca e innovazione di livello nazionale ed internazionale**.
- **Principal investigator di 9** dei 25 progetti con ruolo di redazione delle proposte e coordinamento delle attività principali: Implementing ModernStats Standards, ESSnet Big Data I e II, CORE, CORA, Progetto “Metodi per migliorare l'accuratezza e la consistenza dei dati basati su algoritmi evolutivi”, “Web Object Matching”, ESTEEM, DaQuinCIS.
- Visibilità da parte della comunità di ricerca internazionale, testimoniata dalle numerose *chairship* e partecipazioni a comitati di programma di conferenze internazionali. Si sottolinea la *chairship* della *17th International Conference on Information Quality (ICIQ 2012)*, promossa dal MIT (Massachusetts Institute of Technology) <http://mitiq.mit.edu/ICIQ/2012/> e dell'*International Workshop on Semantic Statistics (semstats.org)* in conjunction con l'International Semantic Web Conference (ISWC). Inoltre, si sottolinea il recente invito come membro del comitato scientifico della *International Scientific Conference on New Techniques and Technologies for Statistics (NTTS)* del 2021.

Ho **capacità di gestione di progetti e di coordinamento di risorse**, come testimoniato dalla responsabilità di progetti e dai ruoli di coordinamento rivestiti:

- Dal novembre 2007 ad oggi ha rivestito in Istat il ruolo di **responsabile di risorse**, prima con un incarico specifico, poi come dirigente di unità operative e da Febbraio 2017 come responsabile di servizio.
- Ruolo di **coordinatore o responsabile** in diversi gruppi di lavoro Istat sia a carattere metodologico sia legati alla produzione statistica. Alcuni gruppi hanno previsto un numero di risorse maggiore di 20 (ad esempio [GDL 30], [GDL12], [GDL23] e [GDL 20]).
- Ho coordinato **progetti di grande impatto per l'Istat**: il data warehouse del XV Censimento generale della popolazione e delle abitazioni, il portale dei Linked Open Data, le prime statistiche sperimentali sui Big Data, le prime realizzazioni del Sistema Integrato dei Registri, l'impianto del primo Laboratorio per l'Innovazione dell'Istat.

Ho caratteristiche di **leadership** sia attitudinali sia legate all'esperienza maturata. Al Report di valutazione della dirigenza Istat basato sulla metodologia di sviluppo manageriale Feedback 360° (del 4/6/2018), ho riportato una valutazione sempre superiore a 4 su tutte le dimensioni (valutate su scala da 1 a 5), con punte di 4.8 su “Comunicazione”, 4.7 su “Motivazione e sviluppo delle persone” e 4.7 su “Competenza: Ottimizzazione: tempi/costi/benefici”.

Ho capacità di **gestione di gruppi inter-istituzionali ed europei** e so rivestire **ruoli di rappresentanza istituzionale**:

- Sono stata **project e technical manager** di diversi **progetti** in ambito **Eurostat** e **UNECE**.
- Ho avuto numerosi **ruoli di rappresentanza** per Istat in **task force e board** in ambito **Eurostat** e **UNECE**.
- Ho rappresentato Istat in **gruppi inter-istituzionali** di collaborazione con **AgID** e **Presidenza del Consiglio dei Ministri**
- Sono stata membro di **commissioni** di concorso per posizioni dirigenziali dell'**Agenzia delle Entrate** (area data science) e del **Comune di Milano** (area integrazione).

Sono dotata di **capacità didattiche**, testimoniate da docenze universitarie continuative a partire dall'anno accademico 2002-2003 e dalle numerose docenze istituzionali Istat (in numero totale di 37). Inoltre la sottoscritta è stata relatore di 15 tesi di laurea.

Ho spiccate **capacità di sistematizzazione e di analisi**, testimoniate dalla pubblicazione di due libri di cui sono autore insieme al prof. Carlo Batini:

- *Data Quality: Concepts, Methodologies and Techniques*. Data-Centric Systems and Applications, Springer 2006, ISBN 978-3-540-33172-8, che ha venduto ad oggi **oltre 1500 copie**.
- *Data and Information Quality - Dimensions, Principles and Techniques*. Data-Centric Systems and Applications, Springer 2016, ISBN 978-3-319-24104-3, pp. 1-449. L'e-book ha avuto più di **40.000 download** di capitoli dal 2016 (fonte Springer Bookmetrix).

Infine si rileva che l'attività di ricerca della sottoscritta ha avuto impatto su due **brevetti industriali**:

- *Methods and Systems for Managing Data Quality*, Patent number: 7676523, Filing date: Apr 20, 2007 by SAP, Issue date: Mar 9, 2010, Application number: 11/785,929, articolo: M. Scannapieco and C. Batini, *Completeness in the Relational Model: A Comprehensive Framework, Proceedings of the International Conference on Information Quality (ICIQ 2004), Boston, MA, USA, 2004*.
- *Method and system for including data quality in data stream*. Patent number: 7676522, Filing date: Apr 20, 2007 by SAP, Issue date: Mar 9, 2010, Application number: 11/785,928, articolo: M. Scannapieco and C. Batini, *Completeness in the Relational Model: A Comprehensive Framework, Proceedings of the International Conference on Information Quality (ICIQ 2004), Boston, MA, USA, 2004*.

Pubblicazioni

Libri Scritti

[B3] C. Batini and M. Scannapieco: *Data and Information Quality - Dimensions, Principles and Techniques*. Data-Centric Systems and Applications, Springer 2016, ISBN 978-3-319-24104-3, pp. 1-449.

[B2] C. Batini e M. Scannapieco, *Qualità dei dati: Concetti, Metodi e Tecniche*, Springer, 2008.

[B1] C. Batini e M. Scannapieco, *Data Quality: Concepts, Methodologies and Techniques . Data-Centric Systems and Applications*, Springer 2006, ISBN 978-3-540-33172-8.

