



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Linee guida per il deposito di dataset e software in AMS Acta

Documento redatto da **AlmaDL**

AlmaDL – Settore Gestione e sviluppo della biblioteca digitale d'Ateneo
Area Patrimonio Culturale (ARPAC)
Alma Mater Studiorum – Università di Bologna

Versione 1, dicembre 2025

Sommario

INTRODUZIONE	1
VALUTAZIONI PRELIMINARI AL DEPOSITO IN AMS ACTA	1
SCELTA DEL TIPO DI VALORIZZAZIONE DEI RISULTATI DELLA RICERCA	1
TITOLARITÀ DEI DIRITTI PER IL DEPOSITO	1
DATI PERSONALI E SENSIBILI	2
SCELTA DEL LIVELLO DI ACCESSO E DELLE LICENZE D'USO DEI CONTRIBUTI	2
LA PROCEDURA DI DEPOSITO DI CONTRIBUTI IN AMS ACTA	3
ACCESSO AD AMS ACTA E ACCETTAZIONE DELLE CONDIZIONI D'USO DEL SERVIZIO	3
GESTIONE DEI CONTRIBUTI E FLUSSO DI VALIDAZIONE	3
PREPARAZIONE DEI CONTRIBUTI PER IL DEPOSITO IN AMS ACTA	4
METADATI NECESSARI ALL'IDENTIFICAZIONE DEI CONTRIBUTI	4
<i>Titolo</i>	5
<i>Responsabilità autoriali e affiliazioni</i>	5
<i>Contact Person ed e-mail per contatti</i>	6
<i>Data</i>	6
<i>Abstract</i>	7
<i>Licenza d'uso</i>	7
<i>Citazione raccomandata e indicatori persistenti (PID)</i>	8
<i>Finanziamento</i>	8
<i>Risorse collegate</i>	8
SPECIFICHE PER I DATASET E I SOFTWARE	9
<i>Dai dati ai dataset</i>	9
Formati dei file	10
Nomenclatura di file e cartelle e strutturazione del dataset	11
Organizzazione dei dati all'interno dei file	12
Dati & Codice	13
Volume dei file che compongono il dataset	13
"Underlying data" collegati a una pubblicazione	13
Documentazione a supporto da allegare ai dati (README file)	14
<i>Software</i>	14
Quali Contributi è possibile depositare come "software"?	14
Documentazione da depositare insieme al Software	15
CHECKLIST PER IL DEPOSITO DI DATASET	16
CHECKLIST PER IL DEPOSITO DI SOFTWARE	16
MODELLI FACSIMILARI & GUIDE	17
CRONOLOGIA DELLE REVISIONI E APPROVAZIONE	17



Questo lavoro è soggetto a licenza [Creative Commons Attribuzione 4.0 Internazionale \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

Introduzione

[AMS Acta](#) è il repository istituzionale dell'Alma Mater Studiorum – Università di Bologna per la **conservazione, la pubblicazione e la disseminazione dei dati, dei software e dei risultati inediti della ricerca**, in linea con i principi FAIR e di Open Science, nel rispetto degli standard nazionali e internazionali.

Gli Utenti e le tipologie documentali ammessi al servizio di deposito sono definiti nella [Carta del servizio di deposito in AMS Acta dei dati della ricerca, del software e dei risultati della ricerca inediti](#) [Carta del Servizio].

Queste Linee Guida riguardano i dati della ricerca e il software [Contributi] e in particolare illustrano:

- gli aspetti da valutare prima di procedere al deposito
- la procedura di login in AMS Acta e l'accettazione delle Condizioni d'uso del Servizio;
- l'area di lavoro per la gestione dei Contributi e le procedure di validazione;
- i metadati necessari all'identificazione dei Contributi;
- i requisiti specifici dei Contributi per l'ammissione al deposito;
- una checklist di controllo per il deposito;
- esempi e modelli facsimilari (es. README file, guida per l'organizzazione e nomenclatura dei file e delle cartelle)

Valutazioni preliminari al deposito in AMS Acta

Scelta del tipo di valorizzazione dei risultati della ricerca

La disseminazione in accesso aperto deve essere preceduta da una **valutazione del potenziale di valorizzazione economica, commerciale o industriale dei risultati della ricerca**. Qualora emerga la possibilità di tutela o sfruttamento competitivo, la pubblicazione in open access potrà avvenire solo successivamente all'adozione delle misure di protezione opportune, garantendo il corretto bilanciamento tra principi di trasparenza, condivisione della conoscenza e salvaguardia degli interessi strategici e istituzionali dell'Università di Bologna.

⇒ Per maggiori informazioni sulle possibilità di sfruttamento competitivo e le misure di tutela consultare le pagine informative nella intranet dedicate alla [tutela della proprietà intellettuale e valorizzazione dei risultati della ricerca](#) o contattare l'Unità di processo Knowledge Transfer Office - KTO (kto@unibo.it).

Titolarità dei diritti per il deposito

La responsabilità di quello che viene depositato e disseminato attraverso il repository è in capo a **chi effettua il deposito** che **deve assicurarsi di avere sufficienti diritti o di avere acquisito i permessi**

necessari per autorizzare l'Ateneo alla conservazione dei Contributi e alla loro eventuale disseminazione in forma aperta.

La disseminazione dei Contributi **non deve ledere eventuali interessi di terzi**, quali, ad esempio, i diritti di proprietà intellettuale che possono riguardare contenuti soggetti al diritto d'autore, dati confidenziali, segreti commerciali e industriali, invenzioni oggetto di richiesta di brevetto, *background knowledge* di progetto. I Contributi inoltre non debbono violare eventuali diritti alla **tutela di dati personali e sensibili**.

Dati personali e sensibili

AMS Acta non può essere utilizzato per condividere Contributi che contengano informazioni personali e/o sensibili riconducibili ai soggetti interessati senza espressa autorizzazione dei soggetti stessi. Le **informative e i moduli di consenso informato** utilizzati per la raccolta dei dati devono pertanto richiedere espressamente ai soggetti interessati l'autorizzazione alla pubblicazione dei dati almeno in forma anonima e/o aggregata. È, inoltre, necessario **rimuovere dai dati anonimizzati e/o aggregati anche gli elementi identificativi indiretti** prima di effettuare il deposito.

