

RAPPORTO TECNICO GESTIONALE

La protezione dei dati personali by design e by default nell'elaborazione, gestione e diffusione del questionario per il "CENSIMENTO SUGLI ARCHIVI LETTERARI DIGITALI NATIVI" nel progetto di ricerca ALDiNa – Archivi Letterari Digitali Nativi e impostazione della privacy policy del progetto.

Autrici: Roberta Ottaviani e Roberta Bianca Luzietti

Emissione: settembre 2024

Revisione: 00

Sommario

1. IL PROGETTO	3
2. IL CONTESTO	4
1.1.1. PANORAMA NAZIONALE ED EUROPEO DI RIFERIMENTO	4
1.1.2. L'ISTITUTO	6
1.1.3. ILC, CLARIN ERIC E CLARIN IT	7
1.1.4. IL PNRR: FOCUS SULLE IR	11
1.1.5. IL PROGETTO H2IOSC	11
3. IL QUESTIONARIO	14
4. LA VALUTAZIONE DEL QUESTIONARIO E LA PRIMA COLLABORAZIONE CON H2IOSC (TRA WP2 E WP8)	15
5. GLI INPUT RELATIVI ALLA PROTEZIONE DEI DATI PERSONALI	17
6. CONCLUSIONE	20
7. BIBLIOGRAFIA	21

1. Il progetto

ALDiNa è un progetto di ricerca nato per raccogliere informazioni e garantire una documentazione e valorizzazione adeguata del *nativo digitale d'autore*, per promuovere l'identificazione e la condivisione di metodologie e buone pratiche che garantiscano la preservazione, l'accesso e lo studio del patrimonio, per stimolare la comprensione teorica del problema e il dibattito intellettuale conseguente.

Molte istituzioni conservano documenti digitali (e.g. testi, immagini, audio, video, e-mail, post, etc.) su supporti diversi (magnetici, ottici, ma anche cloud) e/o in differenti formati, spesso proprietari e privi di adeguata documentazione. La mancanza di sensibilità e di cura filologica verso il digitale, la fragilità dei supporti e l'obsolescenza di hardware e di software aumentano il rischio di vuoti di memoria. Casi paradigmatici sono la riformattazione di molti personal computer di autrici e autori, la smagnetizzazione dei floppy disk (o altri tipi di supporto) o l'incompatibilità di vecchi formati di documento con i moderni word processors.

Il primo obiettivo di ALDiNa è quello di realizzare una mappatura dei materiali conservati presso centri di ricerca, fondazioni, università, biblioteche, archivi e musei, ed eventualmente, in futuro, presso case editrici e sedi di riviste, e altre realtà che conservano a vario titolo questa tipologia di materiale. Riteniamo infatti che un primo sondaggio sul territorio sia una condizione essenziale per aprire un dibattito sul tema, ormai urgente e non più rinviabile.

Il progetto partirà dunque da un *questionario* che coinvolgerà un *ampio numero di soggetti conservatori*; il lavoro sarà integrato da consultazioni e verifiche dirette presso archivi e istituzioni. Si realizzerà una base di dati, modellata sulla struttura del questionario e degli archivi di cui si ha già conoscenza. Si andrà quindi a creare un catalogo per la raccolta e la preservazione di metadati relativi a contenuti testuali e multimediali, dispositivi hardware eventualmente conservati (computer, floppy disk, hard disk ecc.) e stato di preservazione e funzionamento¹.

¹ <https://www.aiucd.it/aldina-archivi-letterari-digitali-nativi/> (ultimo accesso 26/11/2024)

Proprio in relazione alla diffusione del questionario e alla gestione dei dati generati dalle risposte è emersa preponderante la questione della gestione dei dati ed in particolare dei dati personali di cui tratta questo Rapporto tecnico-gestionale.

2. Il contesto

1.1.1. Panorama nazionale ed europeo di riferimento

Nel nostro paese la digitalizzazione del patrimonio culturale è iniziata negli anni Novanta del Novecento e, come si può facilmente intuire, non ha seguito un percorso unico in termini di metodologie e di attori coinvolti, ma sia a livello di Istituti Centrali del Ministero della cultura che a livello regionale, ovvero da parte delle istituzioni del territorio, c'è stato uno sforzo significativo che ha prodotto risultati importanti, anche se in un quadro frammentato². Fin da questi primi momenti di attività è emerso in maniera chiara che il patrimonio culturale digitale che, secondo quanto riportato nel Piano Nazionale di digitalizzazione del Patrimonio Culturale, *“rappresenta una grande opportunità per creare un ecosistema della cultura capace di incrementare la domanda potenziale e ampliare l'accessibilità per diversi segmenti di pubblico, raggiungere target generazionali e geografici difficilmente coinvolgibili e tessere nuove relazioni fra i beni culturali e le persone”*³. Non stupisce, quindi, che da alcuni anni l'Istituto centrale per la digitalizzazione del patrimonio culturale, cioè la Digital Library del Ministero della cultura, abbia avviato una strategia di armonizzazione degli strumenti e dei metodi utilizzati per raggiungere una maggiore omogeneità è provare a raggiungere un'interoperabilità dei sistemi informativi a livello non solo nazionale, ma anche europeo (Casellato 2022).

È in questo clima nazionale (e non solo), sempre più attento al digitale, che un gruppo di studiosi appartenenti all'Associazione AIUCD- Associazione di Informatica umanistica e cultura digitale, ha

² Si veda ad esempio il lavoro portato avanti all'interno del progetto Gra.fo in Calamai et al. (2013).

³ Ministero della cultura | Piano nazionale di digitalizzazione del patrimonio culturale 2022-2023 | Versione 1.1

iniziato a lavorare insieme sul tema degli archivi letterari nati digitalmente, con un'attenzione sempre crescente verso la necessità di sensibilizzare le comunità scientifiche di riferimento e di “collaborare su questi temi per individuare e condividere metodologie e buone pratiche di acquisizione, gestione e preservazione, garantendo l'accesso e lo studio di questi materiali anche confrontandosi con esperienze diverse, sia nazionali che internazionali⁴”.

