

## PNRR GRINS -Growing Resilient, Inclusive and Sustainable Spoke

### 3: Households' Sustainability

#### WP2: Database on university students

**Progetto di ricerca:** L'influenza del percorso scolastico, universitario e delle istituzioni sulle condizioni lavorative future

Data inizio prevista: 09/07/2025

Data fine prevista: 31/12/2030

##### 1. Introduzione: breve stato dell'arte e giustificazione teorica

Il progetto di ricerca mira ad esaminare le determinanti delle scelte entro il percorso formativo scolastico e universitario e le conseguenze di queste scelte su specifici aspetti di disuguaglianza nell'accesso ai percorsi di formazione e sulle conseguenze nel mercato del lavoro.

In particolare, interessa:

- a) ricostruire l'influenza del percorso universitario sugli apprendimenti, le condizioni lavorative future e sullo sviluppo locale, esaminando le traiettorie degli studenti;
- b) riconoscere i fattori di rischio per l'abbandono universitario, in modo da poter sviluppare politiche per prevenire e contrastare il *drop-out* dai percorsi universitari;
- c) esaminare la mobilità geografica degli studenti nelle diverse fasi di transizione del percorso scolastico, con particolare attenzione alla transizione da laurea triennale a magistrale;
- d) analizzare le ricadute dei test di ammissione sulla scelta dei percorsi formativi;
- e) descrivere il processo di formazione di valore associato alle traiettorie formative;
- f) documentare la relazione tra tasse universitarie, borse di studio e strumenti di supporto finanziario per la formazione terziaria, gli apprendimenti e le condizioni lavorative future.

L'obiettivo è comprendere i fattori chiave che influenzano alcuni esiti di particolare interesse, riconducibili a due categorie principali: da una parte, gli **esiti universitari**, e dall'altra gli **esiti lavorativi**. Con riferimento ai primi effettueremo una ulteriore differenziazione, analizzando separatamente gli esiti riguardanti i casi di completamento del percorso di studi, quali il voto finale ed i tempi di conclusione (desumibili dai dati amministrativi di Ateneo) e la soddisfazione riguardo al percorso di studi (desumibile da Almalaurea), e quelli di abbandono precoce del percorso di formazione universitaria (o "drop-out"). Per quanto riguarda il drop-out, il progetto vuole identificare, in particolare, quali siano i fattori che aumentano il rischio di abbandono, individuati tra elementi socio-demografici, ambientali, personali, del contesto accademico e del mercato del lavoro locale, per sviluppare un sistema di "early warning" che permetta di identificare e classificare questi fattori di rischio, distinguendo tra: i) rischio legato a

caratteristiche socio-demografiche, familiari, di contesto e di background scolastico all'ingresso; ii) rischio legato alla performance universitaria nel I semestre di frequenza dei corsi universitari. Il progetto mira a studiare i fattori maggiormente predittivi dell'abbandono universitario e l'influenza delle condizioni del mercato del lavoro locale (provinciale e/o regionale) sulle scelte dello studente e delle famiglie.

Allo stesso modo, con riferimento agli esiti sul mercato del lavoro dei laureati, differenzieremo le analisi per condizione occupazionale, in termini di occupato/non occupato/in cerca di occupazione, per tipologia di contratto lavorativo (a tempo determinato/indeterminato, part-time/full-time), per retribuzione annua netta e per soddisfazione relativa alla condizione lavorativa.

Esamineremo le conseguenze delle scelte educative e delle scelte lavorative sul territorio e sul mercato del lavoro locale, con riferimento in particolare alla mobilità studentesca sia in entrata che in uscita, all'attrattività geografica nella transizione scuola-università e laurea triennale – master e con riferimento alle ricadute della formazione terziaria sullo sviluppo locale in termini di capitale umano, innovazione, e allineamento tra domanda e offerta. Verranno esaminate le traiettorie percorse dagli studenti, esaminando i flussi in entrata ed uscita e l'evoluzione nel tempo delle traiettorie prevalenti, con particolare attenzione alle conseguenze della pandemia COVID-19, al ruolo delle caratteristiche socio-demografiche e di background familiare, delle caratteristiche del territorio (quali accessibilità del servizio abitativo, qualità della vita, struttura urbana), delle caratteristiche dell'offerta formativa (in termini di attività curriculari ed extracurriculari) e dell'Ateneo eventualmente dedotte da basi di dati distinte, ed altre eventualmente raccolte all'interno della piattaforma AMELIA sviluppata nell'ambito del progetto GRINS entro l'orizzonte temporale rilevante. Si studierà la relazione tra mobilità geografica e performance scolastica, accademica e lavorativa.