Libri Editi

[Ed6] S. Capadisli, F. Cotton, A. Haller, E. Kalampokis, M. Scannapieco, R. Troncy: Proceedings of the 4th International Workshop on Semantic Statistics, co-located with 15th International Semantic Web Conference (ISWC 2016), Kobe, Japan October 18, 2016. CEUR Workshop Proceedings 1654, CEUR-WS.org 2016

[Ed5] S. Capadisli, F. Cotton, A. Haller, E. Kalampokis, M. Scannapieco, R. Troncy: Proceedings of the 3rd International Workshop on Semantic Statistics co-located with 14th International Semantic Web Conference (ISWC 2015), Bethlehem, Pennsylvania, USA, October 11th, 2015. CEUR Workshop Proceedings 1551, CEUR-WS.org 2015

[Ed4] L. Berti-Equille, I. Comyn-Wattiau, M. Scannapieco: Proceedings of the 17th International Conference on Information Quality, IQ 2012, Paris, France, November 16-17, 2012. MIT 2012

[Ed3] E. Bertino, A. Maurino, M. Scannapieco: Guest Editors' Introduction: Data Quality in the Internet Era. *IEEE Internet Computing* 14(4): 11-13 (2010).

[Ed2] T. Catarci, H. Panetto, M. Scannapieco: Editorial. *Inf. Syst.* 33(7-8): 595-596 (2008).

[Ed1] F. Naumann and M. Scannapieco (eds.), *IQIS 2004, International Workshop on Information Quality in Information Systems*, 18 June 2004, Paris, France (SIGMOD 2004 workshop), ACM, 2004.

Capitoli In Libri

[CH9] P.D. Falorsi, B. Liseo, M. Scannapieco: *Dealing with Privacy Issues in Data Integration Systems*, versione estesa dei proceedings di 'Law via the Internet 2018', *Knowledge of the Law in the Big Data Era*, Series *Frontiers in Artificial Intelligence and Applications*, ISBN 978-1-61499-984-3, IOS Press, 2019.

[CH8] M. Scannapieco, *Data Quality Models*. In L. Liu and M. Tamer Ozsü (Ed.) *Encyclopedia of Database Systems* (2nd edition), Springer, 2018.

[CH7] M. Scannapieco, *Object Matching: New Challenges for Record Linkage*, in L. Floridi, P. Illari (Ed.) *The Philosophy of Information Quality*, Springer, 2014.

[CH6] D. Zardetto, M. Scannapieco: *Optimal Decision Rules for Constrained Record Linkage: an Evolutionary Approach*, versione estesa per i post-conference proceedings di CLADAG 2011, Springer Series "Studies in Classification, Data Analysis, and Knowledge Organization", Springer-Verlag, Heidelberg, Germany, 2012.

[CH5] C. Di Ciccio, M. Mecella, M. Scannapieco, D. Zardetto, T. Catarci: "MailOfMine – Analyzing Mail Messages for Mining Artful Collaborative Processes", versione estesa per i post-conference proceedings di SIMPDA 2011, Springer Series "Lecture Notes in Business Information Processing", Springer-Verlag, Heidelberg, Germany, 2012.

[CH4] M. Scannapieco, *Data Quality Models*. In L. Liu and M. Tamer Ozsü (Ed.) *Encyclopedia of Database Systems*, Springer, 2009.

[CH3] C. Aiello, T. Catarci, P. Ceravolo, E. Damiani, M. Scannapieco, and M. Viviani, *Emergent Semantics in Distributed Knowledge Management*. In L. Jain

(Ed.) Evolution of WEB in an Artificial Intelligence Environment, Springer-Verlag, 2008.

[CH2] D. Milano, M. Scannapieco, and T. Catarci, Quality of XML Data: Measurement and Diffusion in a Peer-to-Peer Architecture. In A. Gunasekaran (Ed.), Global Implications of Modern Enterprise Information Systems: Technologies and Applications, IGI Publishing, 2008.

[CH1] C. Marchetti, M. Mecella, M. Scannapieco, and A. Virgillito, Data Quality in Cooperative Information Systems, Encyclopedia of Data Warehousing and Mining (ISBN 1-59140-557-2) (J. Wang, ed.), Idea Group Reference, June 2005.

Articoli su Riviste Nazionali ed Internazionali e Monografie

[J25] D. Srivastava, M. Scannapieco, T.C. Redman: Ensuring High-Quality Private Data for Responsible Data Science: Vision and Challenges. J. Data and Information Quality 11(1), 2019.

[J24] G. Barcaroli, M.Scannapieco, Integration of ICT Survey data and Internet data from enterprises websites at the Italian National Institute of Statistics, Accepted for Publication, Statistical Journal of the IAOS - Journal of the International Association for Official Statistics, ISSN 1874-7655, 2019.

[J23] P. Ceravolo, A. Azzini, M. Angelini, T. Catarci, P. Cudre-Mauroux, E. Damiani, A. Mazak, M. Van Keulen, M. Jarrar, G. Santucci, M. Scannapieco, K. Sattler, M. Wimmer, R. Wrembel, F. Zaraket, Big Data Semantics, Journal of Data Semantics, <https://doi.org/10.1007/s13740-018-0086-2>, Springer Berlin Heidelberg, 2018.

[J22] G. Barcaroli, M. Scannapieco, D. Summa: On the Use of Internet as a Data Source for Official Statistics: a Strategy for Identifying Enterprises on the Web, RIEDS - Rivista Italiana di Economia, Demografia e Statistica - Italian Review of Economics, Demography and Statistics, Volume 70, issue 4, 2016.

[J21] D. Firmani, Mecella M., M. Scannapieco, Batini C.: On the Meaningfulness of "Big Data Quality", Data Science and Engineering, Volume 1, issue 1, 2016.

[J20] M. Bruno, A. Duma, M. Silipo, M. Scannapieco, G. Vaste: CORE: A concrete implementation of the CSPA architecture, Statistical Journal of the IAOS, vol. 32, no. 4, pp. 591-596, 2016.

[J19] Batini C., Rula A., M. Scannapieco, Viscusi G.: From Data Quality to Big Data Quality. Journal of Database Management, Volume 26, issue 1, 2015.

[J18] G. Barcaroli, Nurra A., Salamone S., M. Scannapieco, Scarnò M., D. Summa, Internet as Data Source in the Istat Survey on ICT in Enterprises, Journal of Austrian Statistics, Volume 44, issue 2, 2015.

[J17] R. Aracri, S. De Francisci, A. Pagano, M. Scannapieco, L. Tosco, L. Valentino, Integrating Statistical Data with the Semantic Web: The Istat Experience, Mondo Digitale, Volume 49, 2014.

[J16] N.Cibella, M. Scannapieco, L.Tosco, T. Tuoto, L.Valentino, Record Linkage with RELAIS: Experiences and Challenges, Estadística Espanola, Volume 54, 2012.