⇒ Per maggiori informazioni sugli obblighi previsti dalla normativa in vigore e per supporto nella predisposizione dei moduli per il consenso informato consultare le pagine informative nell'intranet dedicate alla [privacy nella ricerca](#) e al [trattamento di dati personali per finalità di ricerca scientifica](#) o contattare l'**Unità Professionale "Protezione dei dati personali"** (privacy@unibo.it).

Scelta del livello di accesso e delle licenze d'uso dei Contributi

Una volta identificato quali Contributi possono essere destinati al deposito, occorre individuare per ogni file o cartella compressa il livello di accesso e la licenza d'uso più idonei, tenuto conto di eventuali vincoli con terze parti o di obblighi derivanti dalle policy di Ateneo e dagli enti di finanziamento della ricerca. I livelli di accesso supportati da AMS Acta sono: **accesso libero, accesso libero con embargo, accesso riservato al solo staff amministrativo (AlmaDL)**.

Si raccomanda, quando possibile, di selezionare l'accesso libero affinché i Contributi depositati possano essere immediatamente accessibili a qualsiasi utente terzo secondo i principi dell'Open Science: *as open as possible as closed as necessary*.

In casi circoscritti e motivati **è possibile applicare una restrizione temporanea (embargo) ai file** dei Contributi. Ad esempio, se si deposita un dataset collegato ad un articolo in corso di pubblicazione, è possibile impostare un embargo in attesa dell'uscita ufficiale della pubblicazione.

È consentito caricare dataset ad accesso riservato, e quindi non pubblici, solo se nella documentazione esplicativa allegata al dataset (README file) è presente una dichiarazione in cui vengono definite le condizioni di disponibilità dei Contributi e viene indicata la e-mail di una persona di contatto.

Quando pubblicati in accesso riservato o con embargo, i Contributi sono accessibili solo all'Utente che ha effettuato il deposito e allo staff di AlmaDL. In questi casi gli utenti del repository interessati ai risultati di ricerca possono scrivere alle persone di contatto indicate per i Contributi attraverso un form

online che si visualizza tramite il pulsante “Richiedi copia” presente sulla scheda bibliografica pubblica del Contributo. Le eventuali richieste di accesso devono essere gestite direttamente dalle persone di contatto indicate per i Contributi. **AMS Acta non supporta infatti modalità di accesso ristretto a utenti selezionati.**

Oltre al livello di accesso è necessario **attribuire ad ogni file caricato in AMS Acta una licenza d’uso** che definisca i termini e le condizioni che gli utenti del repository devono rispettare nel riutilizzare i Contributi. Nella scelta delle licenze d’uso è bene tenere in considerazione eventuali restrizioni determinate da accordi con terze parti e gli obblighi previsti dagli enti finanziatori della ricerca. Si veda per maggiori informazioni la sezione dedicata “[Licenza d’uso](#)”.

La procedura di deposito di Contributi in AMS Acta

Accesso ad AMS Acta e Accettazione delle Condizioni d’uso del Servizio

Ciascun Utente, per avviare la procedura di deposito di un Contributo, dovrà effettuare l’accesso con le proprie credenziali di Ateneo dalla [homepage di AMS Acta](#), cliccando sul link “**Login per gli autori**”.

Al primo inserimento di un nuovo Contributo di ricerca è necessario prendere visione e accettare online le [Condizioni d’uso del Servizio](#). L’accettazione ha effetto per tutti i Contributi depositati in AMS Acta, inclusi quelli successivi al primo. **La mancata accettazione non consente il caricamento e la pubblicazione dei Contributi tramite AMS Acta.**

⇒ Per supporto riguardante l’abilitazione al deposito dei Contributi contattare lo staff di AlmaDL scrivendo a almadl@unibo.it.

Gestione dei Contributi e flusso di Validazione

Effettuato il primo login, e accettate le [Condizioni d’uso del Servizio](#), si accede alla propria area di lavoro chiamata “**Gestione dei contributi**”. In questa sezione è possibile attivare una nuova procedura di deposito, cliccando sul bottone “**Inserisci un nuovo contributo di ricerca**”, modificare Contributi già caricati, ma non ancora depositati, visualizzare i propri Contributi già pubblicati.

Ad ogni Contributo corrisponde una scheda bibliografica online [**Scheda**] contenente le informazioni necessarie a descriverlo univocamente [**metadati**] e il link ai file che lo compongono. Ogni Scheda è contraddistinta da un numero progressivo automaticamente assegnato dall’Archivio [**ID**]. La compilazione della Scheda avviene attraverso delle maschere guidate per l’inserimento dei metadati e per il caricamento dei file o delle cartelle compresse che costituiscono il Contributo [**maschera metadati**].

Nell’Area di lavoro è possibile verificare e gestire lo stato delle Schede create. Le Schede possono trovarsi nei seguenti stati:

- “**Area di lavoro dell’utente**”. Il Contributo è in forma di bozza, deve essere ancora formalmente depositato ed è accessibile solo all’Utente.

- **“In corso di revisione”**. Il Contributo è stato depositato dall’Utente ed è in fase di revisione [Validazione] da parte di AlmaDL. In questo stato il Contributo non è più modificabile dall’Utente.
- **“Archivio in linea”**. Il Contributo è stato validato e pubblicato da AlmaDL e la sua Scheda è visibile e accessibile dal web. In questo caso la riga relativa al Contributo verrà evidenziata in verde.
- **“Ritirato”**. Il Contributo validato e pubblicato è stato rimosso.

I Contributi in Area di lavoro possono essere modificati fino a quando non vengono depositati. Una volta depositati, possono essere modificati solo da AlmaDL prima della loro pubblicazione. I Contributi depositati non sono automaticamente pubblicati nell’archivio in linea, ma sono trasferiti in un’area di revisione per il controllo formale [Validazione] da parte di AlmaDL e vengono visualizzati come “In corso di revisione”.