La **creazione di un database che integra i dati amministrativi sulle carriere degli studenti con i dati Almalaurea** consentirà di individuare i fattori predefiniti che contribuiscono a determinare gli esiti osservati, superando così i limiti imposti dalla letteratura esistente che si concentra principalmente su fattori come genere, età e area geografica (residenza e/o studio) a causa della mancanza di basi dati adeguate. Avendo a disposizione nei dati amministrativi informazioni anche sulla carriera scolastica pregressa, potremo valutare in che modo la provenienza educativa intesa come formazione scolastica, in termini di indirizzo e votazione finale complessiva, possa influenzare l'abbandono universitario o la votazione finale/tempo di completamento per coloro che concludono il Corso di Studi, al fine di poter suggerire l'adozione di misure volte a favorire il contrasto del fenomeno dell'abbandono o ad individuare particolari criticità nel percorso universitario anche per chi completa gli studi universitari.

Un altro aspetto rilevante che il progetto intende considerare è il ruolo dei pari. Per quanto riguarda l'effetto dei pari nella formazione e sulle scelte del percorso formativo, il gruppo dei pari verrà identificato entro le diverse attività di progetto come: a) il gruppo dei compagni di classe/compagni di corso; b) il gruppo di studenti iscritti alla stessa scuola/università; c) il gruppo di studenti provenienti dalla stessa città/area di origine. Per le analisi, verranno utilizzati dati relativi alla composizione del gruppo dei pari ed alle caratteristiche del gruppo dei pari e dati relativi alla posizione relativa dello studente entro il gruppo dei pari per la definizione a).

## 2. Obiettivi e risultati attesi

Questo progetto di ricerca si fonda sull'osservazione del percorso accademico e lavorativo degli studenti e delle studentesse iscritti/e presso 6 Università Italiane dal 2008/2009 (o dal primo a.a. per il quale sono disponibili dati omogenei in Esse 3) al 2024/25.

Questa mappatura sarà possibile attraverso l'**integrazione di 2 fonti di dati**:

- 1) **dati amministrativi relativi al percorso accademico** della popolazione degli studenti e delle studentesse iscritti ai corsi di studi negli a.a. sopra indicati;
- 2) **dati relativi alla collocazione lavorativa a 1, 3 e 5 anni dopo la laurea, raccolti dal consorzio Almalaurea** per ciascuno Ateneo entro il consorzio, ed utilizzabili con riferimento ai propri neo-laureati e neo-diplomati da ciascun Ateneo per le finalità previste dall'art. 2 dello Statuto di Almalaurea, coerenti con gli obiettivi di questo progetto di ricerca.

Il dataset risultante – indispensabile per perseguire gli obiettivi del progetto - consentirà di **analizzare il percorso universitario fino alla loro collocazione nel mondo del lavoro dopo la laurea**, rispondendo ad articolate ed integrate domande di ricerca promosse dal progetto PE9-GRINS Growing Resilient, Inclusive and Sustainable, Finanziato dal PNRR, Missione 4 (Infrastruttura e Ricerca), Componente 2 (Dalla Ricerca all'Impresa), Investimento 1.3 (Partnership Estese), Tematica 9 (Sostenibilità economica di sistemi e territori).

Il data base costruito per il progetto consentirà di esaminare come caratteristiche istituzionali quali, per esempio, i requisiti di ingresso, grade inflation (ovvero l'aumento della media dei voti alla quale non corrisponde un incremento delle competenze), le caratteristiche dei corsi di studio e degli esami universitari, le tasse universitarie e finanziamenti legati al diritto allo studio, possano influenzare le prospettive accademiche ed occupazionali degli studenti. L'integrazione della base di dati sulle traiettorie formative degli studenti con i contenuti – in termini di competenze generali e trasversali – del percorso formativo universitario permetterà di identificare il set di competenze che ciascun curriculum è in grado di trasferire ed il contributo delle stesse agli esiti sul mercato del lavoro.