[J15] S. Montanelli, D. Bianchini, C. Aiello, R. Baldoni, C. Bolchini, S. Bonomi, S. Castano, T. Catarci, V. De Antonellis, A. Ferrara, M. Melchiori, E. Quintarelli, M. Scannapieco, F. Schreiber, and L.Tanca, The ESTEEM platform: enabling P2P semantic collaboration through emerging collective knowledge, Journal of Intelligent Information Systems, Volume 26, no. 2, 2011.

- [J14] D. Bianchini, C. Aiello, R. Baldoni, C. Bolchini, S. Bonomi, S. Castano, T. Catarci, V. De Antonellis, A. Ferrara, M. Melchiori, S. Montanelli, E. Quintarelli, M. Scannapieco, F. Schreiber, and L. Tanca, Emergent Semantics and Cooperation in Multi-knowledge Communities: the ESTEEM Approach, *World Wide Web*, Volume 10, issue 1-2, 2010.
- [J13] D. Zardetto, M. Scannapieco: A Novel Suite of Methods for Mixture Based Record Linkage, in “*Rivista di Statistica Ufficiale*”, n. 2-3/2010, Istat, pagg. 31-58, ISSN 1828 1982, 2010.
- [J12] G. Lebanon, M. Scannapieco, Mohamed R. Fouad, E. Bertino: Beyond k-Anonymity: A Decision Theoretic Framework for Assessing Privacy Risk. *Transactions on Data Privacy* 2(3): 153-183, 2009.
- [J11] D. Milano, M. Scannapieco, and T. Catarci, Measuring and Diffusing Data Quality in a Peer-to-Peer Architecture, *International Journal of Enterprise Information Systems* (ISSN 1548-1115), Volume 3, issue 1, 2007.
- [J10] N. Cibella, M. Fortini, M. Scannapieco, R. Spina, L. Tosco, T. Tuoto, RELAIS: an open Source toolkit for record linkage, in “*Rivista di Statistica Ufficiale*”, n. 2-3/2007, Istat, 2007.
- [J9] D. Milano, M. Scannapieco, and T. Catarci, Structure Aware XML Object Identification, *Data Engineering Bulletin - Special Issue on Data Quality*, Volume 29, 2006.
- [J8] D. Milano, M. Scannapieco, and T. Catarci, P2P Data Quality Improvement in the DaQuinCIS System, *Journal of Digital Information Management* (ISSN 0972-7272), Volume 3, issue 3, 2005.
- [J7] M. Scannapieco, P. Missier, and C. Batini, Data Quality At A Glance, *German Database Journal - Datenbank-Spektrum*, Volume 14, 2005.
- [J6] M. Bertoletti, P. Missier, M. Scannapieco, P. Aimetti, and C. Batini, Improving Government-to-Business Relationships through Data Reconciliation and Process Re-engineering, *Advances in Management Information Systems-Information Quality Monograph (AMIS-IQ) Monograph* (R.Y. Wang, E.M. Pierce, S.E. Madnick, and Fisher C.W., eds.), Sharpe, M.E., 2005.
- [J5] M. Scannapieco, B. Pernici, and E.M. Pierce, IP-UML: A Methodology for Quality Improvement based on IP-MAP and UML, *Advances in Management Information Systems-Information Quality Monograph (AMIS-IQ) Monograph* (R.Y. Wang, E.M. Pierce, S.E. Madnick, and Fisher C.W., eds.), Sharpe, M.E., 2005.
- [J4] M. Scannapieco, A. Virgillito, M. Marchetti, M. Mecella, and R. Baldoni, The DaQuinCIS Architecture: a Platform for Exchanging and Improving Data Quality in Cooperative Information Systems, *Information Systems*, Volume 29 issue 7, 2004.
- [J3] M. Mecella, M. Scannapieco, A. Virgillito, R. Baldoni, T. Catarci, and C. Batini, The DaQuinCIS Broker: Querying Data and their Quality in Cooperative Information Systems, *Journal on Data Semantics*, Volume 1, 2003.
- [J2] B. Pernici and M. Scannapieco, Data Quality in Web Information Systems, *Journal on Data Semantics*, Volume 1, 2003.
- [J1] T. Catarci and M. Scannapieco, Data Quality under the Computer Science Perspective, *Archivi & Computer*, Volume 2, 2002.

Conferenze e Workshop Internazionali

- [C75] C. Daraio, R. Bruni, G. Catalano, G. Matteucci, A. Daraio, M. Scannapieco, D. Wagner-Schuster, B. Lepori, European Tertiary Education Register (ETER): Evolution of the Data Quality Approach, 7th Int. Conf. on Scientometrics & Informetrics, 2-5 Sett.2019.
- [C74] M. Scannapieco, N.R. Fazio, Big Data Architectures @ Istat, NTTS – New Techniques and Technologies for Statistics 2019, Brussels, 11-15 March 2019.
- [C73] M. Scannapieco, P. Struijs, G. Stateva, Estimating enterprise characteristics from web data: Achievements and future developments, NTTS – New Techniques and Technologies for Statistics 2019, Brussels, 11-15 March 2019.
- [C72] E. Bernasconi, F. Pugliese, D. Zardetto and M. Scannapieco, Satellite-Net: Automatic extraction of land cover indicators from satellite imagery by deep learning, NTTS – New Techniques and Technologies for Statistics 2019, Brussels, 11-15 March 2019.
- [C71] P.D. Falorsi, B. Liseo, M. Scannapieco: Dealing with Privacy Issues in Data Integration System, Law via the Internet, Florence, Italy, 11-12 Ott. 2018.
- [C70] M. Bruno, G. Ruocco, O. Luzi, M. Scannapieco: Standardization of Business Statistics Processes in Istat, European Conference on Quality in Official Statistics, Q2018, Krakov, Poland, 2018.
- [C69] P.D. Falorsi, R. Radini, M. Scannapieco, L. Tosco, Modernizing Data Integration Systems at Istat, European Conference on Quality in Official Statistics, Q2018, Krakov, Poland, 2018.
- [C68] R. Aracri, A. Bianco, R. Radini, M. Scannapieco, L. Tosco, F. Croce, M. Lenzerini, D. F. Savo: On the Experimental Usage of Ontology-based data access for the Italian Integrated System of Statistical Registers: Quality Issues, European Conference on Quality in Official Statistics, Q2018, Krakov, Poland, 2018.
- [C67] Capezzuoli A., Della Ratta F., Macchia S., Murgia M., Scannapieco M., Zardetto D.: The use of textual sources in Istat: an Overview. Proceedings of the 14th International Conference on Statistical Analysis of Textual Data, Rome, 12-15 June 2018.
- [C66] T. Catarci, M. Scannapieco, M. Console, C. Demetrescu, My (fair) big data, 2017 IEEE International Conference on Big Data, BigData 2017, Boston, MA, USA, December 11-14, 2017
- [C65] G. Barcaroli, M. Scannapieco, D. Summa: Massive Web Scraping of Enterprises Web Sites: Experiences and Solutions, 61st ISI World Statistics Congress, Marrakesh, Marocco, 2017
- [C64] M. Scannapieco: The Role of Ontologies for Official Statistics, 61st ISI World Statistics Congress, Marrakesh, Marocco, 2017
- [C63] R. Aracri, A. Bianco, R. Radini, M. Scannapieco, L. Tosco: Using Ontologies for Official Statistics: The Istat Experience, Current Trends in Web Engineering - {ICWE} 2017 International Workshops Practi-o-Web, Rome, Italy, June 5-8, 2017.
- [C62] R. Aracri, G. Corcione, A. Pagano, P. Pizzo, M. Scannapieco, L. Tosco, L. Valentino: Highlighting the added-value of Statistical Linked Open Data, NTTS, Brussels, 2017.
- [C61] R. Radini, M. Scannapieco, Garofalo, G.: The Italian Integrated System of Statistical Registers: On the Design of an Ontology-based Data Integration Architecture, NTTS, Brussels, 2017.