Guardando al contesto europeo, già nel 2014 il Consiglio dell'Unione Europea aveva *incluso il digitale nativo all'interno del concetto di Cultural Heritage: il patrimonio culturale – si legge nel [documento](#) – è costituito dalle risorse ereditate dal passato, in tutte le forme e gli aspetti – materiali, immateriali e digitali (prodotti originariamente in formato digitale e digitalizzati), ivi inclusi i monumenti, i siti, i paesaggi, le competenze, le prassi, le conoscenze e le espressioni della creatività umana, nonché le collezioni conservate e gestite da organismi pubblici e privati quali musei, biblioteche e archivi. Esso ha origine dall'interazione nel tempo fra le persone e i luoghi ed è in costante evoluzione. Dette risorse rivestono grande valore per la società dal punto di vista culturale, ambientale, sociale ed economico e la loro gestione sostenibile rappresenta pertanto una scelta strategica per il 21° secolo⁵*

Seguendo questo orientamento, anche in Italia ha iniziato a diffondersi sempre più tra i ricercatori del settore la consapevolezza della necessità del superamento della *“funzione ancillare del bene digitale come replica o copia dell'originale fisico e afferma la legittimità di un percorso di conoscenza autonomo, peculiare e connotato da originalità. Originalità che non discende dall'oggetto, ma dalla relazione intellettuale da cui il bene digitale prende forma e da cui attinge nuovi significati trasmissibili e non solo “pensabili””*⁶

⁴ <https://aiucd.github.io/aldina/> (ultimo accesso 26/11/2024)

⁵ Conclusioni del Consiglio del 21 maggio 2014 relative al patrimonio culturale come risorsa strategica per un'Europa sostenibile (2014/C 183/08)

⁶ Ministero della cultura | Piano nazionale di digitalizzazione del patrimonio culturale 2022-2023 | Versione 1.1

Da qui l'origine del Piano nazionale, che costituisce la visione strategica con la quale il Ministero della cultura, d'intesa con le Regioni, intende promuovere e organizzare il processo di trasformazione digitale nel quinquennio 2022-2026 nei diversi settori dell'ecosistema culturale, rivolgendosi in prima istanza ai musei, agli archivi, alle biblioteche, alle soprintendenze, agli istituti e ai luoghi della cultura pubblici che conservano, tutelano, gestiscono e/o valorizzano beni culturali; per questo costituisce anche il contesto strategico, intellettuale e tecnico di riferimento per la realizzazione degli obiettivi del Piano Nazionale di Ripresa e Resilienza (PNRR).

1.1.2. L'istituto

L'Istituto di Linguistica Computazionale "Antonio Zampolli" (CNR-ILC) del CNR è uno dei maggiori centri di ricerca italiano nel settore della Linguistica Computazionale (LC). L'Istituto afferisce al Cnr-Dipartimento Scienze Umane e Sociali, Patrimonio Culturale.

La missione di CNR-ILC è contribuire allo sviluppo scientifico e tecnologico di settori strategici della Linguistica Computazionale attraverso:

- *l'integrazione di ricerca di base e ricerca applicata*, con particolare attenzione alle possibili *ricadute sociali, economiche e culturali*, in linea con gli obiettivi di uno sviluppo sostenibile ed equo, come delineato dal Piano Nazionale della Ricerca 2021-2027 e con gli obiettivi dell'UE e i *Sustainable Development Goals* delle Nazioni Unite per ridurre diseguaglianze di ogni tipo e contrastare l'esclusione sociale
- la *collaborazione* con Istituti di ricerca, Università, Enti pubblici, piccole e medie imprese (PMI) e industrie nell'ambito di progetti e accordi di collaborazione scientifica internazionali, europei, nazionali e regionali
- la *formazione* di studenti condotta con attività di docenza presso università italiane e straniere e con il *coinvolgimento* di dottorandi, laureandi e stagisti nelle attività di ricerca
- il *trasferimento tecnologico* di risorse e tecnologie linguistiche sviluppate verso sia PMI che grandi industrie nazionali e multinazionali (come ad esempio Google, LinkedIn, Vodafone, IBM, Microsoft, Meta, Noema Life, Loquendo, Zanichelli, etc.) per valorizzare i risultati delle

attività di ricerca finanziata con fondi pubblici/privati e massimizzare la portata dell'output dei progetti di ricerca

- l'attenzione al *social engagement* e alle attività di promozione e *divulgazione scientifica* con valore educativo e culturale rivolte non solo agli esperti del settore, ma anche alle scuole e al grande pubblico

Le attuali macro-aree di ricerca di CNR-ILC sono:

- Informatica umanistica
- Trattamento Automatico della Lingua e Gestione della Conoscenza
- Risorse Linguistiche, Standard e *Infrastrutture di Ricerca*
- Modelli (Bio-)Computazionali dell'Uso Linguistico

Per quanto riguarda l'interesse e l'attività dell'Istituto nell'ambito delle Infrastrutture di Ricerca, queste sono state entrambe crescenti e hanno portato ad un coinvolgimento sempre maggiore dal punto di vista sia delle risorse umane, che strumentali, con particolare riferimento all'Infrastruttura CLARIN ERIC.

1.1.3. ILC, CLARIN ERIC e CLARIN IT

L'infrastruttura CLARIN⁷ mira ad integrare a livello europeo *risorse linguistiche e strumenti avanzati* di ricerca, di accesso e di analisi dei contenuti digitali, supportando i ricercatori nel campo delle Scienze Umane e Sociali (SSH). CLARIN persegue questo obiettivo attraverso la *creazione di una rete di centri nazionali distribuiti* sul territorio europeo che forniscono depositi di dati linguistici digitali (*repository*), strumenti altamente avanzati e archivi fisici e digitali all'avanguardia che, tramite un *servizio di accesso unificato*, sono messi a disposizione della comunità scientifica del settore per condurre ricerche di alto livello a sostegno dell'innovazione e della competitività.

⁷ <https://www.clarin.eu> (ultimo accesso 26/11/2024)

La partecipazione formale del nostro paese a CLARIN è stata attivata grazie all'avvio della procedura di adesione da parte del Ministero della Università nella prima parte del 2015: la conclusione di questo iter ha portato l'Italia, dal 1° ottobre 2015, a diventare Paese Membro a pieno titolo del Consorzio CLARIN-ERIC. Questa adesione formale è giunta in ritardo rispetto agli altri paesi fondatori, soprattutto se si considera il lungo arco di tempo nel quale il coinvolgimento dei ricercatori italiani alle fasi preparatorie è stato di primo piano. Basta citare il coinvolgimento dell'Istituto di Linguistica Computazionale "Antonio Zampolli" all'interno del progetto europeo FP7-Infrastructures – CLARIN-212230 (2008–2011), uno studio pilota di fattibilità che ha gettato le basi per la definizione di CLARIN e la sua costituzione come ERIC (avvenuta nel 2012). Tale ruolo strategico dell'Italia e del CNR nel processo di creazione di CLARIN, sul piano politico, organizzativo, tecnologico e scientifico, testimonia la centralità del nostro paese e il ruolo precursore dei suoi ricercatori nel panorama internazionale delle Scienze Umane e Sociali.

Come ogni Paese Membro, con il pagamento della tassa di adesione annuale, l'Italia ha acquisito il diritto di beneficiare del supporto da parte degli Uffici centrali di CLARIN ERIC per lo *sviluppo tecnico del paese* e la sua integrazione nella federazione e di una serie di vantaggi che vanno dalla formazione, inserimento, coinvolgimento e partecipazione attiva della comunità nazionale nel consorzio internazionale.