Inoltre, i risultati attesi della ricerca comprenderanno:

- a) la **costruzione di indicatori aggregati** a partire dai dati individuali che potranno essere resi disponibili alla governance di Ateneo coinvolto nel progetto per guidare le scelte di gestione dei percorsi di formazione offerti, con potenziali ricadute positive per gli studenti, le famiglie degli studenti e gli Atenei;
- b) la **produzione di articoli scientifici** su tematiche rilevanti per il disegno dei percorsi formativi ed il miglioramento dell'efficacia degli stessi ed efficienza nell'utilizzo delle risorse, in particolare con riferimento al rischio di abbandono nelle fasi di transizione del percorso formativo, il ruolo di differenze di genere, le ricadute di politiche relative al sostegno finanziario degli studenti (definizione delle tasse universitarie ed aiuti finanziari), il valore aggiunto di particolari contenuti dei percorsi formativi – sia con riferimento ai contenuti disciplinari che alle competenze trasversali;

c) la **produzione di reports ed articoli divulgativi** orientati al pubblico generale;

d) la possibilità di produrre **analisi di approfondimento (policy briefs)**, orientando l'analisi ad aspetti che per le singole governance presentano potenziali criticità e fornendo strumenti all'avanguardia nella letteratura scientifica utili a supportare il disegno di eventuali politiche e/o scelte di programmazione con riferimento a contenuti informativi e politiche di sostegno finanziario.

Il progetto nasce dall'interazione tra 6 diverse Università Italiane che collaborano entro il progetto PE9-GRINS Growing Resilient, Inclusive and Sustainable, Finanziato dal PNRR, Missione 4 (Infrastruttura e Ricerca), Componente 2 (Dalla Ricerca all'Impresa), Investimento 1.3 (Partnership Estese), Tematica 9 (Sostenibilità economica di sistemi e territori): Università degli studi di Bologna (Capofila), Università degli studi di Bari, Università degli studi di Bergamo, Università degli studi di Cagliari, Università degli studi di Napoli – Federico II, Università di Roma Tor Vergata. In futuro, questo progetto potrà rappresentare il primo step per la costituzione di un dataset integrato fra questi Atenei e per una ulteriore partnership virtuosa tra queste istituzioni con potenziali rilevanti ricadute sul sistema di formazione nazionale. Diversamente da altre progettualità, il presente progetto è disegnato in modo da poter essere facilmente scalabile.

### 3. Metodologia

La preparazione del dataset prevede la definizione di un codebook con una lista di indicatori di modo che le variabili utilizzate siano codificate in maniera omogenea e confrontabile nel tempo e nello spazio (Vedi Allegato 1).

La metodologia prevede l'utilizzo di metodi statistici tradizionali, tra cui, analisi descrittive, modelli di regressione lineare multivariata, modelli a variabili dipendenti binarie, creazione di indicatori per le università. Gli indicatori aggregati prodotti dal progetto nei diversi ambiti di interesse verranno resi pubblicamente accessibili tramite la piattaforma AMELIA, sviluppata entro il progetto PE9-GRINS.

Per quanto riguarda i modelli di stima e gli strumenti di elaborazione dei dati primari, tramite analisi esplorative di fattibilità e precisione, si adotterà il modello più adeguato a seconda nel tipo di dato analizzato, tra:

1. Linear Regression Model
2. Probability Model (probit/logit or tobit)
3. LASSO (model selection algorithm)
4. Support Vector Machine
5. Boosting
6. Logistic regression
7. Classification/decision tree
8. Random forest
9. Neural network

- 10. Transformer-based model
- 11. Dati sintetici

In una parte delle analisi, i risultati verranno presentati selezionando il modello predittivo più adatto, sottolineando i punti di forza e debolezze di ciascuno. Nella scelta del modello, si userà una porzione del dataset e si faranno test di validità e predizione degli outcome di interesse e la “goodness of fit” delle predizioni.

Inoltre, verrà fatta un'analisi di eterogeneità dove si individueranno le categorie di studenti con risultati simili e diversi rispetto a ciascun outcome, in modo da verificare se ci siano o meno eventuali pattern/meccanismi che spiegano i risultati.

In futuro, si utilizzeranno dati ottenuti tramite questionari/interviste che verranno somministrati dalle singole università a studenti e imprese.

