

# Progetto *Carriere Studenti* (analisi dei requisiti)



Paolo Atzeni, Luigi Bellomarini



# Obiettivi

- Realizzare alcuni data mart per la autovalutazione di Ateneo:
  - **carriere degli studenti**
    - ne studieremo i requisiti (oggi, insieme)
    - **progettazione dimensionale (parte I)**
    - **implementazione data mart in Business Analytics (parte II)**
    - **implementazione ETL in Pentaho Data Integration (parte II)**
  - **offerta didattica**
    - lo analizzeremo, ma non lo realizzeremo
- **Un esercizio, ma anche un progetto DWH reale**



# Homework (modalità)

- **Modalità di lavoro**
  - piccoli gruppi (1, 2 o 3 persone)
  - il lavoro può essere svolto in collaborazione
  - gli esperimenti devono essere svolti, almeno in parte, in maniera individuale
  - la consegna è individuale
- **Realizzazione di un progetto DWH composto di 2 parti**
- **Ogni homework ne costituisce una parte**
  - **homework 1: progettazione dimensionale**
    - **non occorre consegnare quanto prodotto**
  - **homework 2: implementazione data mart + ETL**
    - **consegna entro una data di scadenza che sarà comunicata**

# Homework

## (gestione codice parte II)

- ogni **gruppo** crea un repository (pubblico) su GitHub (<https://github.com/>)
- un repository avrà quindi URL:
  - <https://github.com/<nome gruppo>/<nome repo>>
  - es: <https://github.com/myGroup/myRepo>
- nel repository si creano 2 **directory**:
  - **olap/**
  - **etl/**
  - e un file README.md (da salvare nella root del repository, fuori da tutte le directory)
  - contenente **nome e cognome di tutti i membri del gruppo** (senza virgole, su righe diverse in ordine alfabetico)

# Homework

## (consegna parte II)

- per la sola parte 2, entro la data consegna, ogni studente
  - indica su **Moodle**:
    - “Consegna homework: <URL del repository del proprio gruppo>”
    - es: consegna homework: - <https://github.com/bellomarini/myDWproject>
    - quindi gli studenti di uno stesso gruppo indicheranno in Moodle tutti la stessa URL e i loro nomi saranno tutti anche in README.md
- per la sola parte 2, entro la data di consegna, ogni gruppo
  - effettua il PUSH su **GitHub** del codice relativo