- [C60] A. Dreyer, G. Duffes, D. Gillman, M. Scannapieco, L. Tosco: An OWL Ontology for the Generic Statistical Information Model (GSIM): Design and Implementation. International Workshop on Semantic Statistics, SemStats 2016, Kobe, Japan 2016.
- [C59] M. Scannapieco, T. Gjaltema: Classification Explorer: Navigational Querying of Statistical Classifications. International Workshop on Semantic Statistics, SemStats 2016, Kobe, Japan 2016.
- [C58] C. Daraio, A. Gentili, M. Scannapieco: Integrating Microdata on Higher Education Institutions (HEIS) with Bibliometric and Contextual Variables: A Data Quality Approach. ISSI 2015.
- [C57] M. Bruno, A. Duma, M. Silipo, M. Scannapieco, G. Vaste: CORE: a concrete implementation of the CSPA architecture, NTTS, Brussels, 2015.
- [C56] M. Bruno, M. Scannapieco, D. Summa, D. Zardetto: On the Development of a CSPA Error Correction Service: Design and Implementation Issues, NTTS, Brussels, 2015
- [C55] R. Aracri, S. De Francisci, A. Pagano, M. Scannapieco: Official Statistics meets the Semantic Web: How SDMX and RDF can live together, NTTS, Brussels, 2015
- [C54] G. Barcaroli, M. Scannapieco, M. Scarnò, D. Summa: Using Internet as a Data Source for Official Statistics: a Comparative Analysis of Web Scraping Technologies, NTTS, Brussels, 2015
- [C53] M. Scannapieco, G. Barcaroli, S. De Francisci: Big Data Analysis: Experiences and Best Practices in Official Statistics, Conference of European Statistician (CES), Rome, 2014
- [C52] G. Lodi, A. Maccioni, M. Scannapieco, M. Scanu and L. Tosco: Publishing Official Classifications in Linked Open Data, International Workshop on Semantic Statistics (Semstats), Riva del Garda, 2014
- [C51] R. Aracri, S. De Francisci, A. Pagano, M. Scannapieco, L. Tosco, L. Valentino: Publishing the 15th Italian Population and Housing Census as Linked Open Data, , International Workshop on Semantic Statistics (Semstats), Riva del Garda, 2014.
- [C50] M. Scannapieco, D. Summa, D. Zardetto: On Implementing CSPA Specifications for Editing and Imputation Services, Conference of European Statisticians, Work Session on Statistical Data Editing, Paris, France, 2014.
- [C49] G. Barcaroli, S. De Francisci, M. Scannapieco, D. Summa, Dealing with Big Data for Official Statistics: IT Issues; MSIS 2014, Dublin, Ireland, 2014.
- [C48] R. Aracri, S. De Francisci, A. Pagano, M. Scannapieco, L. Tosco, L. Valentino, Publishing Census Data as Linked Open Data, LOD 2014, Rome, Italy, 2014.
- [C47] M. Scannapieco, A. Virgillito, D. Zardetto: Placing Big Data in Official Statistics: a Big Challenge?, NTTS 2013, Brussels, Belgium, 2013.
- [C46] M. Bruno, M. Scannapieco, C. Vaccari, G. Vaste, A. Virgillito, D. Zardetto CORE: a Standard Platform for Statistical Production Processes, NTTS 2013, Brussels, Belgium, 2013.
- [C45] D. Zardetto, M. Scannapieco: “Optimal Decision Rules for Constrained Record Linkage: an Evolutionary Approach”, in: “Book of Abstracts – CLADAG 2011, the 8th Scientific Meeting of the Classification and Data Analysis Group of the Italian Statistical Society”, Pavia University Press, series “Editoria Scientifica”,

pagg. 55-56, Pavia, 7-9 Settembre 2011, ISBN 978 88 96764 22 0, 2011.

[C44] D. Zardetto, M. Scannapieco: “Optimal Decision Rules for Constrained Record Linkage: an Evolutionary Approach”, in: “Book of Short Papers – CLADAG 2011, the 8th Scientific Meeting of the Classification and Data Analysis Group of the Italian Statistical Society”, Pavia University Press, series “Editoria Scientifica”, pagg. 1-6, Pavia, 7-9 Settembre 2011, usb pen ISBN 978 88 906639, 2011.

[C43] C. Di Ciccio, M. Mecella, M. Scannapieco, D. Zardetto, T. Catarci: “MailOfMine – Analyzing Mail Messages for Mining Artful Collaborative Processes”, in: “Proceedings of the 1st International Symposium on Data-Driven Process Discovery and Analysis (SIMPDA 2011)”, pagg. 45-59, Campione d’Italia, 29 Giugno - 1 Luglio 2011, ISBN 978 88 903120 2 1, 2011.