#### 4. Descrizione della procedura

Il codebook definito dal team di ricercatori verrà inviato ai referenti della Direzione sistemi infrastrutture dati, i quali estrarranno e conferiranno i dati in esso indicati al team. Il team procederà all'integrazione dei dati amministrativi con quelli di Almalaurea tramite numero di matricola che sarà poi sostituito con un identificativo numerico casuale per ciascun studente/studentessa (pseudonimizzazione).

I dati pseudonimizzati così ottenuti saranno trattati con le metodologie di analisi sopra descritte per le finalità descritte nel progetto, e coerentemente con le finalità di programmazione, pianificazione, monitoraggio dell'Ateneo.

#### 5. Tipologia dei dati trattati

Il progetto prevede il trattamento di dati personali relativi alle carriere della popolazione studentesca dell'Università degli Studi di Cagliari dall'a.a. 2008-2009 all'a.a. 2024-2025, inclusi i dati di Almalaurea. I criteri di inclusione nel campione sono determinati dall'iscrizione all'Ateneo di Cagliari nel periodo di riferimento indicato nel progetto e/o di eventuali estensioni successivamente concordate. Solo il team di ricercatori, opportunamente istruito, tratterà i dati personali di cui al codebook forniti dai referenti della Direzione sistemi infrastrutture dati.

Non è prevista una forma esplicita di incentivo economico o di altra forma per i partecipanti. In futuro, come previsto dal progetto, anche per rispondere agli esiti della ricerca, le singole Università potrebbero voler attivare approfondimenti specifici.

**Titolare del trattamento:** Università degli studi di Cagliari

**Responsabile della protezione dati:** dpo@unica.it

**Responsabile scientifico UNICA:** Prof.ssa Adriana Di Liberto

## Personale coinvolto

Nome e cognome: Adriana Di Liberto  
Dipartimento di Scienze Economiche ed Aziendali  
Università degli Studi di Cagliari  
Qualifica: Prof.ssa Ordinaria  
Indirizzo: Viale Sant'Ignazio 17, 09124, Cagliari, Italia  
Telefono: +39 3355378693  
E-mail: diliberto@unica.it

Nome e cognome: Silvia Balia  
Dipartimento di Scienze Economiche ed Aziendali  
Università degli Studi di Cagliari  
Qualifica: Prof.ssa Associata  
Indirizzo: Viale Sant'Ignazio 17, 09124, Cagliari, Italia  
Telefono: +39 3286442420  
E-mail: silvia.balia@unica.it

Nome e cognome: Marco Nieddu  
Dipartimento di Scienze Economiche ed Aziendali  
Università degli Studi di Cagliari  
Qualifica: Ricercatore a tempo determinato di tipo B  
Indirizzo: Viale Sant'Ignazio 17, 09124, Cagliari, Italia  
Telefono: +39 3929795591  
E-mail: mgnieddu@unica.it

Nome e cognome: Francesca Ghinami  
Dipartimento di Scienze Economiche ed Aziendali  
Università degli Studi di Cagliari  
Qualifica: Assegnista di ricerca  
Indirizzo: Viale Sant'Ignazio 17, 09124, Cagliari, Italia  
Telefono: +39 3931963410  
E-mail: francesca.ghinami@unica.it

Nome e cognome: Elisa Melis  
Dipartimento di Scienze Economiche ed Aziendali  
Università degli Studi di Cagliari  
Qualifica: Dottoranda di ricerca  
Indirizzo: Viale Sant'Ignazio 17, 09124, Cagliari, Italia  
Telefono: +39 3287840880  
E-mail: elisa.melis94@unica.it

Nome e cognome: Daniele Casula  
Dipartimento di Scienze Economiche ed Aziendali  
Università degli Studi di Cagliari  
Qualifica: Borsista di ricerca  
Indirizzo: Viale Sant'Ignazio 17, 09124, Cagliari, Italia  
Telefono: +39 3453907448  
E-mail: daniele.casula@unica.it

Nome e cognome: Andrea Caria  
Dipartimento di Scienze Economiche ed Aziendali  
Università degli Studi di Cagliari  
Qualifica: Ricercatore a tempo determinato di tipo A  
Indirizzo: Viale Sant'Ignazio 17, 09124, Cagliari, Italia  
Telefono: +39 3476301465  
E-mail: andrea.caria91@unica.it

Nome e cognome: Fiorella Ferro  
Dipartimento di Scienze Economiche ed Aziendali  
Università degli Studi di Cagliari  
Qualifica: Borsista di ricerca  
Indirizzo: Viale Sant'Ignazio 17, 09124, Cagliari, Italia  
Telefono: +39 3463281308  
E-mail: fiorella.ferro@unica.it

## 6. Modalità e procedure

Il progetto NON richiede l'approvazione del comitato etico.  
Il progetto NON comporta rischi per i partecipanti.