Con l'adesione, il Paese Membro si impegna peraltro a contribuire allo sviluppo della infrastruttura e a finanziare, con politiche mirate, un progetto nazionale di:

- *implementazione del nodo nazionale*
- *creazione di un consorzio nazionale* tramite azioni di coinvolgimento degli utenti e di trasferimento di conoscenze

favorendo così lo sviluppo del settore delle Scienze Umane e Sociali⁸. Una conquista molto importante per il nostro paese, che trova riflesso nel Piano Nazionale Infrastrutture di Ricerca (PNIR), dedicato alla programmazione dello sviluppo delle Infrastrutture di Ricerca (IR) e della loro valorizzazione. Il PNIR, emanato dal MUR come parte integrante del Programma Nazionale per la Ricerca (PNR),

⁸ https://office.clarin.eu/v/CE-2011-0001-Update_of_CLARIN_statutes-2017-EC.pdf Chapter 3.

considera le Infrastrutture di Ricerca (IR) un fattore abilitante e uno strumento per sostenere la ricerca di base e prevede un importante investimento su di esse; all'interno del Piano *viene evidenziata la rilevanza di CLARIN-ERIC che risulta inserita nelle 56 infrastrutture prioritarie.*

Negli anni CLARIN ERIC ha, inoltre, preso parte all'organizzazione di eventi formativi e di incontro tra esperti in varie discipline relativamente al tema di conservazione, fruizione e riuso degli archivi orali⁹ portando anche all'instaurazione di un *working group* permanente¹⁰.

In Italia, l'infrastruttura di ricerca CLARIN-ERIC fa capo al Dipartimento Scienze Umane e Sociali, Patrimonio Culturale (DSU) del CNR. Il consorzio italiano CLARIN-IT¹¹ è coordinato per il CNR dalla Dott.ssa Monica Monachini dell'Istituto di Linguistica Computazionale "Antonio Zampolli" (CNR-ILC) di Pisa. Le attività di tecnologiche di implementazione del nodo nazionale sono gestite presso l'Istituto di Linguistica Computazionale "Antonio Zampolli" (CNR-ILC) di Pisa.

Grazie all'adesione a CLARIN si sono avuti importanti risultati strategici, che hanno avuto un impatto significativo sulla comunità italiana di riferimento. Consideriamo in primis che l'infrastruttura tecnologica italiana coordinata attraverso il **data center** nazionale ILC4CLARIN¹², ospitato nell'Area della Ricerca di Pisa presso il CNR-ILC, fornisce all'intera comunità scientifica italiana un *sistema di accesso e autenticazione federata*, secondo le specifiche centrali, così da consentire visibilità, consultazione e accesso sicuro al patrimonio dei dati linguistici digitali prodotti in Italia, nonché la loro reperibilità e preservazione¹³; fornisce inoltre un set di strumenti linguistici digitali sotto forma di servizi di facile utilizzo; inoltre garantisce la consultazione e l'utilizzo dei dati e degli strumenti messi a disposizione da tutti gli altri consorzi nazionali aderenti all'infrastruttura europea CLARIN. Vista

⁹ <https://www.clarin.eu/content/clarin-oral-history-0> (ultimo accesso 26/11/2024)

¹⁰ <https://oralhistory.eu> (ultimo accesso 26/11/2024)

¹¹ <https://www.clarin-it.it/it> (ultimo accesso 26/11/2024)

¹² <https://ilc4clarin.ilc.cnr.it> (ultimo accesso 26/11/2024)

¹³ L'infrastruttura italiana è costituita da due data center: ILC4CLARIN (centro di Tipo B) con il repository nazionale che offre deposito, conservazione, salvaguardia e integrazione di dati linguistici; il repository sviluppato ERCC (Eurac Research CLARIN Centre) è un Centro CLARIN di Tipo C (metadata centre).

l'importanza degli strumenti e dei servizi offerti, non è difficile immaginare che si sia creata attorno al nodo italiano CLARIN-IT una rete di partner¹⁴ che comprende Università, Centri di Ricerca e Organi periferici come la Soprintendenza Archivistica e Biblioteconomica della Toscana, oltre ad un importante network di produttori e utenti di dati e servizi digitali, il cui scopo è rinforzare il rapporto tra il settore delle tecnologie linguistiche e quello delle Scienze Umane e Sociali.

Anche a livello di *Governance* dell'ERIC, il nostro paese ha un forte coinvolgimento e rappresenta gli indirizzi scientifici della comunità italiana, all'interno della Assemblea Generale e dei diversi comitati strategici e nel Forum dei coordinatori nazionali; partecipa attivamente alla redazione dell'Agenda Strategica annuale per il settore, assicurandosi che le linee di ricerca abbiano un'immediata trasposizione nella rete scientifica nazionale.

Il nostro esser membri attivi all'interno della rete di CLARIN-ERIC ha permesso la partecipazione, dell'Italia e del CNR e dell'ILC, ai partenariati dei principali progetti infrastrutturali Europei: ELEXIS¹⁵ che coniuga il settore delle tecnologie linguistiche, Digital Humanities e la lessicografia computazionale; TRIPLE¹⁶ per la creazione di un ambiente integrato di strumenti innovativi e servizi avanzati per le SSH. SSHOC¹⁷ che mira a integrare le varie infrastrutture delle discipline delle Scienze Sociali in un unico ambiente virtuale per favorire ricerche di alta qualità e realizzare la visione dell'Open Science Cloud per il settore. Il coinvolgimento in questi progetti europei innovativi, assicura l'allineamento delle politiche nazionali per la ricerca nel settore alle strategie definite a livello europeo in fatto di scienza aperta e dati FAIR (da spiegare bene) ed infine garantisce la sostenibilità di CLARIN-IT negli anni futuri.

Da questi progetti, che si estendono nell'arco temporale fino al 2023, sono inoltre maturate le condizioni per consentire la messa a punto della proposta progettuale che ha portato al finanziamento del progetto PNRR H2IOSC, in cui l'Istituto CNR-ILC ha un ruolo di primo piano e persegue, tra i vari obiettivi, il potenziamento e l'espansione del data center nazionale CLARIN in Italia.

¹⁴ <https://www.clarin-it.it/it/content/consorzio> (ultimo accesso 26/11/2024)

¹⁵ <https://elex.is/> (ultimo accesso 26/11/2024)

¹⁶ <https://oep.hypotheses.org/2243> (ultimo accesso 26/11/2024)

¹⁷ <https://www.cessda.eu/About/Projects/Current-projects/SSHOC> (ultimo accesso 26/11/2024)

1.1.4. Il PNRR: focus sulle IR

Il Ministero dell'Università e della Ricerca ha emanato un avviso per la presentazione di proposte progettuali per dare attuazione al Piano Nazionale di Ripresa e Resilienza (PNRR), Missione 4 "Istruzione e Ricerca" - Componente 2 "Dalla ricerca all'impresa", Linea di investimento 3.1, "Fondo per la realizzazione di un sistema integrato di infrastrutture di ricerca e innovazione", finanziato dall'Unione europea – NextGenerationEU, azione di riferimento la 3.1.1 "Creazione di nuove IR o potenziamento di quelle esistenti che concorrono agli obiettivi di Eccellenza Scientifica di Horizon Europe e costituzione di reti" di cui al D.M. 7 ottobre 2021, n. 1141.