La fase di Validazione è volta a verificare la correttezza e la completezza delle informazioni descrittive inserite sotto la propria responsabilità dall’Utente durante la procedura di deposito e la conformità con la presente Guida. AlmaDL si riserva di richiedere agli Utenti eventuali modifiche e integrazioni ai Contributi ove necessario. Al termine della fase di Validazione, AlmaDL notifica all’Utente l’avvenuta pubblicazione del Contributo depositato in AMS Acta.

I Contributi validati e pubblicati nell’archivio in linea, non sono più modificabili dall’Utente. L’Utente, se necessario, può creare nuove versioni o copie del Contributo depositato accedendo alle rispettive funzioni (**“Nuova versione”** e **“Crea una copia di questo documento”**), attraverso il tasto **“Visualizza il documento”** che è presente per ogni Contributo



e selezionando successivamente **“Modifica”**.

Nell’area di lavoro è presente anche la funzione **“Richiesta di rimozione”**, tuttavia **i Contributi validati e pubblicati sono eliminabili solo alle condizioni previste nella [Carta del Servizio](#).**

Preparazione dei Contributi per il deposito in AMS Acta

Metadati necessari all’identificazione dei Contributi

Ogni Contributo depositato in AMS Acta deve contenere al suo interno, in maniera chiara e inequivocabile, alcune informazioni minime che permettano di identificarlo senza ambiguità. Le **informazioni fondamentali da includere all’interno di ogni Contributo, e da riportare nella maschera metadati di AMS Acta**, sono le seguenti:

- titolo
- responsabilità autoriali e affiliazioni
- Contact Person e e-mail per contatti
- data
- abstract

- licenza d'uso
- identificativo persistente e citazione raccomandata
- indicazione di un eventuale finanziamento, progetto di ricerca, patrocinio o sponsorizzazione
- riferimenti a eventuali risorse collegate al Contributo

Titolo

Il titolo è un elemento descrittivo obbligatorio. La sua formulazione richiama più o meno direttamente l'argomento o il soggetto del Contributo e ne permette l'identificazione.

Il titolo deve essere scritto in linguaggio naturale e deve seguire possibilmente le convenzioni e la punteggiatura della lingua in cui è scritto¹ e della comunità scientifica a cui il Contributo è destinato.

Il titolo deve essere univoco, e non deve essere uguale a quello di eventuali altre risorse collegate al Contributo. Ad esempio, se un dataset è associato ad una pubblicazione scientifica che presenta e discute i dati contenuti nel dataset, il titolo e l'abstract del dataset devono essere diversi da quelli della pubblicazione.

⇒ Nella maschera metadati di AMS Acta bisogna ricordarsi di **non** inserire il punto finale al termine del titolo.

Responsabilità autoriali e affiliazioni

Gli Autori e le Autrici [Autori o Autore] sono le persone o le istituzioni che hanno avuto un ruolo autoriale primario e diretto nella creazione del Contributo e che sono indicati come tali al suo interno.

I Contributor sono persone o istituzioni che hanno avuto un ruolo autoriale secondario nella creazione del Contributo, contribuendo a diverse parti o fasi della sua realizzazione, ad esempio nella raccolta, nella cura, nella produzione e conservazione dei dati. Per una panoramica delle tipologie di Contributor supportate da AMS Acta, consultare la [sezione “contributorType” delle linee guida di Datacite](#).

Per ogni responsabilità autoriale deve essere indicata l'affiliazione. La dicitura delle istituzioni di affiliazione, dove possibile, deve essere quella ufficiale o quella più comune. Per una lista delle diciture standard, consultare il registro delle organizzazioni [ROR](#). È raccomandato indicare anche il codice identificativo [ORCID](#), se l'Autore o il Contributor ne sono in possesso.

Le responsabilità autoriali sono metadati obbligatori e devono essere presenti nei file che compongono il Contributo. Non è possibile inserire nella Scheda dei metadati di AMS Acta riferimenti a responsabilità autoriali non espressamente indicate nel Contributo.

¹ I titoli in italiano sono in caratteri minuscoli, il titolo e il sottotitolo sono separati da un punto. Ad esempio *Opera aperta. Forma e indeterminazione nelle poetiche contemporanee*. In inglese, solitamente i titoli seguono la convenzione “title caps”, secondo cui tutte le parole del titolo iniziano con una lettera maiuscola, tranne gli articoli e le preposizioni; il sottotitolo è separato dal titolo con i due punti. Ad esempio, *Orientalism: Western Conceptions of the Orient*. AMS Acta non fornisce indicazioni sullo stile da adottare, rimandando alla scelta degli autori e delle autrici l'uso delle convenzioni stilistiche più appropriate.

⇒ Nella maschera metadati di AMS Acta è possibile specificare il ruolo del Contributor selezionandolo da una lista controllata, basata sul Datacite Metadata Schema.

È anche opportuno riportare, oltre all'affiliazione, il codice ORCID di ogni figura con responsabilità autoriale che ne sia in possesso.

Per ogni Autore è possibile inoltre indicare l'e-mail: si consiglia di riportare quella istituzionale o, nel caso la persona non abbia un ruolo strutturato in Ateneo, un'e-mail personale che venga mantenuta nel tempo. Le e-mail indicate non vengono esposte nell'interfaccia pubblica di AMS Acta.

Contact Person ed e-mail per contatti

È necessario specificare all'interno del Contributo o nella documentazione esplicativa allegata l'indirizzo e-mail della persona o delle persone da contattare per qualsiasi informazione o aspetto riguardante il Contributo stesso [**Contact Person**].

Se si sta depositando un dataset o un software, indicare uno o più Contact Person nel README file o in altra documentazione allegata.

Per dipendenti dell'Ateneo a tempo indeterminato, si consiglia di utilizzare l'e-mail istituzionale. Per Autori o Contributor affiliati all'Università di Bologna ma non strutturati (es. borsisti di ricerca, dottorandi, etc.) si suggerisce di indicare anche un indirizzo e-mail che venga stabilmente mantenuto nel tempo.