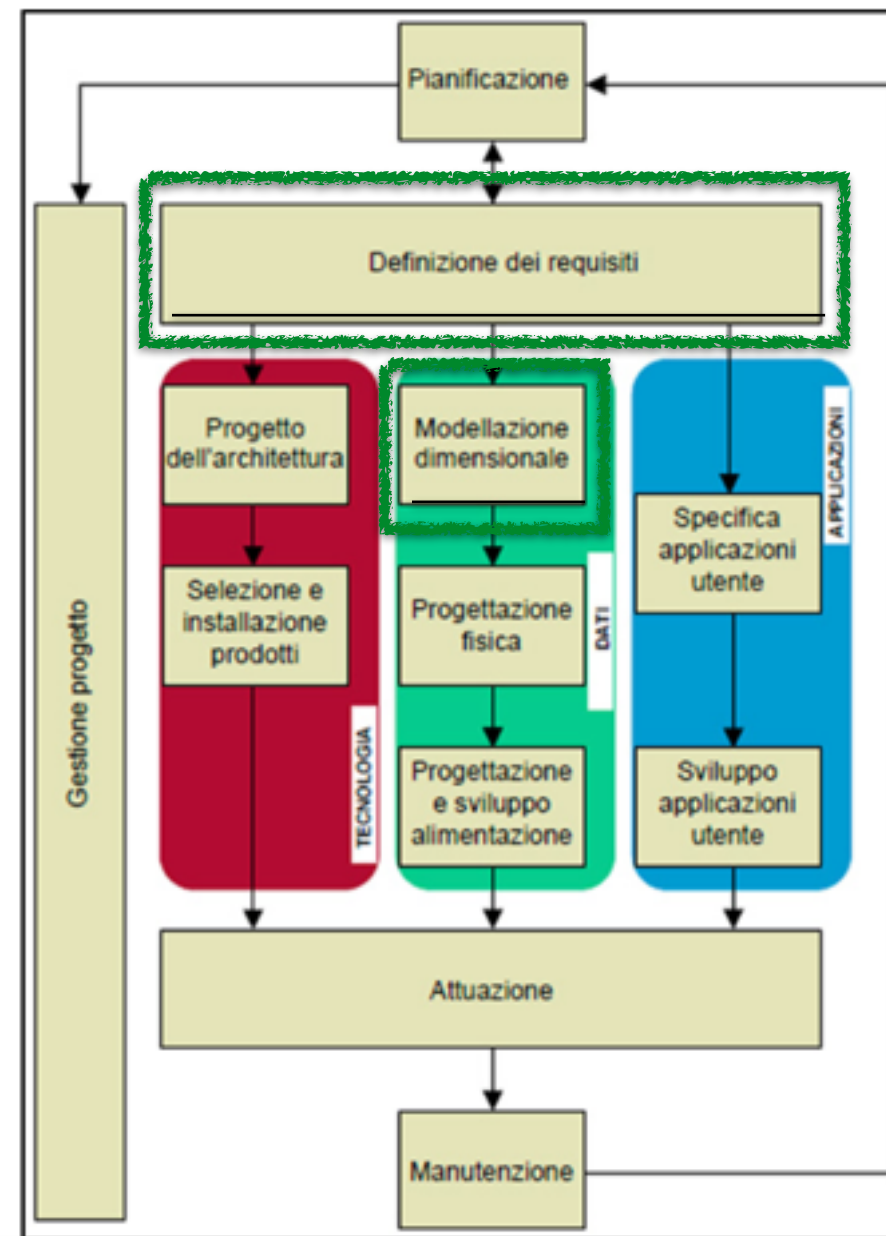
# Homework I

- **Progettazione del data mart relativo alla carriera degli studenti**
- Progettare il data mart per la nostra realtà di interesse (che vedremo oggi)
  - varie fact table
  - varie dimensioni
  - eventualmente correggendo/estendendo l'analisi che vedremo
- La progettazione può essere effettuata utilizzando lo strumento BI-modeler
  - <http://www.bimodeler.com/download/>
- **Non si deve consegnare**
- **Soluzione discussa in aula presumibilmente il 15 marzo**

# Ciclo vita sistemi data warehousing

- Esistono **varie metodologie** di progetto, sviluppo e esercizio dei sistemi di data warehousing.
- **Top-down**
  - Promettono risultati generali.
  - Risultati rilasciati in tempi lunghi.
  - Rischio fallimento (vd. processi a cascata).
- **Bottom-up**
  - Costruzione in modo incrementale del DW.
  - Scope limitato ma risultati immediati.
  - Minore rischio di fallimento (vd. metodi agili).

# Business Dimensional Lifecycle (Kimball, 1998)

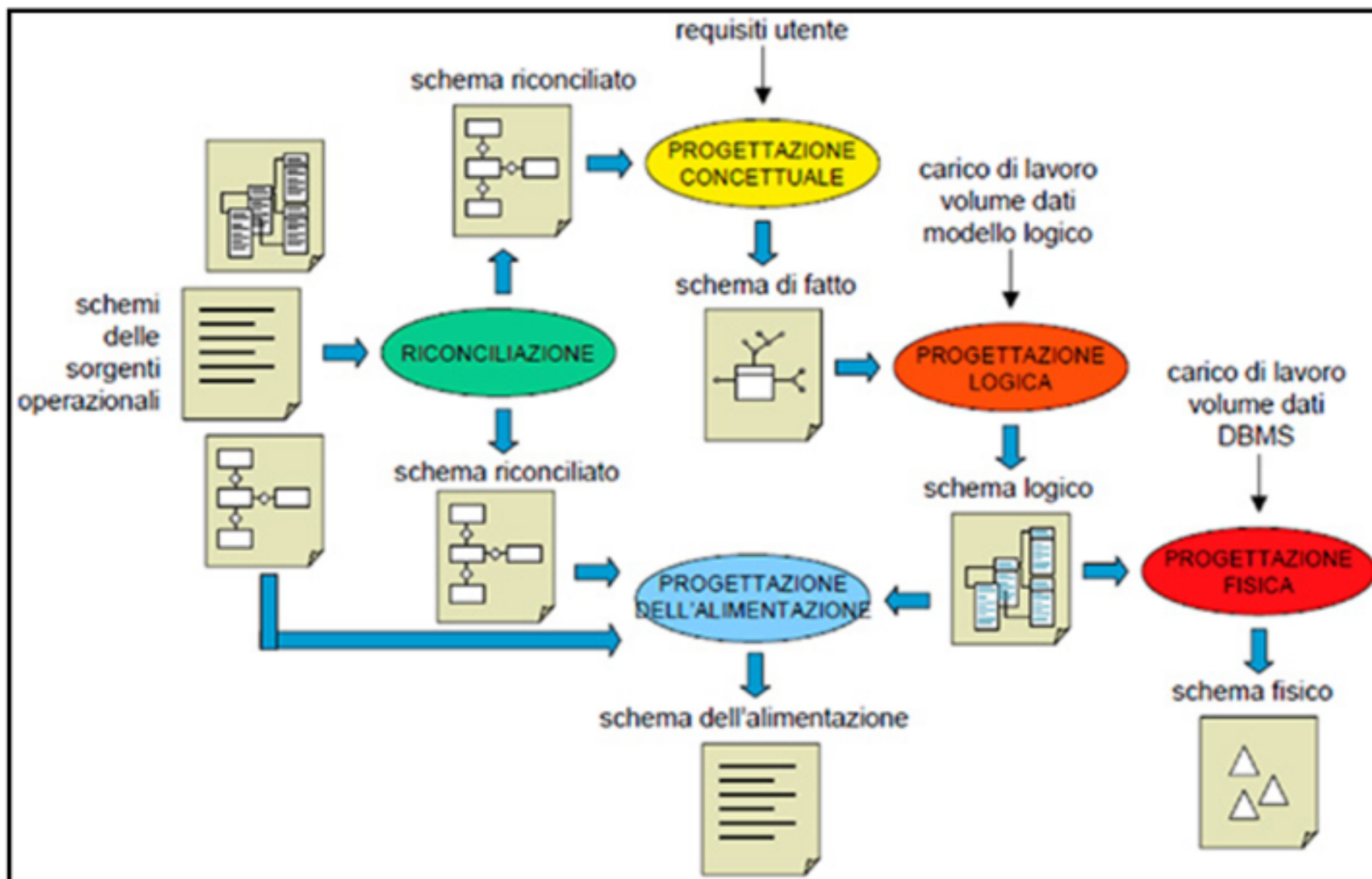




# Realizzazione di data mart

- **Definizione dei requisiti**
  - analisi e riconciliazione delle fonti di dati
  - specifica dei requisiti
  - carico di lavoro preliminare
- **Modellazione dimensionale**
  - Progettazione concettuale
    - guidata dai dati
    - guidata dai requisiti
    - **approccio misto**
  - Analisi del carico di lavoro
  - Progettazione logica
- Progettazione fisica
- Progettazione dell'ETL (o dell'alimentazione)