Con il D.D. 3264 del 28.12.2021 il MUR mira a finanziare almeno 20 progetti dedicati ad IR aventi priorità alta o media, così come individuate nel PNIR; si tratta di "interventi a regia", con modalità di selezione valutativo-negoziale.

L'avviso ha come caratteristiche fondamentali la conformità rispetto al *principio DNSH* e un orientamento fortemente rivolto al *superamento di divari territoriali, di genere e generazionali*, in perfetta coerenza con l'art. 4 del Regolamento (UE) 2021/241 e in linea con il superamento della prima e più importante criticità evidenziata nel nostro PNIR. Questo, infatti, se da un lato sottolinea il *riconoscimento della qualità delle nostre infrastrutture di ricerca in ambito internazionale* testimoniata, ad esempio, dal buon tasso di successo nelle azioni di Horizon 2020 dedicate alle infrastrutture di ricerca, dall'altro espone chiaramente il *problema dello scarso coordinamento tra le IR italiane* ed anche *l'insufficiente coordinamento dei diversi attori* (fra ministeri potenzialmente coinvolti, fra attori pubblici e privati) e delle politiche, nazionali e regionali.

1.1.5. Il Progetto H2IOSC

In questo contesto prende vita l'idea progettuale alla base dell'Humanities and cultural Heritage Italian Open Science Cloud (H²IOSC)¹⁸ che si inserisce armonicamente nella missione del CNR, legata alla valorizzazione delle infrastrutture di ricerca ad alta priorità nell'agenda nazionale definita dal MUR

¹⁸ <https://www.h2iosc.cnr.it> (ultimo accesso 26/11/2024)

nel PNIR 2021-2027, e rappresenta la prosecuzione della strategia infrastrutturale di lungo periodo elaborata dal DSU per accelerare la transizione digitale, a supporto della ricerca e dell'innovazione in ambito umanistico. H²IOSC trae ispirazione dai progetti cluster HORIZON 2020 ed in particolare dall'esperienza del progetto SSHOC, e mira al potenziamento dei nodi nazionali delle quattro Infrastrutture di Ricerca europee afferenti al DSU, dedicate alle Scienze Umane e al Patrimonio Culturale. H²IOSC promuove l'integrazione, in una infrastruttura federata, dei servizi e delle competenze sviluppate presso DARIAH-IT, CLARIN-IT, E-RIHS.it e OPERAS.it. In questo modo sarà possibile fornire ai ricercatori, alle imprese e ai cittadini, un ambiente (il *marketplace*, com'è denominato nel progetto), multidisciplinare, federato e ad accesso libero, in cui poter cercare, condividere e riutilizzare dati, risorse, strumenti e servizi digitali condivisi, basati su buone pratiche e standard FAIR¹⁹. Lo scopo è quello di favorire la ricerca scientifica, lo sviluppo di nuove soluzioni e, in maniera particolare, la formazione degli utenti.

Per il raggiungimento degli scopi del progetto è necessario partire dalla costituzione di una robusta infrastruttura tecnologica. H²IOSC svilupperà una rete di centri di calcolo attrezzati con tecnologie avanzate per assicurare alte prestazioni. Questi centri sono costruiti in pieno rispetto del principio *Do No Significant Harm* (DNSH) il quale assicura che gli interventi previsti dai progetti PNRR non arrechino danno significativo all'ambiente.

L'Unità Operativa dell'ILC di Pisa, guidata dalla Dott.ssa Monachini, grazie al reclutamento di 13 nuove unità di personale altamente qualificato, ha progettato un nuovo Data Center, acquisito tramite una gara europea, basato su tecnologie iperconvergenti per fornire, da un lato, servizi di deposito e archiviazione dati e, dall'altro, importanti risorse di calcolo per applicazioni nell'ambito dell'Intelligenza Artificiale a supporto dell'analisi computazionale di grandi moli di dati e dello sviluppo di tecnologie avanzate.

Un progetto infrastrutturale ha come focus la rete degli utenti: mettere al centro gli attori principali della rete scientifica italiana, università, centri di ricerca, inclusi archivi, musei e biblioteche, per identificarne le rispettive esigenze e analizzarne le richieste e le aspettative (espresse o non espresse),

¹⁹ <https://www.go-fair.org> (ultimo accesso 26/11/2024)

ma anche attraverso la comprensione delle necessità degli *utenti di domani*. L'Unità operativa dell'ILC cura il piano di *engagement*²⁰ degli utenti con una prospettiva di medio e lungo periodo, quindi operando scelte in linea con la sostenibilità dell'infrastruttura per un periodo superiore al decennio che seguirà la fine del progetto (come previsto dai requisiti). La ricognizione e la fotografia dell'esistente costituiscono uno degli strumenti fondamentali a garanzia di una attiva partecipazione e coinvolgimento: l'ILC ha quindi lanciato una campagna di rilevamento per studiare il grado di conoscenza dei servizi delle infrastrutture da parte della comunità, il grado di utilizzo, e il livello di aderenza ai principi FAIR e a quelli della scienza aperta in riferimento ai dati sviluppati. Questa attività ha previsto l'ideazione di un'indagine gestita a più livelli tramite l'impegno di un questionario²¹, interviste, focus group e la ricerca semi-automatica da fonti di informazione quali altri repository, pubblicazioni ed atti di convegni (Luzietti et al. 2024a). Il punto di arrivo di questo sforzo di analisi prevede l'elaborazione di un osservatorio per il monitoraggio anche dopo della fine del progetto delle attività di ricerca ed innovazione delle comunità afferenti ad H2IOSC e di/in linea con un piano di *Training* (Pedonese et al. 2024), dedicato alla formazione delle comunità degli utenti, volto a migliorare le competenze interdisciplinari specifiche del settore e aumentare la consapevolezza dei servizi che le infrastrutture disciplinari possono offrire. Uno degli obiettivi collegati è la formazione di nuove figure professionali specializzate, in grado di formare a loro volta le future generazioni in una prospettiva di *formare i formatori*, mettendo a disposizione materiali didattici comuni facilmente riutilizzabili.

H2IOSC, quindi, ha non solo un alto potenziale tecnico-scientifico, ma costituisce una sfida in quanto trasposizione concreta della *vision* del CNR per i prossimi dieci anni dal punto di vista dello sviluppo delle infrastrutture di ricerca.