⇒ Nella maschera metadati di AMS Acta occorre specificare le figure di contatto indicate nel README file come Contributor con ruolo di tipo "Contact Person".

Nel campo "e-mail per contatti" è possibile indicare solo una e-mail riferita a uno dei Contact Person. Tale e-mail permette di attivare la funzione "Richiedi una copia" quando l'accesso ai file viene ristretto temporaneamente o a tempo indeterminato. La funzione consente di contattare attraverso un form online l'e-mail indicata come contatto. L'e-mail non viene esposta nel form o nell'interfaccia pubblica di AMS Acta.

Data

La data è un'informazione obbligatoria. È possibile specificare se si tratta di una data di creazione o fine stesura oppure della data di rilascio o pubblicazione.

All'interno del Contributo o nella documentazione allegata è possibile specificare tutti i diversi tipi di data, nella maschera metadati di AMS Acta occorre invece indicare quella che si ritiene più significativa.

La data può essere nei formati YYYY, MM/YYYY, DD/MM/YYYY. Ad esempio "2025", "luglio 2025", o "11 luglio 2025".

⇒ Si tenga presente che **la data indicata nella maschera metadati sarà riportata da AMS Acta nella citazione bibliografica del Contributo.**

Abstract

L'abstract è un **metadato obbligatorio** e rappresenta una sintesi breve e strutturata del Contributo, presentandone in modo chiaro e immediato i contenuti essenziali. Ha la funzione di informare rapidamente il lettore sugli obiettivi, il contesto, la metodologia, i principali risultati e l'impatto della ricerca. Deve essere autonomamente comprensibile, conciso e accurato, facilitando l'indicizzazione e la reperibilità del Contributo nei database e nelle banche dati bibliografiche.

Quando si depositano dati che sono alla base di una pubblicazione (c.d. "underlying data"), si raccomanda di inserire un abstract focalizzato sulla descrizione del contenuto del dataset senza riprodurre quello della pubblicazione.

Licenza d'uso

Ad ogni file o cartella compressa che compongono un Contributo è **obbligatorio attribuire una licenza d'uso** che chiarisca i termini e le condizioni di utilizzazione dei contenuti del Contributo stesso da parte di terzi.

Come ricordato in precedenza, **nella scelta della licenza gli Autori devono tenere conto della natura dei dati o dei contenuti e di eventuali vincoli a cui questi sono soggetti**, ad esempio, diritti di terze parti quali diritti di proprietà intellettuale o privacy, clausole contrattuali sottoscritte con gli enti di finanziamento, accordi di confidenzialità, accordi che obbligano l'adozione di licenze restrittive.

Ad esempio, il programma europeo di finanziamento della ricerca, Horizon Europe, richiede che i dati dei progetti finanziati siano rilasciati con licenze Creative Commons [CC BY](https://creativecommons.org/licenses/by/4.0/) and [CCO](https://creativecommons.org/licenses/by/4.0/) (o equivalenti), a meno che non sussistano vincoli o altre restrizioni particolari che debbono essere debitamente documentati.

Per quanto riguarda i software, invece, eventuali librerie, proprietarie o non proprietarie, utilizzate per lo sviluppo potrebbero porre dei vincoli nella scelta della licenza. È fondamentale accertarsi che non ci siano conflitti con la licenza scelta per il rilascio.

Si raccomanda di utilizzare licenze standard come le Creative Commons supportate anche da AMS Acta.

È necessario **indicare la licenza d'uso scelta all'interno dei file del Contributo** attraverso una dichiarazione. Per esempio, nel README file dei dataset è possibile citare le licenze Creative Commons nel seguente modo: *"this dataset is distributed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, <https://creativecommons.org/licenses/by/4.0/>".*

⇒ Nella maschera metadati bisogna **indicare una licenza per ogni file o cartella compressa caricata in AMS Acta**. AMS Acta permette di scegliere diverse tipologie di licenza, da quelle più aperte, come

le Creative Commons CC0 e CC BY², a quelle più restrittive, che permettono l'uso solo per scopi personali, di studio e didattici.

In caso di deposito di Software, se la licenza che si desidera associare al Contributo non è implementata dal repository e non è quindi riscontrabile nella maschera metadati, contattare AlmaDL per valutare la possibilità di inserimento.

Citazione raccomandata e indicatori persistenti (PID)

È consigliabile fornire all'interno del Contributo un suggerimento per citare il Contributo stesso attraverso uno stile citazionale codificato e utilizzato nell'ambito disciplinare di riferimento³.

È necessario indicare nei file che costituiscono il Contributo almeno un identificatore persistente come il DOI. Se il Contributo che si sta depositando non ha già un DOI attribuito, AMS Acta può associarne uno in automatico nella forma **10.6092/unibo/amsacta/ID**, dove ID è il numero identificativo assegnato al momento della creazione della Scheda.

Finanziamento

Se il Contributo è il risultato di una ricerca finanziata è necessario inserire nei file del Contributo un riferimento al programma di finanziamento e al progetto di ricerca, nel formato richiesto dall'ente di finanziamento, includendo, ove necessario, eventuali emblemi o loghi, declaratorie e il link al sito del progetto. Ad esempio:

"This research was funded by European Union's Horizon Europe Program, grant number 1234567890"

"Project info: MADFORWATER (DevelopMent AnD application of integrated technological and management solutions FOR wasteWATER treatment and efficient reuse in agriculture tailored to the needs of Mediterranean African Countries), funded by European Union, Horizon 2020 Programme. Grant Agreement num. 688320, DOI <https://doi.org/10.3030/688320>; <https://www.madforwater.eu>" (in caso di dataset o software)

⇒ Nella maschera metadati di AMS Acta è possibile catturare o inserire manualmente le informazioni riguardanti il finanziamento della ricerca.

Risorse collegate

È raccomandato inserire nel README file di un dataset o di un software i riferimenti a risorse informative collegate. Citare pubblicazioni scientifiche, preregistrazioni, repository di codice o materiali supplementari completi dei loro identificativi persistenti quali DOI o HANDLE consente infatti di contestualizzare meglio i Contributi facilitando la loro comprensione, riutilizzabilità e corretta citazione. In questo modo si rafforza la trasparenza della ricerca, si favorisce la tracciabilità dei risultati e si

² Per maggiori informazioni sulle licenze Creative Commons consultare <https://creativecommons.org/share-your-work/cclicenses/>.