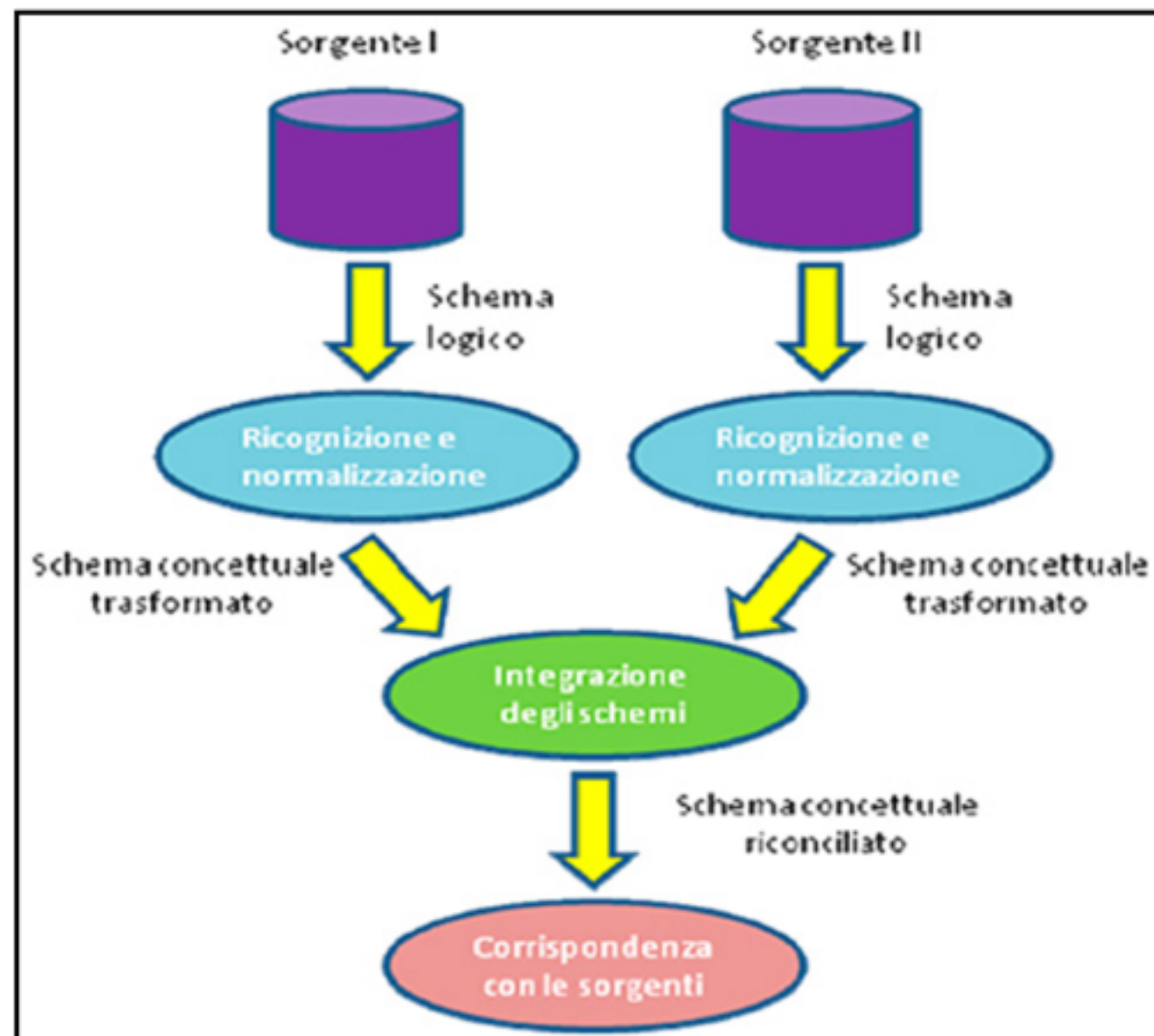
# Realizzazione di un data mart



# Analisi e riconciliazione delle fonti di dati

- Definire e documentare lo schema del livello operativo.
- Analizzare e comprendere gli schemi delle sorgenti disponibili.
  - Spesso a disposizione solo gli schemi logici.
  - Uso di formalismi concettuali.
- Determinarne le porzioni utili.
- Integrazione
  - A livello **intensionale**: schema matching
  - A livello **estensionale**: pulizia, trasformazione, entity resolution, fusion
- Fase complessa

# Analisi e riconciliazione delle fonti di dati



# Specifica dei requisiti

- Il progettista filtra e raccoglie i requisiti degli utenti finali.
  - Quali sono le informazioni di interesse.
  - Informazioni di interesse in linea con gli obiettivi strategici.
  - Si devono produrre delle specifiche sui fatti da modellare.
- Vari formalismi per la modellazione dell'organizzazione, più o meno efficaci.
- Linguaggi per modellare l'Enterprise Architecture.
  - Goal, Target, etc.

# Carico di lavoro preliminare

- Indagine sulle interrogazioni principali che saranno effettuate sul sistema.
- Studio degli obiettivi di ciascuna categoria di utenti.
- Studio della reportistica aziendale esistente.
- Studio delle applicazioni cruscotto esistenti.
- Integrazione dell'analisi della reportistica con i risultati delle interviste.