²⁰ D1.3 – version M18- H2IOSC Training Activities Coordination and Management Procedures: H2IOSC Engagement and Training Plan

²¹ Al momento della redazione di questo rapporto tecnico l'indagine è ancora attiva e compilabile per mezzo del seguente link: <https://survey.cnr.it/index.php?r=survey/index&sid=825638&lang=it> (ultimo accesso 26/11/2024)

3. Il questionario

Come primo passo verso la raccolta e documentazione di archivi digitali nativi d'autore, i membri di Aldina hanno iniziato stilando una prima breve versione di questionario concepito attorno a tre punti/aspetti fondamentali:

1. Identificazione del soggetto conservatore e del materiale nativo digitale
2. Quantità, tipo di supporto e tipologia del contenuto digitale
3. Informazioni su piani di acquisizione, catalogazione e accesso ai fondi

Nel dare vita a questo strumento di indagine i membri del progetto Aldina si sono rivolti ad alcuni dei membri del progetto H2IOSC i quali all'interno dell'attività di Landscaping and Building Communities hanno lavorato alla costruzione e disseminazione del sopra menzionato questionario. Tale collaborazione ha riguardato il confronto su termini di percorsi formativi, servizi offerti di supporto non solo per le questioni tecniche di gestione del questionario e di protezione dei dati personali, ma anche per avere una partenza con la raccolta dei dati fin dall'origine impostata per essere adeguata agli standard di deposito in CLARIN.

Il gruppo di ricerca del progetto ALDINA si è rivolto al nodo nazionale di CLARIN, ed in particolare al gruppo di ricercatori e tecnologi impegnati nel progetto infrastrutturale H2IOSC, per un supporto ed una condivisione di metodologie e strumenti di lavoro utilizzati, per risolvere una serie di problematiche che erano state già affrontate e superate nel corso del primo anno del progetto H2IOSC.

In particolare, il progetto ALDINA, una volta predisposto il draft del **"CENSIMENTO SUGLI ARCHIVI LETTERARI DIGITALI NATIVI"**, ha individuato la necessità di definire:

- Lo strumento idoneo per la somministrazione digitale del questionario
- Le strategie per renderlo maggiormente fruibile e ridurre i tempi di compilazione
- Gli accorgimenti per una maggiore chiarezza espositiva e per evitare errori nella compilazione
- I documenti necessari per una corretta impostazione della protezione dei dati personali nella fase di raccolta dei dati

- Le policy legate alla protezione dei dati personali per i trattamenti successivi (post raccolta dati) necessari per il buon andamento del progetto di ricerca

4. La valutazione del questionario e la prima collaborazione con H2IOSC (tra WP2 e WP8)

Per la validazione tecnica della struttura del questionario la consulenza si è articolata in un incontro in cui attraverso la rilettura congiunta del questionario sono state valutati e proposti miglioramenti in merito a la scelta di tipologia di domande, l'inserimento di suggerimenti per incrementare la comprensibilità delle domande poste, l'utilizzo di esempi e l'individuazione di un gruppo di rispondenti target potenzialmente interessati a partecipare/collaborare.

Per vocazione intrinseca il progetto H2IOSC è orientato all'analisi dei bisogni delle comunità di riferimento e all'azione in risposta a quanto analizzato ed indentificato come necessità; spesso questo si traduce nell'elaborazione di moduli formativi da erogare o altri servizi da offrire attraverso l'azione diretta della compagine di progetto o sfruttando le piattaforme ed i servizi realizzati nel corso del progetto H2IOSC (attualmente in fase di perfezionamento finale).

In questo caso c'è stato un contatto diretto tra la rappresentante del Comitato Organizzatore e del Comitato Scientifico del progetto ALDINA e alcuni ricercatori strutturati dell'Istituto di Linguistica Computazionale "A. Zampolli" del CNR e, questi ultimi, hanno indirizzato la ricercatrice verso due risorse a tempo determinato appositamente reclutate e formate nell'ambito del progetto H2IOSC: la ricercatrice Roberta Bianca Luzietti e la tecnologa Roberta Ottaviani, che si erano occupate del questionario lanciata all'interno di H2IOSC, sia dal punto di vista della concezione teorica del questionario, che da quello pratico dell'*implementazione tramite l'applicativo Lime Survey*, che, infine, delle *questioni legate alla protezione dei dati personali* connesse a questa attività.

Durante i primi contatti c'è stato uno scambio di e-mail finalizzato alla condivisione della documentazione di progetto e alla definizione degli obiettivi della collaborazione, in risposta alle necessità rappresentate dalla coordinatrice del progetto Aldina.

È poi seguita una fase di studio critico e valutazione del questionario posto alla base del censimento, sotto due profili diversi:

- Ottimizzazione del questionario per la sua riproduzione tramite l'applicativo proposto: Lime Survey
- Segnalazione delle tematiche inserenti la gestione dei dati personali e del flusso documentale necessari sia per la fase di raccolta dei dati e che quella successiva, in considerazione delle finalità del progetto di ricerca condivise

La restituzione di un feedback documentale è stata accompagnata da un incontro in cui sono stati analizzati i punti aperti che dovranno essere decisi dal gruppo di ricerca impegnato nel progetto ALDINA e sono state fornite le indicazioni provenienti dall'esperienza H2IOSC.

Le questioni rimandate alla compagine di progetto sono state le seguenti:

- Sistematizzazione del gruppo di lavoro, definizione dei ruoli
- strutturazione dell'idea progettuale, definizione dei *work package* di progetto, del cronoprogramma e del budget (almeno per macrocategorie)
-

Per quanto riguarda, invece, la condivisione delle *best practices* già applicate in H2IOSC, siamo partiti dalla rilettura di ciascuna domanda alla luce dei template disponibili in Lime Survey. Per ciascuna di queste è stata scelta la modalità migliore di presentazione e, soprattutto, è stato valutato caso per caso l'opportunità di lasciare campi a risposta aperta a fronte della maggiore difficoltà ad analizzare e standardizzare i dati.

È stato mostrato il software in modalità di creazione del questionario, per far vedere quali sono le possibilità offerte dallo strumento e la modalità operativa di configurazione del questionario.

I 3 focus principali hanno riguardato i seguenti elementi da utilizzare:

- Incoraggiare l'uso di domande a risposta multipla: rispetto a quelle a risposta aperta, con un duplice obiettivo: da un lato, facilitare risposte rapide e poco impegnative per il rispondente; dall'altro, limitare le categorie di risposte da analizzare durante la fase di elaborazione.
- Utilizzare il campo "esempio": per fornire soluzioni di risposta nelle domande particolari, al fine di semplificare il processo di compilazione.

- Valorizzare il campo "note": ogni volta che sia necessario aggiungere informazioni per chiarire la logica della domanda e contestualizzare le risposte, in modo da delimitare meglio l'ambito del quesito e prevenire equivoci o risposte fuori tema.
- Limitare l'uso dell'opzione "Altro": e, in ogni caso, associarla alla necessità di fornire una specificazione della risposta in un campo di testo a compilazione libera.