³ Stili citazionali codificati. Es. citazione dataset in formato APA: Amurri, A., Giachino, E., Misuraca, E., Peroni, S. (2025). *UNIBO IRIS Bibliographic Data Dump, Dated 30 May 2025, Updated on 3 July 2025* [Dataset]. University of Bologna. <https://doi.org/10.6092/unibo/amsacta/8427>

aumenta l'impatto complessivo del lavoro, rendendo più agevole la collaborazione tra comunità scientifiche e interdisciplinari.

⇒ Nella maschera metadati di AMS Acta è possibile inserire gli identificativi persistenti delle risorse collegate specificando la natura della relazione. Per una panoramica delle tipologie di "relazioni" supportate da AMS Acta, consultare la [sezione "relationType" delle linee guida di Datacite](#).

Specifiche per i dataset e i software

La seguente sezione fornisce indicazioni su come strutturare i dataset e i software da depositare.

Dai dati ai dataset

La condivisione dei dati della ricerca, che siano o meno a supporto di articoli scientifici, è considerata un fondamento per la realizzazione dei principi della Scienza Aperta, per garantire trasparenza, riproducibilità e integrità nelle pratiche di ricerca e della comunicazione scientifica. Alcuni enti per il finanziamento della ricerca, come la Commissione Europea o la Wellcome Foundation, impongono ai destinatari dei finanziamenti la condivisione dei dati della ricerca sotto forma di dataset.

Un dataset è una collezione di dati relazionati fra loro, raccolta all'interno di una struttura ben definita.

Un dataset può contenere dati di qualsiasi tipo (grezzi, elaborati, ecc.) o natura (numerici, testuali, immagini...) e i criteri per la loro aggregazione possono essere molteplici: stesso scopo di ricerca; stesse tecniche o metodologie usate per produrli; stessa area geografica di riferimento; stesso argomento o classe concettuale; stesso fenomeno di riferimento; stesso autore o gruppo di autori, ecc.

I dati sono spesso, ma non necessariamente, codificati in formato numerico. Anche bibliografie ragionate, dati relativi ad interviste, collezioni di immagini o di audiovisivi o altri tipi di evidenze possono essere raccolte, depositate e disseminate nella forma di dataset.

I dataset devono includere documentazione a supporto (ad esempio README file, guide, codebook, template di questionari, protocolli sperimentali, ecc.) **ed eventuale codice preparato ad hoc per processare e/o analizzare i dati** (ad esempio, script in Python, R o altri linguaggi di programmazione ideati per accorpate i dati da diverse fonti, "pulire" e/o normalizzare i dati, eseguire analisi statistiche, visualizzare i dati, ecc.).

⇒ Per maggiori indicazioni sulle buone pratiche di gestione dei dati della ricerca consultare le [Linee guida di Ateneo per la gestione dei dati della ricerca](#)⁴; per un supporto in fase di progettazione per la gestione dei dati della ricerca e la redazione del Data Management Plan contattare i [Data Stewards](#) dell'Università di Bologna (aric.datasteward@unibo.it).

⁴ Alma Mater Studiorum - Università di Bologna (2024) Linee guida di Ateneo per la gestione dei dati della ricerca. Bologna: University of Bologna, DOI <https://10.6092/unibo/amsacta/7835>.

Formati dei file

I file che compongono il Contributo devono essere salvati in formati che permettano l'accessibilità, il riuso e l'interoperabilità del dato nel lungo periodo.

Sono quindi da prediligersi **formati basati su standard ben documentati e di uso comune, possibilmente aperti, non crittografati e non compressi**, ad esempio semplici file di testo codificati in formati standard (UTF-8, Windows-1252, ecc.) e file .csv per le tabelle.

⇒ Per una panoramica sui formati più raccomandati per la preservazione consultare ad esempio le pagine del UK Data Service⁵, della Library of Congress⁶, dell'Archivio nazionale olandese dei dati di ricerca (DANS)⁷ e dell'Agenzia per l'Italia Digitale (AgID)⁸.

In caso di impossibilità a usare formati aperti standard, è buona norma usare formati proprietari diffusi e conosciuti nella comunità di riferimento. Ad esempio, se non si possono usare formati aperti quali .odt/.txt o .ods/.csv, possono essere usati i formati .docx e .xlsx, anche se proprietari, in quanto ampiamente diffusi.

Se invece il formato proprietario non è diffuso, occorre **convertire il file in un formato standard aperto o, quantomeno, in un altro formato standard proprietario ma ampiamente diffuso**. Solo in caso di problemi di conversione si può mantenere il formato originale dei dati, allegando la documentazione relativa alle applicazioni necessarie per la lettura e fruizione del file.

I dati prodotti da strumentazioni e conservati in formati proprietari e specifici per lo strumento devono essere convertiti in formati (aperti) standard o in formati proprietari diffusi e conosciuti. Spesso i software delle strumentazioni hanno funzioni di esportazione dei dati in formati multipli ed è opportuno controllare se tra i formati di esportazione supportati sono inclusi quelli idonei alla conservazione del tipo di dato che si sta depositando. In alternativa occorre **indicare nella documentazione allegata se i file in formato proprietario possono essere comunque aperti con software libero o open source**. È necessario **includere nel dataset l'eventuale codice sviluppato ad hoc per la fruizione dei dati**.

È possibile fornire i dati all'interno dello stesso dataset in più formati, ad esempio in formato proprietario (.xlsx), fornendo anche una conversione in un formato aperto standard (.csv).

I dati inoltre devono essere preservati nel formato più idoneo alla loro natura.

Ad esempio, **dati numerici tabellari** non devono essere forniti come tabella in un documento di testo tipo .odt o .docx, ma in formato più idoneo al dato tabellare (.csv, .ods o .xlsx).