[C42] D. Zardetto, L. Valentino, M. Scannapieco: “MAERLIN: New Record Linkage Methods At Work”, in: “Proceedings of the 6th International Conference on New Techniques and Technologies for Statistics (NTTS 2011)”, pagg. 1-10, Brussels, 22-24 Febbraio 2011, DOI 10.2901/Eurostat.C2011.001, 2011

[C41] M. Scannapieco, C. Vaccari. Standardizing European Statistical Processes: CORA and CORE Projects, 5th International Conference on Methodologies, Technologies and Tools enabling e-Government, Camerino, Italy, 2011

[C40] N. Cibella, M. Scannapieco, T. Tuoto: Open source software: a way to enrich local solutions, supporting paper at the International Meeting on the Management of statistical information systems (MSIS 2011), Luxembourg, 2011.

[C39] D. Zardetto, M. Scannapieco, T. Catarci (2010) “Effective Automated Object Matching”, nel volume: “Proceedings of the 26th IEEE International Conference on Data Engineering (ICDE 2010)”, pagg. 757-768, Long Beach, California, USA, 1-6 Marzo 2010, ISBN 978 1 4244 5444 0. Full Research Paper, Acceptance Rate = 69/570 ~ 12%

[C38] M. Di Gioia, D. Beneventano, M. Scannapieco: Multi-Source Object Identification With Constraints. International Conference on Information Quality (ICIQ 2009): 266-267, 2009.

[C37] N. Cibella, G. L. Fernandez, M. Fortini, M. Guigo, F. Hernandez, M. Scannapieco, L. Tosco and T. Tuoto, Sharing Solutions for Record Linkage: the RELAIS Software and the Italian and Spanish Experiences. Proc. of the 31st New Techniques and Technologies for Statistics (NTTS 2009), Brussels, B, 2009.

[C36] N. Cibella, M. Fortini, M. Scannapieco, L. Tosco, T. Tuoto, C. Vaccari and L. Valentino: Open Source nella Pubblica Amministrazione (OSPA 2009), Rome, 2009.

[C35] M. T. Saccoccio, M. Scannapieco, C. Vaccari and A. Virgillito: SMART: a Web Architecture for Long Running Applications. Proc. of Int. Conf. on Methodologies, Technologies and Tools enabling e-Government, (Metteg 2009), Vigo, 2009.

[C34] F. Cusmai , C. Aiello, M. Scannapieco, T. Catarci: EARL: an Evolutionary Algorithm for Record Linkage, Proc. of the 6th International Workshop on Quality in Databases (QDB 2008), in conjunction with VLDB 2008, August 25th, Auckland, New Zealand, 2008.

[C33] F. Cusmai , C. Aiello, M. Scannapieco, T. Catarci: Record Linkage as a Multiobjective Optimization Problem solved by Evolutionary Computation, in Proc of the First International Workshop on Emergent Semantics and cooperation in open systems (ESTEEM 2008), In cooperation with the 2nd International Conference on Distributed Event-Based Systems (DEBS 2008) Rome, Italy, 2008

- [C32] N. Cibella, M. Fortini, M. Scannapieco, T. Tosco, T. Tuoto, Theory and Practice in Developing a Record Linkage Software, in Proc. of the International Workshop on Integration Methodologies, Vienna, Austria, 2008.
- [C31] A. Inan, M. Kantarcioglu, E. Bertino, M. Scannapieco, A Hybrid Approach to Private Record Linkage, Proc. of the 24th International Conference on Data Engineering (ICDE 2008), April 7-12, Cancun, Mexico, 2008.
- [C30] M. Scannapieco, I. Figotin, E. Bertino and A.K. Elmagarmid, Privacy Preserving Schema and Data Matching, Proc. of the 27th ACM SIGMOD Conference (SIGMOD 2007), Pechino, Cina, 2007.
- [C29] G. Barcaroli, M. Scannapieco, C. Vaccari, and D. Zardetto, Migrating a Critical Application Towards Open Source: the Istat Experience, Proc. of the 1st International Conference on Methodologies, Technologies and Tools Enabling eGovernment (METTEG 2007), Camerino, Italy, 2007
- [C28] T. Tuoto, N. Cibella, M. Fortini, M. Scannapieco, and T. Tosco, RELAIS: Don't Get Lost in a Record Linkage Project, Proc. of the Federal Committee on Statistical Methodologies (FCSM 2007) Research Conference, Arlington, VA, USA, 2007.
- [C27] The ESTEEM Team, Emergent Semantics and Cooperation in MultiKnowledge Environments: the ESTEEM Architecture, Proc. of the VLDB International Workshop on Semantic Data and Service Integration (SDSI 2007), Vienna, Austria, 2007.
- [C26] G. Lebanon, M. Scannapieco, M. Fouad, and E. Bertino, Beyond K-Anonymity: Towards a Framework for Assessing Privacy Risk, Proc. International Conference on Privacy in Statistical Databases (PSD 2006), Roma, Italy, 2006.
- [C25] M. Fortini, M. Scannapieco, T. Tosco, and T. Tuoto, Towards an Open Source Toolkit for Building Record Linkage Workflows, Proc. SIGMOD 2006 Workshop on Information Quality in Information Systems (IQIS'06), Chicago, USA, 2006.
- [C24] D. Milano, M. Scannapieco, and T. Catarci, Structure Aware XML Object Identification, Proc. VLDB 2006 Workshop on Clean Databases (CleanDB'06), Seoul, Korea, 2006.
- [C23] V. Katifori, A. Poggi, M. Scannapieco, T. Catarci, and Y. Ioannidis, OntoPIM: How to Rely on a Personal Ontology for Personal Information Management, Proc. ICSW'05 Workshop on The Semantic Desktop - Next Generation Personal Information Management and Collaboration Infrastructure, Edinburgh, Scotland, 2005.
- [C22] D. Milano, M. Scannapieco, and T. Catarci, Using Ontologies for XML Data Cleaning., Proc. OTM Workshop on Inter-organizational Systems and Interoperability of Enterprise Software and Applications (MIOS+INTEROP), Larnaca, Cyprus, 2005.
- [C21] D. Milano, M. Scannapieco, and T. Catarci, Design and Implementation of a P2P Data Quality Broker, Proceedings of the International Conference on Interoperability of Enterprise Software and Applications (INTEROP 2005), Geneva, Switzerland, 2005.
- [C20] H.V. Jagadish, L.V.S. Lakshmanan, M. Scannapieco, D. Srivastava, and N. Wiwatwattana, Colorful XML: One Hierarchy Isn't Enough, Proceedings of the 2004 ACM SIGMOD Conference (SIGMOD 2004), Paris, France, 2004
- [C19] D. Milano, M. Scannapieco, and T. Catarci, Quality-driven Query Processing of XQuery Queries, Proceedings of the International Workshop on Data

and Information Quality (DIQ 2004), Riga, Latvia, 2004.