# DWH

## *Carriere Studenti*



# Analisi e riconciliazione delle fonti di dati

- Le fonti di dati **reali**, nel nostro caso, sono:
  - file XML o CSV spediti dall'Ateneo per alimentare l'**Anagrafe Nazionale degli Studenti**
- Saranno rilevanti quando realizzeremo l'ETL. Per il momento assumiamo di avere:
  - base di dati operativa **carriere degli studenti**
  - base di dati operativa **offerta didattica**
- Sistemi operazionali non direttamente accessibili, ma **operational data store** ipotetico:
  - validare l'operational data store con il carico di lavoro
  - validare lo schema concettuale progettato con il carico di lavoro



# Specifica dei requisiti

- In generale, l'obiettivo è semplificare la **autovalutazione dell'offerta formativa e della didattica**
- miglioramento dell'efficienza e della qualità dei servizi offerti.
- Produzione di report (per l'**ANVUR**) per la valutazione periodica dell'attività di Ateneo.
- Produzione di **report estemporanei** per terze parti.
- Interrogazioni estemporanee da parte della **dirigenza** o dei **nuclei per la valutazione**.

# Glossario

- **Corso di Studi (Cds)**
  - CdL, CdLM (è, in un certo senso, il tipo di Cds)
- **Coorte (statistica)**
  - Insieme di individui facenti parte di una popolazione che hanno sperimentato uno stesso evento in un particolare intervallo di tempo.
  - Nel nostro caso, l'insieme degli studenti iscritti per la prima volta al primo anno di un corso di laurea in un certo anno accademico.
- **CFU**
- **ECTS**
  - Sistema di conversione europeo per il trasferimento dei risultati accademici ottenuti dagli studenti.

# ANVUR



- Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca
- Sistema **AVA**
  - **Autovalutazione, Valutazione periodica e Accreditamento.**
  - Insieme delle attività ANVUR in attuazione della legge n.240 del 20/12/2010 e del D.Lgs. n.19 del 27/10/2012.
  - Prevede, tra l'altro, la verifica e il monitoraggio di parametri e indicatori.
  - Ripartizione periodica delle risorse assegnate (quota premiale) alle università.

# Carico di lavoro preliminare

- Abbiamo accesso ai report che attualmente è possibile ottenere.
- Normalmente si procede per interviste agli utenti.
- **Analisi** dei report prodotti per comprendere, inizialmente, le esigenze informative.
- Nella **progettazione** si cercherà di produrre schemi dimensionali per rispondere a queste esigenze.
- Schemi sufficientemente generali per saper rispondere anche ad interrogazioni non (completamente) previste.

# Report sulla carriera degli studenti

- Tutti i report sono per Cds o per tutti i Cds e per Anno Accademico di riferimento.
- Domande di preimmatricolazione e immatricolazione (RC.1)
- Attrattività del Cds (Followup) (RC.2)
- Motivazioni di abbandono (RC.3)
- Esiti del percorso formativo (per range CFU) (RC.4.1)
- Esiti del percorso formativo (medie) (RC.4.2)
- Tipo Maturità (RC.5.1)
- Voto Maturità (RC.5.2)
- Votazione riportata nel titolo di primo livello (RC.6)
- Dati sui laureati del CdS (RC.7)
- Numero medio di CFU conseguiti dagli iscritti per Anno Accademico (RC.8)
- Propensione dei laureati triennali ad immatricolarsi ad un Corso di Laurea Magistrale (RC.9)

# Domande di preimmatricolazione e immatricolazione (RC.I)

- **AA COORTE:** A.A. coorte (fino a quello di riferimento)
- **POSTI O UTENZA SOST:** N° programmato/utenza sostenibile
- **DOMANDE PREIMM:** numero di domande di preimmatricolazione con opzione per il CdL/CdLM
- **DI CUI PARTECIPANTI TEST:** numero di preimmatricolati che ha partecipato al test di ingresso
- **DI CUI IMMATRICOLATI:** numero di preimmatricolati che si è immatricolato al CdL/CdLM
- **MEDIA PUNTEGGIO:** punteggio medio dei partecipanti al test di ingresso (ponderato in 100-mi)
- **MEDIA PUNTEGGIO IMM:** punteggio medio degli immatricolati al CdS (ponderato in 100-mi)

# Attrattività del Cds (RC.2)

- **AA COORTE:** A.A. coorte (fino a quello di riferimento)
- **IMM TOT:** ° di immatricolati
- **PERC TRIENNALE ALTRO ATENEO:** percentuale di studenti che ha conseguito il titolo di 1° livello in ateneo diverso da Roma Tre (per i CdLM)
- **ISCR II ANNO INCORSO:** numero di studenti che nell'a.a successivo frequenta lo stesso corso al 2° anno in corso
- **ISCR I ANNO RIP:** numero di studenti che nell'a.a successivo frequenta lo stesso corso al 1° anno ripetente
- **TASSO ABBANDONO:** tasso di abbandono al 1° anno (il tasso di abbandono è calcolato escludendo a numeratore e denominatore gli studenti che si sono laureati in un numero di anni inferiore alla durata legale).