Riguardo la scelta di Lime Survey, sono state rappresentate le opportunità importanti offerte dell'applicativo, sia dal punto di vista della configurazione del questionario, che della distribuzione e dell'analisi delle risposte, in maniera compliance rispetto alle regole del GDPR. La facile fruibilità del QR code per la distribuzione e dell'utilizzo degli strumenti di analisi dei risultati è stata subito apprezzata dalla rappresentante del progetto.

5. Gli input relativi alla protezione dei dati personali

Nella parte di confronto relativa alla privacy nel progetto di ricerca, si è voluto riportare l'accento sulla necessità di capire quali sono gli obiettivi scientifici del progetto di ricerca, sulla delimitazione del perimetro dei dati personali che è necessario trattare per raggiungere gli scopi prefissati, sulla corretta gestione degli stessi fin dalla definizione dell'idea progettuale e, soprattutto, sulla proporzionalità ed adeguatezza degli strumenti da adottare rispetto ai rischi che possono presentarsi per i dati personali. L'approccio, quindi, vuole scardinare l'idea di una "burocrazia difensiva" e passare, invece, ad una gestione consapevole e corretta di quanto necessario in base ad una progettazione del sistema di protezione dei dati impostato con consapevolezza fin dall'inizio ed una responsabilizzazione di tutti i soggetti coinvolti.

Come primo passaggio è stato necessario fare una ricognizione di tutti i dati necessari (rispettando il principio di proporzionalità e limitazione) e tracciare un primo draft di *Data Flow* per capire il percorso che gli stessi dovranno fare nel corso del progetto. In questo modo è stato possibile far emergere in maniera chiara anche il quadro delle attività e delle responsabilità per ciascuna fase di progetto.

La seconda azione è stata mostrare l'*informativa* predisposta per il questionario distribuito dal progetto H2IOSC e rileggere insieme il testo per mostrare come poteva essere facilmente adattato al

progetto ALDINA. Un elemento molto importante è la possibilità di inserire l'informativa e la richiesta di consenso direttamente all'interno di Lyme Survey come prima pagina disponibile per chi si collega al questionario.

Ciao, grazie per il tuo interesse a partecipare a questa indagine.

Questo questionario è stato sviluppato dal team del **Work Package 2** del progetto PNRR-IR H2IOSC – Humanities and cultural Heritage Italian Open Science Cloud (<https://www.h2iosc.cnr.it/>) per **nessità di conoscere meglio e coinvolgere le comunità delle scienze umane, sociali e del patrimonio culturale** individuate come di riferimento per le infrastrutture di ricerca CLARIN, DARIAH, E-RHIS e OPERAS.

Il questionario è rivolto a **tutti gli studiosi**: studenti, ricercatori, professori, cultori della materia ed esperti delle discipline sopra indicate.

I **temi** su cui ti chiederemo di dare il tuo personale contributo **saranno** in merito a:

- i) uso e/o creazione di risorse dati, tecnologie, software e servizi nella tua area di ricerca;
- ii) conoscenza delle infrastrutture di ricerca;
- iii) bisogni/desideri di formazione;
- iv) abitudini di pubblicazione.

Per rispondere sono previsti massimo di **10 minuti**. Se vorrai, al termine del questionario, potrai decidere se contribuire ulteriormente a questo lavoro di ricerca rispondendo ad una seconda parte dell'indagine.

Per interrompere la compilazione del questionario e continuare in un secondo momento, basterà cliccare sul tasto "salva" in alto a destra. Per riprendere la compilazione sarà necessario accedere **solamente** dal link ricevuto via mail.

Per qualsiasi necessità o dubbio relativamente alla compilazione del questionario è possibile **contattare**:

Monica Monachini WP2 Leader: monica.monachini@ilc.cnr.it
Valeria Quochi WP2: valeria.quochi@ilc.cnr.it
Roberta Bianca Luzietti: robertabianca.luzietti@ilc.cnr.it
Roberta Ottaviani (privacy policy): roberta.ottaviani@ilc.cnr.it

Informativa resa ai sensi dell'art.13 del Regolamento UE n. 2016/679

Gentile rispondente,

desideriamo informarti che il Regolamento UE n. 679/2016 "Regolamento generale sulla protezione dei dati" prevede la tutela delle persone rispetto al trattamento dei dati personali. Secondo questa normativa, il trattamento sarà improntato ai principi di liceità, correttezza e trasparenza, adeguatezza, pertinenza e limitazione, esattezza e aggiornamento, non eccedenza e responsabilizzazione. I dati personali sono trattati, ai sensi dell'art. 6, comma 1 lettera a) del GDPR, previo Suo consenso.

Pertanto, Le forniamo le seguenti informazioni:

1. I dati comuni da Lei forniti appartengono alle seguenti categorie:
 - a. Dati anagrafici: nome, cognome (se desumibili dall'indirizzo e-mail fornito), anno di nascita
 - b. Dati di contatto: indirizzo e-mail (che verrà usato per ricontattarla solo in caso di suo consenso esplicito per questo ulteriore trattamento)
 - c. Dati relativi alla sua professione: livello di carriera, istituto di appartenenza, settore e sottosettore di ricerca.

I dati verranno trattati per finalità di ricerca scientifica, nell'ambito del progetto di ricerca Humanities and cultural Heritage Italian Open Science Cloud (H2IOSC) www.h2iosc.cnr.it.

Progetto IR0000029 - Humanities and cultural Heritage Italian Open Science Cloud (H2IOSC) - Missione 4, "Istruzione e Ricerca" - Missione 4, "Istruzione e Ricerca" - Componente 2, "Dalla ricerca all'impresa" - Linea di investimento 3.1 del PNRR, Azione 3.1.1 "Creazione di nuove IR o potenziamento di quelle esistenti che concorrono agli obiettivi di Eccellenza Scientifica di Horizon Europe e costituzione di reti" - Area ESFR1 S&CL. Finanziato dall'Unione europea - NextGeneration EU (CUP B63C22000730005).