[C18] M. Scannapieco and C. Batini, Completeness in the Relational Model: A Comprehensive Framework, Proceedings of the International Conference on Information Quality (ICIQ 2004), Boston, MA, USA, 2004

[C17] D. Milano, M. Scannapieco, and T. Catarci, A Peer-to-Peer Service Supporting Data Quality: Design and Implementation Issues (Poster), Proceedings of the International Conference on Semantics of a Networked World (SNW 2004), Paris, France, 2004.

[C16] K. Aberer, P. Cudrè-Mauroux, A.M. Ouksel, T. Catarci, M. Hacid, A. Illarramendi, V. Kashyap, M. Mecella, M. Mena, E.J. Neuhold, O. De Troyer, T. Risse, M. Scannapieco, F. Saltor, L. De Santis, S. Spaccapietra, S. Staab, and R. Studer, Emergent Semantics Principles and Issues, Proceedings of the 9th International Conference on Database Systems for Advances Applications (DASFAA 2004), Jeju Island, Korea, 2004.

[C15] H. Panetto, M. Scannapieco, and M. Zelm, INTEROP NoE: Interoperability Research for Networked Enterprises Applications and Software, Proceedings of the On the Move to Meaningful Internet Systems 2004, Cyprus, Greece, 2004.

[C14] K. Aberer, T. Catarci, P. Cudre-Mauroux, T. S. Dillon, S. Grimm, M. Hacid, A. Illarramendi, M. Jarrar, V. Kashyap, M. Mecella, E. Mena, E. J. Neuhold, A. M. Ouksel, T. Risse, M. Scannapieco, F. Saltor, L. de Santis, S. Spaccapietra, S. Staab, Rl Studer, and Ol De Troyer, Emergent Semantics Systems., Proceedings of the International Conference on Semantics of a Networked World (SNW 2004), Paris, France, 2004.

[C13] L. De Santis, M. Scannapieco, and T. Catarci, Trusting Data Quality in Cooperative Information Systems, Proceedings of the 11th International Conference on Cooperative Information Systems (CoopIS'03), Catania, Italy, 2003.

[C12] C. Marchetti, M. Mecella, M. Scannapieco, and A. Virgillito, Enabling Data Quality Notification in Cooperative Information Systems through a Web-service based Architecture (short paper), Proceedings of the 4th International Conference on Web Information Systems Engineering (WISE'03), Roma, Italy, 2003.

[C11] P. Bertolazzi, L. De Santis, and M. Scannapieco, Automatic Record Matching in Cooperative Information Systems, Proceedings of the ICDT'03 International Workshop on Data Quality in Cooperative Information Systems (DQCIS'03), Siena, Italy, 2003.

[C10] C. Cappiello, C. Francalanci, B. Pernici, P. Plebani, and M. Scannapieco, Data Quality Assurance in Cooperative Information Systems: a Multi-dimension Quality Certificate, Proceedings of the ICDT'03 International Workshop on Data Quality in Cooperative Information Systems (DQCIS'03), Siena, Italy, 2003.

[C9] C. Marchetti, M. Mecella, M. Scannapieco, A. Virgillito, and R. Baldoni, Data Quality Notification in Cooperative Information Systems, Proceedings of the ICDT'03 International Workshop on Data Quality in Cooperative Information Systems (DQCIS'03), Siena, Italy, 2003.

[C8] P.D. Falorsi, S. Pallara, A. Pavone, A. Alessandrini, E. Massella, and M. Scannapieco, Improving the Quality of Toponymic Data in the Italian Public Administration, Proceedings of the ICDT '03 Workshop on Data Quality in Cooperative Information Systems (DQCIS '03), Siena, Italy, 2003.

[C7] M. Bertoletti, P. Missier, M. Scannapieco, P. Aimetti, and C. Batini, The Services To Businesses Project: Improving Government-to-Business Relationships in Italy, Proceedings of the 2nd EGOV Conference (EGOV'03), Prague, Czech Republic, 2003.

[C6] M. Scannapieco, B. Pernici, and E.M. Pierce, IP-UML: Towards a Methodology for Quality Improvement based on the IP-MAP Framework, Proceedings of the 7th International Conference on Information Quality (ICIQ'02), Boston, MA, USA, 2002.

[C5] M. Mecella, M. Scannapieco, A. Virgillito, R. Baldoni, T. Catarci, and C. Batini, Managing Data Quality in Cooperative Information Systems, Proceedings of the 10th International Conference on Cooperative Information Systems (CoopIS 2002), Irvine, CA, 2002.

[C4] P. Aimetti, C. Batini, C. Gagliardi, and M. Scannapieco, The “Services To Enterprises” Project: an Experience of Data Quality Improvement in the Italian Public Administration, Experience Paper at the 7th International Conference on Information Quality (ICIQ'02), Boston, MA, USA, 2002.

[C3] B. Pernici and M. Scannapieco, Data Quality in Web Information Systems, Best rated paper in Proceedings of the 21th International Conference on Conceptual Modeling (ER 2002), Tampere, Finland, 2002.

[C2] M. Scannapieco, V. Mirabella, M. Mecella, and C. Batini, Data Quality in e-Business, Proceedings of the Workshop on Web Services, e-Business, and the Semantic Web: Foundations, Models, Architecture, Engineering and Applications (WES 2002), in conjunction with CAiSE 2002, Toronto, Ontario, Canada, 2002.

[C1] P. Bertolazzi and M. Scannapieco, Introducing Data Quality in a Cooperative Context, Proceedings of the 6th International Conference on Information Quality (ICIQ'01), Boston, MA, USA, 2001.