⁵ UK Data Service Recommended formats, <https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/recommended-formats/>

⁶ Library of Congress Recommended Formats Statement 2024-2025, <https://www.loc.gov/preservation/resources/rfs/TOC.html>

⁷ DANS file format page, <https://dans.knaw.nl/en/file-formats/>

⁸ AgID. Linee Guida Open Data. Allegato B - Standard di riferimento e formati aperti, <https://docs.italia.it/AgID/documenti-in-consultazione/Ig-opendata-docs/it/bozza/allegato-b-standard-di-riferimento-e-formati-aperti.html#allegato-b-standard-di-riferimento-e-formati-aperti>

Analogamente, per le **bibliografie estese** sono da preferirsi formati strutturati rispetto ad elenchi di citazioni in documenti di testo. Ad esempio, attraverso l'uso di tabelle (.csv, .ods o .xlsx), in cui ad ogni colonna corrisponde un'informazione relativa alla singola pubblicazione citata (es. colonna "autori", colonna "titolo", colonna "editore", ...), oppure in formati specifici per la gestione dei riferimenti bibliografici (es. .bib e .xml).

Nel caso invece di **grafici, istogrammi**, ecc., il dato da preservare non è l'immagine ma i dati numerici presentati graficamente in essi, che devono essere conservati in formato tabellare.

Nomenclatura di file e cartelle e strutturazione del dataset

Si richiede di adottare una **nomenclatura omogenea per i nomi di file e cartelle** in linea con le buone prassi promosse a livello internazionale⁹ e una **struttura coerente per l'organizzazione dei file in cartelle e sottocartelle all'interno del dataset**. Le convenzioni e l'organizzazione adottata devono essere **documentate quanto più possibile nel README file** per facilitare la condivisione e il riuso dei dati.

⇒ Se i dati sono legati a progetti di ricerca che prevedono la formulazione di un **Data Management Plan** di progetto è importante seguire eventuali regole di naming e strutturazione presenti in questo documento.

In generale per i nomi di file e cartelle si consiglia di:

- scegliere nomi brevi¹⁰ ma che contengano elementi significativi relativi al contenuto;
- evitare gli spazi per separare gli elementi che compongono il nome del file e utilizzare, per esempio trattini (-) e/o underscore (_);
- non utilizzare caratteri speciali;
- in caso di numerazione inserire degli zeri all'inizio della sequenza (es. usare "001, 002, ..., 100" invece di "1, 2, ..., 100");
- indicare le date in formato YYYYMMDD o YYMMDD, ecc.

Si possono depositare per ogni dataset più file o cartelle compresse. Per motivi di usabilità dell'archivio, quando i file sono molto numerosi è consigliabile raggrupparli in un numero congruo di cartelle compresse, a condizione che il volume delle cartelle compresse sia in linea con le indicazioni della sezione "[Volume dei file che compongono il dataset](#)".

⁹Alcuni esempi di linee guida promosse a livello internazionale:

UK Data Service (<https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/organising/>);

CESSDA Training (<https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/2.-Organise-Document/File-naming-and-folder-structure>);

Stanford Libraries training material (<https://guides.library.stanford.edu/data-best-practices/name-files>);

Harvard file naming guidelines (<https://datamanagement.hms.harvard.edu/plan-design/file-naming-conventions>)

¹⁰ Si tenga presente che nei principali sistemi operativi la lunghezza massima consentita del percorso di un file (comprensiva della lunghezza del nome file) è di 255-260 caratteri e che più la struttura del dataset è nidificata in cartelle e sottocartelle, più il percorso dei file nelle sottocartelle sarà lungo. Di conseguenza è consigliabile non eccedere nella lunghezza dei nomi file, mantenendosi intorno ai 30-40 caratteri, quando possibile.

Le cartelle compresse possono essere strutturate al loro interno in diversi livelli di sottocartelle (struttura “nidificata”). **Si raccomanda di non adottare una struttura profondamente nidificata** (più di 4-5 livelli di sottocartelle) e di evitare la creazione di cartelle compresse all’interno di altre cartelle compresse.

⇒ Quando si includono i file in cartelle compresse è bene selezionare i singoli file insieme e dare il comando per comprimerli invece di inserirli tutti in una cartella e poi comprimere quella. Altrimenti si crea una cartella intermedia ridondante che aumenta il livello di nidificazione e allunga inutilmente il percorso file.

All’interno di cartelle zippate generate in sistemi operativi Mac spesso sono presenti cartelle e file residuali di sistema di nome “__MACOSX” e “.DS_Store”, che risultano nascoste nei sistemi Mac ma sono visualizzabili nei sistemi Windows. È consigliabile non includere tali cartelle nel materiale depositato. Si vedano a questo proposito come esempio le [istruzioni sul forum Stack Overflow](#). Similmente, è bene non includere nelle cartelle zippate depositate i file nascosti di tipo “Thumbs.db” generati nei sistemi Windows.

AlmaDL fornisce una breve guida per la nomenclatura di file e cartelle e la strutturazione in cartelle e sottocartelle (vedi sezione “[Modelli facsimilari & Guide](#)”).

Organizzazione dei dati all’interno dei file

I dati all’interno dei file devono essere chiaramente esposti e comprensibili.

In caso di dati numerici occorre indicare, direttamente nel file stesso o almeno nel [README file](#), **unità di misura, eventuali sigle e acronimi a cui vengono associati**. In caso i dati siano presentati in fogli di calcolo (es. .ods o .xlsx) si sconsiglia l’uso di **colori o altre formattazioni del testo**. Eventuali raggruppamenti o parametri di selezione devono essere aggiunti in una nuova colonna.

È sconsigliabile depositare fogli di calcolo (.ods, .xlsx) con **formule attive**, perché potrebbero non essere compatibili tra software diversi o tra diverse versioni dello stesso software. In questi casi è necessario presentare nel foglio di calcolo i valori letterali dei dati elaborati con le formule e descrivere dettagliatamente a parte le formule applicate per ottenerli¹¹.