La ricerca è finalizzata alla creazione di un cluster federato e inclusivo di infrastrutture di ricerca (IR) all'interno del dominio European Strategy Forum on Research Infrastructure (ESFRI) riguardo l'innovazione sociale e culturale per consentire ai ricercatori di varie discipline nei settori delle scienze umane, delle tecnologie linguistiche e dei beni culturali di collaborare alla ricerca intensiva di dati e condurre

Ho letto e compreso il testo dell'informativa

Avanti

Figura 1: Pagina di apertura del questionario sviluppato all'interno del WP2 del progetto H2IOSC

Lo schema di informativa, disponibile in allegato (Allegato 1), è strutturato in alcuni punti fondamentali:

1. Introduzione e rappresentazione del riferimento normativo principale, il Regolamento UE n. 679/2016 "Regolamento generale sulla protezione dei dati", e sintetico riferimento ai principi cui si ispira il trattamento dei dati personali

2. Specificazione delle categorie di dati raccolti (quando possibile è bene fornire un elenco completo ed esaustivo)
3. Indicazione della finalità del trattamento. Nel caso di ricerca scientifica legata ad un progetto specifico, fornire tutte le indicazioni che consentono di identificarlo e una descrizione sintetica e di facile comprensione del progetto stesso
4. La modalità con cui avviene il trattamento (manuale, automatizzata, mista), le garanzie offerte (es. pseudonimizzazione, cifratura, etc.) e il periodo di conservazione dei dati
5. La revocabilità del consenso espresso, tempi e modi per ritirarlo, condizioni particolari
6. Eventuale comunicazione di dati a terzi (definiti o indefiniti) – se già previsto è opportuno citare eventuali accordi di trasferimento dati già da questo momento iniziale
7. Titolare del trattamento con tutti i dati di contatto
8. DPO (dove previsto) e Responsabili Interni del Trattamento – se già previsto, indicare anche gli eventuali contitolari o i soggetti che dovranno esser nominati Responsabili esterni del trattamento
9. Persone autorizzate al trattamento
10. Modalità di esercizio dei diritti dell'interessato
11. Riferimenti dell'Autorità Garante per il trattamento dei dati personali e facoltà di rivolgersi a quest'ultimo per qualsiasi reclamo

In aggiunta al contenuto minimo del documento, sono stati forniti ulteriori spunti:

1. in armonia con la linea guida 5/2020 sul consenso, adottata dal Comitato Europeo per la protezione dei dati personali ai sensi del regolamento (UE) 2016/679, è stato suggerito di richiedere alle persone selezionate un consenso ulteriore (e separato rispetto a quello specifico necessario per il progetto ALDINA) richiedendo la disponibilità ad esser ricontattati in caso di ulteriori trattamenti necessari per il progetto in oggetto oppure per progetti di ricerca ulteriori aventi finalità simili a questo
2. dotarsi di un Data Management Plan fin dal principio, concependolo come un living document oggetto di aggiornamento periodico, che partendo dal Data Flow vada a chiarire meglio gli aspetti fondamentali per la gestione dei dati del progetto. In questo caso si è ritenuto utile

suggerire l'adozione del modello europeo che è di semplice compilazione e molto intuitivo nei contenuti, gestendo il versionamento per garantire il corretto aggiornamento del documento da parte delle persone autorizzate (modifica/verifica/validazione/nuova emissione)

3. valutare la necessità di produrre documenti ulteriori per definire le relazioni con i partner di progetto ed eventuali terze parti esterne (come ad esempio Accordi di Contitolarità, Data Sharing Agreement, etc.
4. valutare fin dal principio le politiche di sicurezza minime da garantire, lo storage dei dati nel corso del progetto, il trasferimento dei dati, l'eventuale uso di Repository certificati per la conservazione di lungo periodo dei dataset e di tutti gli altri prodotti della ricerca

6. Conclusione

Il progetto H2IOSC ha creato una rete di lavoro molto importante in cui le risorse impegnate nel progetto hanno avuto molte occasioni di incontro e confronto con la comunità scientifica di riferimento, con le associazioni e con gli stakeholder. Da questo, oltre che dal programma di *outreach* e alle iniziative di *engagement* e *dissemination*, è nato un grande interesse nei confronti del progetto, che ha avuto modo di aprirsi alla realtà circostante. L'attività di analisi delle esigenze e dei bisogni formativi unita all'elaborazione di una strategia di communities engagement e training (concepito puntando l'accento sull'approccio *train-the-trainers*) ha iniziato a portare molti frutti fin dal secondo anno della vita del progetto.

La collaborazione con il progetto ALDINA (destinata a continuare in futuro) ha creato l'occasione per condividere le buone pratiche messe a punto nella gestione del questionario di H2IOSC e ha valorizzato le competenze acquisite dal personale reclutato sul progetto, che si può proporre come partner per fornire supporto e formazione per colleghi che lavorano su altri progetti di ricerca.

Questo è un esempio di gestione virtuosa di risorse pubbliche ed aumenta l'impatto di entrambi i progetti.

7. Bibliografia

Calamai, S., Bertinetto, P. M., Bertini, C., Biliotti, F., Ricci, I., & Scuotri, G. (2013). Architecture, methods and purpose of the Gra. fo sound archive. In *2013 Digital Heritage International Congress (DigitalHeritage)* (Vol. 2, pp. 439-439). IEEE.

Casellato, A. (2022). Vademecum per il trattamento delle fonti orali: i punti qualificanti. Definizioni e presupposti legali. *IL DE MARTINO*, 33, 76-79.

Luzietti, R. B., Spadi, A., Giampietro, N., Mancuso, G., Caravale, A., D'Eredità, A., Caradonna, M., Moscati, P., Quochi, V., Monachini, M., Degl'Innocenti, E. (2024a). Digital Humanities and Heritage Science: moving from Landscaping to a dynamic research Observatory in an Open Science Cloud. In *Me.Te. Digitali. Mediterraneo in rete tra testi e contesti, Proceedings del XIII Convegno Annuale AIUCD2024*. A cura di: Di Silvestro, Antonio ; Spampinato, Daria Catania: AIUCD, p. 595. ISBN 978-88-942535-8-0. DOI [10.6092/unibo/amsacta/7927](https://doi.org/10.6092/unibo/amsacta/7927). In: Quaderni di Umanistica Digitale

Luzietti, R. B., Del Gratta, R., Quochi, V., Ottaviani, R., Carpita, D., Monachini, M. (2024b) CLARIN in the Italian Open Science Cloud: Landscaping and Community Engagement. CLARIN Annual Conference, Barcelona, Spain.

Pedonese, G., Frontini, F., Ottaviani, R., Boschetti, F., Spadi, A., Francalanci, L., Scognamiglio, A., Restaneo, P., Chaban, A., Striova, J., Benassi, L. (2024). Materiali didattici come oggetti digitali FAIR: una metodologia condivisa per la formazione in H2IOSC. In *Me.Te. Digitali. Mediterraneo in rete tra testi e contesti, Proceedings del XIII Convegno Annuale AIUCD2024*. A cura di: Di Silvestro, Antonio ; Spampinato, Daria Catania: AIUCD, p. 595. ISBN 978-88-942535-8-0. DOI [10.6092/unibo/amsacta/7927](https://doi.org/10.6092/unibo/amsacta/7927). In: Quaderni di Umanistica Digitale

Allegato 1

Informativa resa ai sensi dell'art.13 del Regolamento UE n. 2016/679

Gentile rispondente,

desideriamo informarla che il Regolamento UE n. 679/2016 "Regolamento generale sulla protezione dei dati" prevede la tutela delle persone rispetto al trattamento dei dati personali. Secondo questa normativa, il trattamento sarà improntato ai principi di liceità, correttezza e trasparenza, adeguatezza, pertinenza e limitazione, esattezza e aggiornamento, non eccedenza e responsabilizzazione. I dati personali sono trattati, ai sensi dell'art. 6, comma 1 lettera a) del GDPR, previo Suo consenso.