Conferenze e Workshop Italiani

[IT15] G. Alleva, M. Scannapieco: The Path for Using Big Data Sources in Istat, SIS 2019, Smart Statistics for Smart Applications, Milano, 18-21 Giugno 2019.

[IT14] R. Radini, M. Scannapieco, L. Tosco, D. Zardetto: Intelligenza Artificiale e Statistica Ufficiale: Esperienze in Istat, Convegno Nazionale CINi sull'intelligenza Artificiale (Ital-IA), Roma 18-19 Marzo 2019.

[IT13] F. De Fausti, F. Pugliese, M. Magarò, A. Ghizzoni, D. Zardetto, M. Scannapieco: Studio di Performance della piattaforma NUTANIX su applicazioni di Deep Learning in ISTAT, Poster alla XIII Conferenza Nazionale di Statistica, Roma, 2018

[IT12] R. Aracri, A. Bianco, R. Radini, M. Scannapieco, L. Tosco: Applicazione dell'Approccio di Gestione dei Dati Guidato dalle Ontologie nell'Ambito del Sistema Integrato dei Registri, Poster alla XIII Conferenza Nazionale di Statistica, Roma, 2018

[IT11] P. Struijs, A. Consten, P. Daas, M. Debusschere, M. Ilves, B. Nikic, A. Nowicka, D. Salgado, M. Scannapieco, and N. Swier: The ESSnet Big Data: Experimental Results, SIS 2017

[IT10] S. Salamone, M. Scannapieco, M. Scarnò: Web scraping and web mining: new tools for Official Statistics, SIS 2014

[IT9] D. Zardetto, M. Scannapieco, L. Valentino, T. Catarci: “On Probabilistic Record Linkage: New Methods Compared to the Fellegi-Sunter Approach”, in: “Proceedings of the 19th Italian Symposium on Advanced Database Systems (SEBD 2011)”, pagg. 21-32, Maratea, 26-29 Giugno 2011.

[IT8] M. Di Gioia, M. Scannapieco, D. Beneventano: Object Identification across Multiple Sources. SEBD 2010: 414-425, 2010.

[IT7] N.Cibella, M. Fortini, M.Scannapieco, T. Tuoto: From theory to practice: the software RELAIS as a solution for record linkage. SIS (Società Italiana di Statistica) 2010.

[IT6] N.Cibella, M. Fortini, M. Scannapieco, L. Tosco, T.Tuoto; L.Valentino: RELAIS: Record Linkage At Istat (Poster), Nona Conferenza Nazionale di Statistica, Roma, 2008.

[IT5] D. Milano, M. Scannapieco, and T. Catarci, Data Quality Improvement in the DaQuinCIS System, Proceedings of the National Conference Sistemi Evoluti per Basi di Dati (SEBD 2004), S. Margherita di Pula, Italy, 2004.

[IT4] L. De Santis, M. Scannapieco, and T. Catarci, A Trust Model for Tightly Coupled P2P Systems, Proceedings of the National Conference Sistemi Evoluti per Basi di Dati (SEBD 2003), Cetraro (CS), Italy, 2003.

[IT3] M. Mecella, M. Scannapieco, A. Virgillito, R. Baldoni, T. Catarci, and C. Batini, Managing Data Quality in Cooperative Information Systems (Extended Abstract), Proceedings of the National Conference Sistemi Evoluti per Basi di Dati (SEBD 2002), Isola d'Elba, Italy, 2002.

[IT2] B. Pernici and M. Scannapieco, A Model for Data Quality in Web Information Systems, Proceedings of the National Conference Sistemi Evoluti per Basi di Dati (SEBD 2002), Isola d'Elba, Italy, 2002.

[IT1] D. Natale, M. Scannapieco, Angeletti P., T. Catarci, and G. Raiss, Qualità dei Dati e Standard ISO/IEC 9126: Analisi Critica ed Esperienze nella Pubblica Amministrazione Italiana, Proceedings of the National Workshop on Sistemi in Rete nella Pubblica Amministrazione, Roma, Italy, 2001.

Publicazioni in Collane Istat e Eurostat

[IS7] B.Lepori, A. Bonaccorsi, A. Daraio, C. Daraio, H. Gunnes, E. Hovdhaugen, M. Ploder, M. Scannapieco, D. Wagner-Schuster, ETER project. Handbook for Data Collection, Brussels, June 2014.

[IS6] S. Bergamasco, M. Scannapieco, A. Virgillito (a cura di): CAPITOLO 10, I sistemi di diffusione dei dati. I Sistemi Informatici a Supporto delle Operazioni Censuarie, Atti del 15° Censimento della popolazione e delle abitazioni, 2016

[IS5] M. Scannapieco, D. Zardetto and G. Barcaroli, La Calibrazione dei Dati con R: una Sperimentazione nell'Indagine Forze di Lavoro ed un Confronto con GENESEES/SAS, ISTAT (serie Contributi), no.4, 2007.

[IS4] A. Boggia, A. Pavone, L. Pizzicannella, and M. Scannapieco, Tecniche e Metodi Sperimentali per la Misurazione dell'Accuratezza dei Dati Toponomastici, L'Integrazione di Dati di Fonti Diverse: Tecniche e Applicazioni del Record Linkage e Metodi di Stima Basati sull'Uso Congiunto di Fonti Statistiche e Amministrative (P.D. Falorsi, A. Pallara, and A. Russo, eds.), Franco Angeli, 2006.

[IS3] A. Boggia, F. Crescenzi, O. Gargano and M. Scannapieco, Miglioramento della Qualità dei Dati Interno alle Amministrazioni, in Principi Guida per la Qualità dei Dati Toponomastici nella Pubblica Amministrazione (P.D. Falorsi and M. Scannapieco, eds.), cap. 4, ISTAT (serie Contributi, vol. 12 - http://www.istat.it/dati/pubbsci/contributi/Contr_anno2005.htm), 2005.

[IS2] P.D. Falorsi and M. Scannapieco, Miglioramento della Qualità dei Dati basato sulla Cooperazione, in Principi Guida per la Qualità dei Dati Toponomastici nella Pubblica Amministrazione (P.D. Falorsi and M. Scannapieco, eds.), cap. 5, ISTAT (serie Contributi, vol. 12 - http://www.istat.it/dati/pubbsci/contributi/Contr_anno2005.htm), 2005.