# Motivazioni dell'abbandono (RC.3)

- **AA COORTE:** A.A. coorte (fino a quello di riferimento)
- **IMM TOT:** N° di immatricolati
- **PCT PASS INT:** Percentuale di studenti che è passato ad altro CdL/CdLM nell'ambito dell'offerta dello stesso Dipartimento di origine
- **PCT PASS EST:** Percentuale di studenti che è passato ad altro CdL/CdLM nell'ambito dell'offerta di un Dipartimento diverso da quello di origine
- **PCT RIN:** Percentuale di studenti che ha effettuato rinuncia formale agli studi
- **PCT TRASF:** Percentuale di studenti che si è trasferito presso altro ateneo
- **PCT MANC RINN:** Percentuale di studenti che non ha rinnovato l'iscrizione per l'a.a. successivo
- **PCT LAU:** Percentuale di studenti che ha conseguito il titolo
- Le percentuali sono riferite a coloro che NON frequentano lo stesso corso di studi l'a.a. successivo a quello di immatricolazione.
- Il loro numero è ricavabile per differenza sulla base del report Attrattività del CdS (RC.2).



# Esiti del percorso formativo (per range CFU) (RC.4.I)

- **AA COORTE:** A.A. coorte (fino a quello di riferimento)
  - **IMM TOT:** N° di immatricolati
  - **N STUD INCORSO:** N° di studenti che nell'a.a. successivo frequenta lo stesso CdL/CdLM in corso
  - **N STUD RIP:** N° di studenti che nell'a.a. successivo frequenta lo stesso CdL/CdLM ripetente
  - **PCT 0:** Percentuale di studenti che hanno maturato 0 CFU durante il primo anno accademico
  - **PCT FINO 10:** Percentuale di studenti che hanno maturato 1-10 CFU durante il primo anno accademico
  - **PCT FINO 20:** Percentuale di studenti che hanno maturato 11-20 CFU durante il primo anno accademico
  - **PCT FINO 30:** Percentuale di studenti che hanno maturato 21-30 CFU durante il primo anno accademico
  - **PCT FINO 40:** Percentuale di studenti che hanno maturato 31-40 CFU durante il primo anno accademico
  - **PCT FINO 50:** Percentuale di studenti che hanno maturato 41-50 CFU durante il primo anno accademico
  - **PCT FINO 60:** Percentuale di studenti che hanno maturato 51-60 CFU durante il primo anno accademico
  - **PCT FINO 70:** Percentuale di studenti che hanno maturato 61-70 CFU durante il primo anno accademico
  - **PCT FINO 80:** Percentuale di studenti che hanno maturato 71-80 CFU durante il primo anno accademico
  - **PCT FINO 90:** Percentuale di studenti che hanno maturato 81-90 CFU durante il primo anno accademico
  - **PCT FINO 100:** Percentuale di studenti che hanno maturato 91-100 CFU durante il primo anno accademico
  - **PCT OLTRE 100:** Percentuale di studenti che hanno maturato oltre 100 CFU durante il primo anno accademico
- 
- La ricognizione del numero di CFU maturati è effettuata con riferimento all'a.a., senza restrizioni in relazione alla data di superamento degli esami. Sono quindi incluse anche le prove superate nei primi mesi dell'a.a. successivo.

# Esiti del percorso formativo (medie) (RC.4.2)

- **AA COORTE:** A.A. coorte (fino a quello di riferimento)
- **IMM TOT:** N° di immatricolati
- **N STUDENTI SOPRAVV:** N° di studenti che nell'a.a. successivo frequenta lo stesso CdL/CdLM
- **MEDIA CFU STESSO CDS II ANNO:** Numero medio di CFU maturato in corrispondenza del primo a.a. di carriera dagli studenti che frequentano lo stesso CdL/CdLM nell'a.a. successivo (in corso e ripetenti)
- **MEDIA CFU IN CORSO:** Numero medio di CFU maturato in corrispondenza del primo a.a. di carriera dagli studenti che frequentano al 2° anno in corso lo stesso CdL/CdLM nell'a.a. successivo
- **MEDIA CFU RIPETENTI:** Numero medio di CFU maturato in corrispondenza del primo a.a. di carriera dagli studenti che frequentano come ripetenti al 1° anno lo stesso CdL/CdLM nell'a.a. successivo

# Tipo maturità (RC.5.1)

- **AA COORTE:** A.A. coorte (fino a quello di riferimento)
- **IMM TOT:** N° di immatricolati
- **PCT CLASSICA:** Percentuale di immatricolati con maturità classica
- **PCT SCIENT:** Percentuale di immatricolati con maturità scientifica
- **PCT TECNICA:** Percentuale di immatricolati con maturità tecnica
- **PCT PROF:** Percentuale di immatricolati con maturità professionale
- **PCT MAGIST:** Percentuale di immatricolati con maturità magistrale
- **PCT LING:** Percentuale di immatricolati con maturità linguistica
- **PCT ART:** Percentuale di immatricolati con maturità artistica
- **PCT STRANIERA:** Percentuale di immatricolati con maturità straniera
- **PCT NON DISP:** Percentuale di immatricolati con maturità non disponibile

# Voto maturità (RC.5.2)

- **AA COORTE:** A.A. coorte (fino a quello di riferimento)
  - **IMM TOT:** N° di immatricolati
  - **PCT 60 70:** Percentuale di immatricolati con votazione 60-70
  - **PCT 71 80:** Percentuale di immatricolati con votazione 71-80
  - **PCT 81 90:** Percentuale di immatricolati con votazione 81-90
  - **PCT 91 100:** Percentuale di immatricolati con votazione 91-100
  - **PCT NON DISP:** Percentuale di immatricolati con maturità straniera
  - **PCT STRANIERA:** Percentuale di immatricolati con votazione non disponibile
- 
- Le votazioni conseguite dagli studenti che hanno conseguito il titolo di scuola secondaria con un sistema di votazione diverso dai 100mi devono essere opportunamente ponderate.