Pertanto, Le forniamo le seguenti informazioni:

1. I dati comuni da Lei forniti appartengono alle seguenti categorie:
 - a. Dati anagrafici: nome, cognome (se desumibili dall'indirizzo e-mail fornito), anno di nascita
 - b. Dati di contatto: indirizzo e-mail (che verrà usato per ricontattarla solo in caso di suo consenso esplicito per questo ulteriore trattamento)
 - c. Dati relativi alla sua professione: livello di carriera, istituto di appartenenza, settore e sottosettore di ricerca.

I dati verranno trattati per finalità di ricerca scientifica, nell'ambito del progetto di ricerca Humanities and cultural Heritage Italian Open Science Cloud (H2IOSC) www.h2iosc.cnr.it.

Progetto IR0000029- Humanities and cultural Heritage Italian Open Science Cloud (H2IOSC)- Missione 4, "Istruzione e Ricerca" - Missione 4, "Istruzione e Ricerca" - Componente 2, "Dalla ricerca all'impresa" - Linea di investimento 3.1 del PNRR, Azione 3.1.1 "Creazione di nuove IR o potenziamento di quelle esistenti che concorrono agli obiettivi di Eccellenza Scientifica di Horizon Europe e costituzione di reti"- Area ESFRI S&CI. Finanziato dall'Unione europea – NextGeneration EU (CUP B63C22000730005).

La ricerca è finalizzata alla creazione di un cluster federato e inclusivo di infrastrutture di ricerca (IR) all'interno del dominio European Strategy Forum on Research Infrastructure (ESFRI) riguardo l'innovazione sociale e culturale per consentire ai ricercatori di varie discipline nei settori delle scienze umane, delle tecnologie linguistiche e dei beni culturali di collaborare alla ricerca intensiva di dati e condurre analisi complesse.

Il secondo work package (WP2) del progetto è incaricato della conduzione di un'indagine dettagliata del panorama delle tecnologie linguistiche, delle scienze umane e del patrimonio in Italia, prendendo in considerazione i progetti esistenti, le risorse, gli strumenti, le comunità, le pratiche e gli standard in uso non ancora indicizzati/inseriti in alcun repository nazionale.

Il risultato sarà un osservatorio del panorama della comunità di interesse per H2IOSC in Italia, che comprenda i progetti esistenti, le risorse, gli strumenti, le comunità, le migliori pratiche, gli standard utilizzati e le esigenze da inserire all'interno di un marketplace a disposizione della comunità scientifica e non solo.

Per questo vi chiediamo di compilare il questionario che abbiamo predisposto (suddiviso in un modulo a cui si aggiunge un secondo modulo facoltativo) nell'ambito del progetto di ricerca. Anche in caso di compilazione parziale i dati forniti verranno inseriti nella ricerca.

2. Il trattamento dei dati personali sarà effettuato con modalità automatizzata. Il titolare del trattamento metterà in atto misure tecniche e organizzative adeguate, quali la cifratura e la pseudonimizzazione, volte ad attuare in modo efficace i principi di protezione dei dati e a tutelare i diritti degli interessati in conformità con gli artt. 25 e 32 del GDPR. Al momento del coinvolgimento a ciascun partecipante è assegnato un codice, per impedire l'identificazione dei partecipanti alla ricerca, i documenti/file contenenti i codici identificativi saranno conservati separatamente dai documenti/file che contengono le altre informazioni acquisite. I dati saranno conservati per il tempo necessario alle finalità e cancellati dopo 10 anni dalla conclusione programmata della ricerca.

3. La sua partecipazione a questo progetto è libera e volontaria. Il conferimento dei dati è obbligatorio e l'eventuale rifiuto di fornire tali dati ha come conseguenza la non partecipazione al progetto di ricerca. Lei potrà inoltre decidere di cambiare la propria opinione dopo aver prestato il consenso, purché lei esprima questa volontà prima che i suoi dati vengano completamente anonimizzati, in quanto a seguito di questo processo non sarà più possibile risalire all'origine dei dati.

4. I dati raccolti tramite questionario non saranno comunicati a soggetti terzi definiti, ma potranno essere diffusi esclusivamente in forma anonima e aggregata.

5. Il Titolare del trattamento è Consiglio Nazionale delle Ricerche nella sua articolazione organizzativa Istituto di Linguistica Computazionale "Antonio Zampolli", Area della Ricerca del CNR di Pisa, Via Giuseppe Moruzzi, 1- 56124 Pisa (PI), rappresentato dalla direttrice f.f. Simonetta Montemagni.

6. Il Delegato al trattamento è Simonetta Montemagni .

7. Il Responsabile scientifico della ricerca è Monica Monachini Istituto di Linguistica Computazionale (ILC)- Area della Ricerca di Pisa.

8. Le Persone autorizzate al trattamento dei dati sono: Emiliano Degl'Innocenti CNR-OVI ricercatore (Project leader; WP1 leader), Monica Monachini CNR-ILC dirigente di ricerca (WP2 leader), Valeria Quochi CNR-ILC ricercatrice (WP2 task leader), Roberta Bianca Luzietti ricercatrice CNR-ILC (WP2), Marta Caradonna CNR-ILIESI (WP2), Paola Moscati CNR-ISPC (WP2), Alessandra Caravale CNR-ISPC (WP2), Alessia Spadi CNR-OVI (WP2), Alessandra Chirivì CNR-ISPC (WP2), Francesca Frontini CNR-ILC (WP8), Roberta Ottaviani CNR-ILC (WP8), Daniele Carpita CNR-ILC (WP1) .

9. Presso il Titolare del trattamento è presente il Responsabile della protezione dei dati, nominato ai sensi dell'art. 37 del Regolamento UE 2016/679. Il responsabile della protezione dei dati può essere contattato ai seguenti indirizzi: e-mail: rpdcnr@cnr.it pec: rpdcnr@pec.cnr.it

10. In ogni momento potrà esercitare i Suoi diritti di cui alla sezione 2, 3 e 4 del capo III del Regolamento UE n. 679/2016 (es. diritti di informazione e accesso, di rettifica e cancellazione, di limitazione e di opposizione al trattamento, di portabilità dei dati personali).

11. Lei ha il diritto di presentare un reclamo all'autorità di controllo (in Italia l'Autorità Garante per la protezione dei dati personali, www.garanteprivacy.it, e-mail: protocollo@gpdp.it pec: protocollo@pec.gpdp.it, centralino tel. 06696771).



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



Consiglio Nazionale
delle Ricerche