[IS1] P.D. Falorsi and M. Scannapieco (eds.), Principi Guida per la Qualità dei Dati Toponomastici nella Pubblica Amministrazione, ISTAT (serie Contributi, vol. 12 - http://www.istat.it/dati/pubbsci/contributi/Contr_anno2005.htm), 2005.

Lavori (note e relazioni tecniche)

[RT 40] P.D. Falorsi, M.Scannapieco: “Big Data: Proposta di Roadmap per il triennio 2019-2021”. Documento interno Istat, approvato dal Comitato di Presidenza, 2019.

[RT 39] G. Bruno and D. Nicoletti (Bank of Italy), M. De Cubellis, F. De Fausti, M. Scannapieco and D. Zardetto (Istat): Experiences in Setting up Private Data Sharing Protocols, Maggio 2018

[RT38] Il Sistema Integrato dei Registri- Architettura Informativa e Metodologica, Marzo 2018

[RT37] Orsini E., Grippo G., Scannapieco M.: The Istat Innovation Lab, Relazione alla seconda riunione plenaria del Big data Committee, Gennaio 2018.

[RT36] M. Scannapieco, D. Summa, A. Virgillito, D. Zardetto: “Istat’s Reference Architecture for Internet as a Data Source for Official Statistics”. Big data committee – Annual Report 2017

[RT35] UNECE Machine Learning in Official Statistics, July 2018, available at <https://statswiki.unece.org/pages/viewpage.action?pageId=156336592>

[RT34] ESSnet Big Data: Deliverable 2.4 Final Report, May 2018, available at: https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/index.php/WP2_Reports_milestones_and_deliverables

[RT33] REPORT FINALE Metodologia e risultati della fase sperimentale della mappatura dei processi ai fini della prevenzione della corruzione, Marzo 2018. Task force costituita con Deliberazione n. 428/PER del 13 giugno 2017 available at: https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/index.php/WP2_Reports_milestones_and_deliverables

[RT 31] M. Bruno, G. Ruocco, M.Scannapieco: Linee Guida Enterprise Architecture dell’Istat, Febbraio 2018

[RT 30] ESSnet Big Data: Deliverable 2.1 Legal aspects related to Web scraping of Enterprise Web Sites, February 2017

[RT 29] Implementing ModernStats Standards – Linked Open metadata Design Guidelines, December 2016.

[RT28] Il Registro Base degli Individui Ottobre 2016

[RT 27] La progettazione del Sistema Integrato dei Registri, Luglio 2016

[RT26] E. Baldacci, M. Scannapieco: Istat’s Roadmap for the Adoption of Big Data for Official Statistics, Aprile 2016.

[RT25] Il Sistema Integrato dei Registri nell’ambito del processo di modernizzazione dell’Istat, Febbraio 2016

[RT24] EMS: Sintesi 2010

[RT23] I.stat: Piano di diffusione file di input 2009

[RT22] CORE Deliverable 3.1, Technical Specification vs1, 2011

[RT21] CORE Deliverable 1.3, Management Final Report, 2011

[RT20] CORE Deliverable 1.2, Management

- [RT32] ESSnet Big Data: Deliverable 2.2 Methodological and IT Issues, July 2017
- [RT19] CORE Deliverable 1.1, Management Preliminary Report, 2011
- [RT18] CORE Proposal, 2010
- [RT17] CORA Deliverable 5.2, Report on Dissemination Activities, 2010
- [RT16] CORA Deliverable 5.1, Final Report, 2010
- [RT15] CORA Deliverable 1.3, Management Final Report, 2010
- [RT14] CORA Deliverable 1.2, Management Intermediate Report, 2010
- [RT13] CORA Deliverable 1.1, Management Preliminary Report, 2010
- [RT12] CORA Proposal, 2009
- [RT11] RELAIS 2.3 –User Guide, preprocessing, 2011
- [RT10] “RELAIS 2.2 – User Guide, 2011.
- [RT9] “RELAIS 2.1 – User Guide, 2010.
- [RT8] “RELAIS 2.0 – User Guide, 2009.
- [RT7] “RELAIS”- Guida Utente All’Utilizzo – Versione 1.0, 2008.
- [RT6] M. Scannapieco, and C. Vaccari, Open Source e Statistica nella Pubblica Amministrazione, Rivista del Sistema Statistico Nazionale (Sistan), 2007.
- [RT5] Deliverable CENEX - Integration of Survey and Administrative Data – WP3 – Software Tools for Integration Methodologies, 2004
- [RT4] M. Scannapieco, DaQuinCIS: Exchanging and Improving Data Quality in Cooperative Information Systems, Ph.D. Thesis in Computer Engineering , Università di Roma “La Sapienza”, Dipartimento di Informatica e Sistemistica
- [RT3] M. Scannapieco, Data Quality in Cooperative Information Systems, Doctoral Poster at the 28th Very Large Databases Conference (VLDB 2002), Hong Kong, 2002.
- [RT2] P. Bertolazzi, M.G. Fugini, M. Mecella, B. Pernici, G. Plebani, and M. Scannapieco, Supporting Trusted Data Exchanges in Cooperative Information Systems, Technical Report 30-2001, Dipartimento di Informatica e Sistemistica, Università di Roma “La Sapienza”, Roma, Italy, 2001.
- [RT1] P. Missier, M. Scannapieco, and C. Batini, Cooperative Architectures: Introducing Data Quality, Technical Report 14-2001, Dipartimento di Informatica e Sistemistica., Università di Roma “La Sapienza”, Roma, Italy, 2001.

Il sottoscritto è a conoscenza che, ai sensi dell'art. 76, del D.P.R. n. 445 del 28 dicembre 2000, le dichiarazioni di cui al presente curriculum sono da considerare come rese a pubblico ufficiale e che le dichiarazioni mendaci, la falsità negli atti e l'uso di atti falsi sono puniti ai sensi del codice penale e delle leggi speciali in materia

Dichiara, altresì, di essere informato che, ai sensi e per gli effetti di cui all'art. 13 del D. Lgs. n. 196/2003, i dati personali raccolti saranno trattati, anche con strumenti informatici, esclusivamente nell'ambito del procedimento per il quale la presente dichiarazione viene resa.

Roma, 02/11/2019

Firma
MONICA SCANNAPIECO

(per esteso)

Allega copia fotostatica leggibile fronte-retro non autenticata di un documento d'identità in corso di validità, debitamente sottoscritta.