# Votazione riportata nel titolo di primo livello (RC.6)

- **AA COORTE:** A.A. coorte (fino a quello di riferimento)
- **IMM TOT:** N° di immatricolati
- **PCT 66 90:** Percentuale di immatricolati con votazione 66-90 su 110
- **PCT 91 100:** Percentuale di immatricolati con votazione 91-100 su 110
- **PCT 101 105:** Percentuale di immatricolati con votazione 101-105 su 110
- **PCT 106 110:** Percentuale di immatricolati con votazione 106-110 su 110
- **PCT NON DISP:** Percentuale di immatricolati con votazione non disponibile
- **PCT TIT STRANIERO:** Percentuale di immatricolati con titolo di studio straniero
- Le votazioni conseguite dagli studenti che hanno conseguito il titolo di 1° livello con un sistema di votazione diverso dai 110mi devono essere opportunamente ponderate.

# Dati sui laureati del Cds (RC.7)

- **AA COORTE:** A.A. coorte (fino a quello di riferimento)
- **IMM TOT:** N° di immatricolati
- **LAU TOT:** N° di studenti della coorte laureatisi nel CdL/CdLM di partenza
- **PERC IN CORSO:** Percentuale di studenti laureati nella durata legale del CdL/CdLM
- **PERC 1 ANNO FC:** Percentuale di studenti laureati con 1 anno f.c. rispetto alla durata legale del CdL/CdLM
- **PERC 2 ANNI FC:** Percentuale di studenti laureati con 2 anni f.c. rispetto alla durata legale del CdL/CdLM
- **PERC 3 ANNI FC O PIÙ:** Percentuale di studenti laureati con 3 anni f.c. o più rispetto alla durata legale del CdL/CdLM
- **MEDIA DURATA CARR CDS:** Durata media della carriera degli studenti della coorte laureati nel CdL/CdLM
- **MEDIA DURATA CARR STR:** Durata media della carriera degli studenti laureati nei CdL/CdLM del Dipartimento nell'ambito della stessa coorte
- **MEDIA VOTO:** Voto medio di laurea
- La durata media delle carriere a livello di dipartimento è calcolato con riferimento agli studenti della stessa coorte di origine nell'ambito dello stesso livello, valutando solamente coloro che hanno conseguito il titolo nello stesso CdL/CdLM di immatricolazione.

# Numero medio di CFU conseguiti dagli iscritti per Anno Accademico (RC.8)

- **AA:** Anno Accademico di riferimento
- **N STUDENTI:** N° di iscritti all'AA di riferimento
- **MEDIA CFU GENERALE:** Numero medio di CFU maturato nell'A.A. (tutti gli anni di corso)
- **MEDIA CFU 1 ANNO:** Numero medio di CFU maturato nell'A.A. (1° anno di corso)
- **MEDIA CFU 2 ANNO:** Numero medio di CFU maturato nell'A.A. (2° anno di corso)
- **MEDIA CFU 3 ANNO:** Numero medio di CFU maturato nell'A.A. (3° anno di corso)
- **MEDIA CFU 4 ANNO:** Numero medio di CFU maturato nell'A.A. (4° anno di corso)
- **MEDIA CFU 5 ANNO:** Numero medio di CFU maturato nell'A.A. (5° anno di corso)
- **MEDIA CFU STR:** Numero medio di CFU maturati nell'A.A. dagli iscritti ai corsi dello stesso livello e dipartimento

# RC8 : note

- Le carriere attive nel corso sono ricostruite conteggiando anche gli studenti trasferiti e rinunciatari, qualora la data di chiusura carriera sia successiva al 31 gennaio.
- Si devono conteggiare anche gli studenti che hanno effettuato un passaggio da un corso ex D.M. 509, valutando tutti i CFU maturati nell'ambito del corso di destinazione nell'A.A. di riferimento.
- Quindi si conteggiano i CFU maturati a fronte di esami superati realmente nell'A.A., mentre vanno esclusi quelli derivati da esami convalidati da precedenti carriere.
- L'indicatore costituisce quindi una misura della produttività media degli studenti, non della loro posizione rispetto ai CFU teorici da conseguire.
- Sono inclusi anche i CFU attribuiti all'atto della discussione della tesi e le idoneità, se conseguite nell'A.A. di riferimento.



# Propensione dei laureati triennali ad immatricolarsi ad un Corso di Laurea Magistrale (RC.9)

- **AA LAUREA:** A.A. laurea
- **TOT LAU:** N° di laureati
- **PCT NO IMM:** Percentuale di laureati non immatricolati ad un CdLM
- **PCT DOPO 1 ANNO:** Percentuale di laureati immatricolati ad un CdLM l'a.a. successivo a quello di laurea
- **PCT DOPO 2 ANNI:** Percentuale di laureati immatricolati ad un CdLM dopo due a.a.
- **PCT DOPO 3 ANNI:** Percentuale di laureati immatricolati ad un CdLM dopo tre a.a. o più
- I laureati sono quelli complessivi, senza restrizioni in relazione all'iter universitario (coorte, provenienza).

# Report sugli insegnamenti impartiti e sui docenti

- Docenti (RI.1.1)
- Docenti esterni (RI.1.2)
- Insegnamenti e tipologie di copertura (RI.2)
- Esami di profitto relativi agli insegnamenti del Cds (RI.3)
- Statistiche della distribuzione dei voti (RI.4.1)
- Statistiche ECTS della distribuzione dei voti (5 valori) (RI.4.2).