Il progetto nel complesso può portare notevoli benefici: a) direttamente ai soggetti coinvolti (per esempio identificando fattori di rischio per il drop-out e direzioni per la programmazione di misure correttive e di supporto); b) direttamente alla governance dei singoli Atenei ed indirettamente agli studenti, dando riscontri sull'efficacia di politiche di supporto all'accesso alla formazione universitaria per eventuale ri-progettazione e miglioramento; c) fornendo indicazioni utili alle famiglie sulle diverse caratteristiche dei percorsi formativi in termini di competenze disciplinari e trasversali; d) alla governance degli Atenei ed alla comunità scientifica nazionale ed internazionale, costruendo un data base integrato utile per le attività di ricerca con accesso controllato nel rispetto della regolamentazione nazionale ed europea in tema di privacy, utilizzando tecnologie all'avanguardia per la protezione dei dati personali quali i dati sintetici.

Lo studio è no-profit e osservazionale e viene utilizzata l'assicurazione d'Ateneo.

L'interessato, informato del progetto, potrà, in ogni momento, richiedere l'esclusione dallo stesso e la cancellazione dei propri dati personali raccolti per le finalità del progetto fino a quando il team di ricercatori non avrà provveduto all'anonimizzazione irreversibile degli stessi.

## 7. Base giuridica e finalità della ricerca

Il team di ricerca analizza i dati relativi al percorso accademico e lavorativo degli studenti e delle studentesse iscritti/e dell'Università degli Studi di Cagliari dall'a.a. 2008/2009 al a.a. 2024/2025 nell'ambito dello svolgimento di un compito di interesse pubblico di cui all'art. 6, par. 1 lett. e) per finalità di ricerca scientifica volte allo studio di politiche di orientamento, didattica, placement e internazionalizzazione.

Gli interessati sono informati mediante avviso nella sezione relativa al progetto della pagina web del Dipartimento di Scienze Economiche e Aziendali di UniCa, laddove possono consultare anche l'informativa al trattamento dei dati personali.

## 8. Comunicazione e diffusione dei dati

I dati personali (non anonimi o aggregati) non saranno diffusi o trasferiti all'estero.

## 9. Misure di sicurezza

I dati personali raccolti nell'ambito del progetto sono pseudonomizzati dal team di ricercatori mediante l'utilizzo di codici numerici univoci per ciascun interessato. Solo il responsabile scientifico del progetto può collegare i codici all'identità degli interessati.

Per anonimizzare o aggregare i dati, anche in un momento successivo alla raccolta, i dati personali sono sostituiti da uno o più identificatori, che possono essere utilizzati per un set di dati o per ogni singolo dato.

Il database integrato pseudonimizzato è conservato e utilizzato su server dell'Ateneo o su pc forniti dall'Ateneo ai ricercatori coinvolti.

Il team di ricercatori è istruito e responsabile per il trattamento e la conservazione dei dati secondo la normativa vigente e le policy privacy di Ateneo.

## 10. Conservazione

I dati saranno conservati per 10 anni dopo il termine della formazione universitaria per ciascuno studente iscritto in ciascun anno accademico considerato nel progetto, in modo da permettere l'integrazione con esiti sul mercato del lavoro rilevati dal questionario Almalaurea a 1, 3 e 5 anni - tenendo conto dei tempi

relativi a consolidamento del dato, ed eventuali verifiche di qualità e/o ratifiche in caso i soggetti decidessero essere esclusi.

Al termine di questo periodo i dati saranno conservati in forma anonima in formato digitale.

## 11. Informazioni aggiuntive

Il progetto NON prevede:

- l'autorizzazione di altri Enti/soggetti terzi per l'accesso ai dati o per il coinvolgimento di partecipanti;
- ai sensi della normativa vigente, interventi che richiedono specifiche professionalità (ad es. medico, psicologo, infermiere, ecc.)