È altresì sconsigliabile depositare fogli di calcolo contenenti **codici macro**^{11, 12} per problemi di compatibilità e sicurezza. In questi casi si consiglia di conservare il codice macro commentato in un file testuale a parte.

Si sconsiglia inoltre di inserire all’interno del foglio di calcolo la **rappresentazione grafica (es. grafici e istogrammi) dei dati**. Se si ritiene sia necessario corredare i dati con questo tipo di informazione visiva per aumentarne la comprensibilità, la fruizione e il riutilizzo, è bene allegarla in un file a parte all’interno della cartella del dataset in formato immagine (es. .png, .jpeg. etc.).

¹¹ Artefactual Systems and the Digital Preservation Coalition (2021) *Preserving Spreadsheets. Data Types Series*. Digital Preservation Coalition. DOI: <http://doi.org/10.7207/twgn21-09>.

¹² Codice macro. Codice scritto generalmente in Visual Basic o Javascript che permette effettuare azioni ripetitive all’interno del foglio di calcolo.

Se i dati sono presentati all'interno di una pubblicazione e sono mostrati in determinate figure e tabelle **non usare i riferimenti dell'articolo es. "figura 1, 2, ..." nel dataset**, ma piuttosto descrivere il tipo di dati e le unità di misura in cui sono espressi. È fondamentale, infatti, che il contenuto del dataset sia comprensibile anche senza prendere visione di un'eventuale pubblicazione associata.

Dati & Codice

Quando i dati sono il risultato dell'elaborazione mediante un codice, è **opportuno che il dataset contenga il codice sorgente degli script utilizzati per l'elaborazione, i dati di input e i dati elaborati**, specialmente se l'elaborazione è difficilmente riproducibile o computazionalmente impegnativa.

Nel caso in cui i dati di input siano già esistenti e liberamente accessibili in repository o altre piattaforme esterne che ne assicurino la conservazione nel tempo, è opportuno citarli all'interno del README file e nel metadato "Risorse Collegate" di AMS Acta, specificando la relazione opportuna.

Volume dei file che compongono il dataset

AMS Acta non impone limiti al numero o alle dimensioni dei file e delle cartelle compresse da caricare per ogni record. **Per quanto riguarda la dimensione del singolo file o della singola cartella compressa, in generale si suggerisce di non superare i 3GB.**

Tuttavia, se il volume complessivo dei dati da depositare è molto consistente, ad esempio è superiore ai 30 GB, **possono essere depositate cartelle compresse di volume maggiore, fino a 10 GB.** In queste situazioni, si richiede di contattare AlmaDL prima del deposito.

Nel caso di cartelle compresse, è buona prassi indicare nella documentazione allegata al dataset l'algoritmo e il livello di compressione utilizzato.

"Underlying data" collegati a una pubblicazione

Se i dati sono associati ad una pubblicazione e presentati in essa (cd. **"underlying data"**), titolo e abstract del dataset e della pubblicazione devono differenziarsi. Dati e pubblicazione, per quanto collegati, sono infatti "oggetti" distinti: l'abstract deve essere focalizzato sulla descrizione del contenuto del dataset e il titolo deve riflettere il contenuto del dataset.

Gli "underlying data" possono avere responsabilità autoriali differenti da quelli della pubblicazione.

Si raccomanda di depositare gli "underlying data" prima dell'uscita ufficiale della pubblicazione a cui si riferiscono per poter inserire la citazione del dataset completa di DOI attribuito da AMS Acta nella pubblicazione stessa.

Inoltre, **si raccomanda di compilare l'apposita sezione "Dati FAIR" presente nella scheda bibliografica in IRIS della pubblicazione di ricerca con il DOI del dataset depositato**, per collegare la pubblicazione descritta in IRIS con i dati di riferimento.

Documentazione a supporto da allegare ai dati (README file)

Ogni dataset depositato in AMS Acta deve essere corredato da **file di documentazione che permettano di comprenderne il contenuto e le metodologie utilizzate per raccogliere e produrre i dati, senza dover chiedere chiarimenti agli autori.**

AlmaDL richiede di allegare al dataset almeno un README file che riporti, oltre ai metadati fondamentali (titolo, responsabilità autoriali, data, licenza, informazioni su eventuali finanziamenti e risorse collegate), anche informazioni su:

- contenuto del dataset (file che compongono il dataset, il loro nome, formato ed eventuale strutturazione in cartelle e sottocartelle);
- descrizione dettagliata del dataset
 - abstract;
 - descrizione del contenuto dei singoli file;
 - istruzioni/informazioni tecniche relative al formato dei file e ai software. Ad esempio, indicazione dei software necessari per l'apertura dei file o per la loro conversione;
 - informazioni sulle fonti dei dati;
 - metodologia con cui i dati sono stati raccolti e generati, se necessario per il loro riutilizzo;
 - significato di sigle e acronimi, se presenti

AlmaDL fornisce un modello di README file da completare e allegare ai dati (vedi sezione "[Modelli facsimilari & Guide](#)").

⇒ **Alcune delle informazioni da inserire nel README file coincidono con quelle da inserire nella maschera metadati in AMS Acta in fase di deposito e devono quindi corrispondere.**

La documentazione a supporto da allegare ai dati non è limitata al README file, ma può includere anche altri tipi di documenti quali, ad esempio, guide, codebook, template di questionari, protocolli sperimentali, script ecc.

Software

Quali Contributi è possibile depositare come "software"?

In AMS Acta è possibile depositare software strutturato sia in forma di **file eseguibile (.exe)**, sia nella forma di **codice sorgente**.

⇒ Se il Contributo da depositare consiste in uno o più script per processare/analizzare dati e necessita di essere depositato insieme ai dati di input e output, si suggerisce di depositare il Contributo come dataset e non come software.

Le seguenti indicazioni per il deposito di software in AMS Acta sono in linea con le buone pratiche riconosciute e adottate a livello internazionale¹³.

Documentazione da depositare insieme al Software

In AMS Acta, il software depositato deve essere corredato almeno da due file fondamentali: README file e LICENSE file.