# Docenti (RI.1.1)

- **DOCENTE:** Nominativo del Docente
- **RUOLO:** Ruolo del Docente (Ordinario, Associato, Ricercatore, Assistente, Assegnista di Ricerca, Ricercatore a Tempo Determinato)
- **SSD DOCENTE:** Settore Scientifico Disciplinare di appartenenza del docente (es: ING-INF/05, ING-INF/03)
- **IMPEGNO:** Percentuale di impegno stimato per lo specifico CdL/CdLM. La percentuale deve essere stimata ripartendo l'impegno in percentuale uguale per il numero di CdL/CdLM ai quali sono associate attività didattiche reali (che non sono destinatarie di mutuaione/condivisione) in cui il docente è coinvolto a vario titolo.
- Le attività didattiche **reali** sono derivabili dalla voce **Unità Didattica** del report sugli Insegnamenti (RI.2).

# Docenti esterni (RI.1.2)

- **DOCENTE:** Nominativo del Docente
- **RUOLO:** Ruolo del Docente (Docente Esterno; Esperto della materia)
- **IMPEGNO:** Percentuale di impegno stimato per lo specifico CdL/CdLM. La percentuale è stimata ripartendo l'impegno in percentuale uguale per il numero di CdL/CdLM ai quali sono associate attività didattiche reali (che non sono destinatarie di mutuaione/condivisione) in cui il docente è coinvolto a vario titolo.
- Le attività didattiche **reali** sono derivabili dalla voce **Unità Didattica** del report sugli Insegnamenti (RI.2).

# Insegnamenti e tipologie di copertura (RI.2)

- **AD DES:** Nome dell'Attività Didattica (es: “Algoritmi e strutture dati”)
- **AD COD:** Codice dell'Attività Didattica
- **UD DES:** Nome dell'Unità Didattica (es: “Primo modulo”). Un'attività didattica (AD) può essere costituita da più unità didattiche (UD) vale a dire da più moduli
- **PARTIZIONAMENTO:** Suddivisione in canali (es: “Canale A-L”)
- **SSD:** Settore Scientifico Disciplinare dell'Unità Didattica (es: “INF-INF/05”)
- **SSD DOCENTE:** Settore Scientifico Disciplinare del docente
- **DOCENTE:** Nominativo del Docente
- **COPERTURA:** Tipologia di copertura (es: “Compito didattico”, “Esterno”)
- **CFU AD:** Crediti Formativi Universitari associati all'Attività Didattica
- **CFU UD:** Crediti Formativi Universitari associati all'Unità Didattica
- **ORE LEZIONE FRONTALE:** Ore di lezione frontali associate all'Unità Didattica
- **CONDIVISIONE:** In caso di condivisione dell'unità didattica con altri insegnamenti, viene riportato il codice dell'Attività Didattica e del CdS che eroga l'attività reale

# Esami di profitto relativi agli insegnamenti del Cds (RI.3)

- **AD:** Nome dell'Attività Didattica
- **AD COD:** Codice dell'Attività Didattica
- **DOCENTE:** Nominativo del Docente presidente della Commissione d'Esame
- **STESSO CDS:** SI: il numero di studenti che ha sostenuto l'esame e che apparteneva allo stesso CdL/CdLM in cui l'Attività Didattica era erogata. NO: il numero di studenti che ha sostenuto l'esame e che non apparteneva allo stesso CdL/CdLM in cui l'Attività Didattica era erogata.
- **2011:** Esami sostenuti nel terzultimo anno solare precedente a quello di riferimento (ad es. 2014)
- **2012:** Esami sostenuti nel penultimo anno solare precedente a quello di riferimento
- **2013:** Esami sostenuti nel ultimo anno solare precedente a quello di riferimento
- **2014:** Esami sostenuti nell'anno solare di riferimento
- Gli esami vengono conteggiati sulla base del numero di atti verbalizzazioni di attività didattiche.
- Vengono conteggiati anche eventuali esami non presenti nella carriera dello studente per problemi amministrativi (iscrizione non regolare, propedeuticità non rispettata, etc.)

# Statistiche della distribuzione dei voti (RI.4.1)

- **VOTO:** Voto espresso in 30-esimi (la granularità delle fasce è quella del voto).
- **PERCENTUALE:** Percentuale di esami che ricadono nel voto in esame.
- **NUMERO ESAMI:** Numero di esami che ricadono nella fascia in esame.
- I dati sono relativi all'intervallo statistico:
  - ultimi tre anni per le lauree Triennali
  - ultimi due anni per le lauree Magistrali
  - ultimi cinque anni per le lauree Magistrali a Ciclo Unico

# Statistiche ECTS della distribuzione dei voti (5 valori) (RI.4.2)

- **VOTO ECTS:** Voto espresso secondo lo standard ECTS (decrescente da A ad E)
- **PERCENTUALE:** Percentuale di esami che ricadono nella fascia associata al livello ECTS
- **VOTO MASSIMO:** Voto massimo della fascia associata al livello ECTS (limite superiore dell'intervallo)
- **VOTO MINIMO:** Voto minimo della fascia associata al livello ECTS (limite inferiore dell'intervallo)
- Intervalli statistici:
  - ultimi tre anni per le lauree Triennali
  - ultimi due anni per le lauree Magistrali
  - ultimi cinque anni per le lauree Magistrali a Ciclo Unico