Nel **README file**, oltre ai metadati necessari all'identificazione dei Contributi (titolo, responsabilità autoriali e affiliazioni, informazioni di contatto, data, abstract, licenza d'uso, informazioni su eventuale finanziamento, citazione raccomandata e indicatori persistenti, risorse collegate), vanno fornite anche informazioni relative a:

- versione del software e eventuali note relative ad essa (es. change log e bugfixes);
- ambiente di run-time (il linguaggio di programmazione e la sua versione, la piattaforma di run-time, i sistemi operativi in cui il software può essere installato, eventuali dipendenze da librerie o altri software);
- eventuale piattaforma su cui è stato sviluppato il software (es. GitHub, GitLab, ecc.);
- link al sito web e alla documentazione del software o del progetto che ha portato al suo sviluppo (se esiste);
- eventuale descrizione delle funzionalità principali del software;
- istruzioni per la compilazione/installazione del software (o un riferimento su dove reperire un file contenente queste informazioni, tipicamente chiamato INSTALL);
- istruzioni per l'uso del software (o un riferimento a file di documentazione per il suo utilizzo).

Il file **LICENSE.txt** deve contenere per esteso i termini della licenza del progetto. Nel caso di licenze software open source, si raccomanda di usare i nomi standard e i testi delle licenze reperibili nella [SPDX License List](#). Ulteriori informazioni relative alle licenze software open source e a come applicarle possono essere reperite sul sito della [Open Source Initiative](#) e sul [license selection tool di GitHub](#).

⇒ Nel caso in cui il software sia stato sviluppato utilizzando **librerie proprietarie**, è fondamentale accertarsi che la licenza a cui si desidera associarlo non sia in conflitto con i termini d'uso a cui sono soggette le suddette librerie.

In linea con le pratiche emergenti di preservazione dei software¹³, che suggeriscono di fornire i metadati citazionali e gli altri metadati che descrivono il Contributo in formato machine readable, è anche possibile codificare la citazione raccomandata per il software in Citation File Format ([CFF](#)) e i metadati descrittivi in formato JSON-LD secondo lo schema [CodeMeta](#). I tool [CFFINIT](#) e [CodeMeta generator](#)

¹³ Si veda come esempio la guida per l'archiviazione di software di ricerca di Software Heritage: Roberto di Cosmo (2020) "How to use Software Heritage for archiving and referencing your source code: guidelines and walkthrough" (2020), <https://annex.softwareheritage.org/public/guidelines/archive-research-software.pdf>

permettono di creare file CITATION.cff e codemeta.json che possono essere depositati insieme al Contributo.

Checklist per il deposito di dataset

- **Dispongo delle autorizzazioni necessarie per depositare questo dataset?**
vedi sezione [Valutazioni preliminari al deposito in AMS Acta](#)
- **Il dataset contiene informazioni personali o informazioni che permettono indirettamente di risalire all'identità delle persone coinvolte?**
vedi sezione [Dati personali e sensibili](#)
- **Ho stabilito un livello di accesso e una licenza d'uso?**
vedi sezione [Scelta del livello di accesso e delle licenze d'uso dei Contributi](#)
- **Il livello di accesso e la licenza sono in linea con le politiche del mio ente finanziatore?**
vedi sezioni [Scelta del livello di accesso e delle licenze d'uso dei Contributi](#) e [Licenza d'uso](#)
- **I file sono in un formato standard, aperto e non proprietario?**
vedi sezione [Formati dei file](#)
- **Nel caso in cui i dati non siano in un formato aperto, ho allegato una conversione degli stessi in un formato aperto? Ho fornito indicazioni sul software da utilizzare per aprire i file?**
vedi sezione [Formati dei file](#)
- **I nomi dei file e delle cartelle non contengono caratteri speciali, seguono una convenzione e sono organizzati in una struttura documentata?**
vedi sezioni [Nomenclatura di file e cartelle e strutturazione del dataset](#) e [Organizzazione dei dati all'interno dei file](#)
- **Se il dataset è a corredo di una pubblicazione, ho attribuito i metadati descrittivi richiesti e sono in grado di inserire la citazione nella pubblicazione?**
vedi sezioni [Citazione raccomandata e indicatori persistenti \(PID\)](#), [Risorse collegate](#) e ["Underlying data" collegati ad una pubblicazione](#)
- **Ho predisposto e allegato almeno un README file di documentazione con i metadati e le informazioni richieste dal repository?**
vedi sezione [Documentazione a supporto da allegare ai dati \(README file\)](#)

Checklist per il deposito di software

- **Il Contributo che devo depositare è un software?**

vedi sezioni [Dati & Codice](#) e [Quali Contributi è possibile depositare come "software"?](#)

- **Dispongo delle autorizzazioni necessarie per depositare questo software?**

vedi sezione [Valutazioni preliminari al deposito in AMS Acta](#)

- **Il livello di accesso e la licenza sono in linea con le politiche del mio ente finanziatore?**

vedi [Scelta del livello di accesso e delle licenze d'uso dei Contributi](#) e [Licenza d'uso](#)

- **Il software è stato sviluppato usando librerie proprietarie?**

vedi [Licenza d'uso](#) e [Documentazione da depositare insieme al Software](#)

- **Ho rispettato i termini d'uso di eventuali moduli e librerie open source che ho utilizzato?**

Vedi sezione [Licenza d'uso](#)

- **Ho predisposto e allegato almeno un README file di documentazione con i metadati e le informazioni richieste dal repository e un file LICENSE, che riporta il testo integrale della licenza?**

vedi sezione [Documentazione da depositare insieme al Software](#)

Modelli facsimilari & Guide

Modelli facsimilari e guide (es. README file, guida per l'organizzazione e nomenclatura dei file e delle cartelle) sono disponibili nella [pagina che descrive queste Linee Guida sul sito del Sistema Bibliotecario di Ateneo](#).

Cronologia delle revisioni e approvazione

20/10/2025	AlmaDL	prima stesura
30/10/2025	Gruppo di Lavoro Open Science (GLOS)	revisione
06/11/2025	Comitato di Indirizzo del Sistema Bibliotecario di Ateneo	approvazione
11/12/2025	AlmaDL	pubblicazione