# Operational data store



# Definizione dell'ODS

- Prescindiamo per il momento dalle sorgenti dati reali e modelliamo un ODS a partire dal carico di lavoro preliminare.
- Progettazione **basata sui requisiti**
  - Si progettano gli schemi dimensionali in maniera indipendente dagli schemi operazionali.
- Progettazione **basata sui dati**
  - Apparentemente semplice.
  - Alta probabilità di fallimento.
- Progettazione con **approccio misto**
  - Fondamentalmente guidato dai requisiti
- L'ODS predispone ad una progettazione con approccio misto (requisiti/dati) anche in assenza della specifica dei sistemi OLTP
  - Aiuta nella comprensione del dominio
  - Verifica nel sistema operativo della disponibilità dei dati necessari

# Progetto dell'architettura

- Impatto della scelta per la riconciliazione delle fonti sull'architettura del DWH.
- La scelta fatta di utilizzo di un ODS supporta (suggerisce) una architettura a tre livelli per il DWH:
  - Sistema operativo (file e db OLTP)
  - Operational data Store
  - Data warehouse

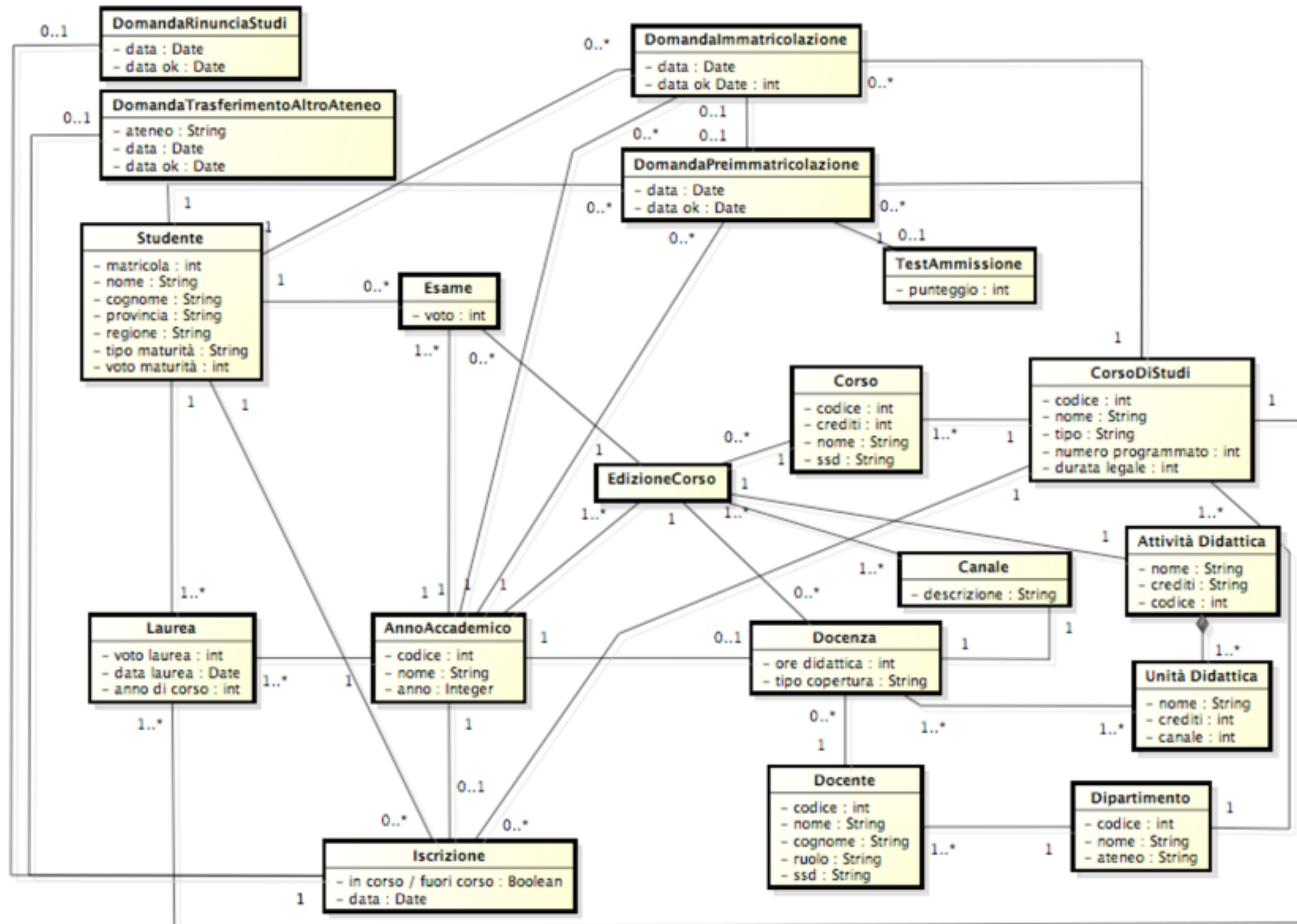
# Operational data store

- Materializzazione dei dati operazionali a valle di integrazione e ripulitura dei dati sorgente.
- Modello di dati comune, di riferimento.
- Separazione tra schemi dimensionali e schemi operazionali.
- Disaccoppiamento della progettazione di schemi e flussi ETL.
- In realtà spesso esiste anche se non materializzato.
  - ODS virtuali.

# Modellazione dell'ODS

- Decomposizione funzionale del dominio di interesse.
  - Tecniche utilizzate per la definizione di modelli di dominio.
  - Schemi E/R.
- Utilizzo di un formalismo concettuale.
  - Ad esempio modelli di dominio UML.
- Applicazione (“esecuzione”) dei report sull'ODS per verificare la sua completezza informativa.
- **Validazione del carico di lavoro** sullo schema concettuale.

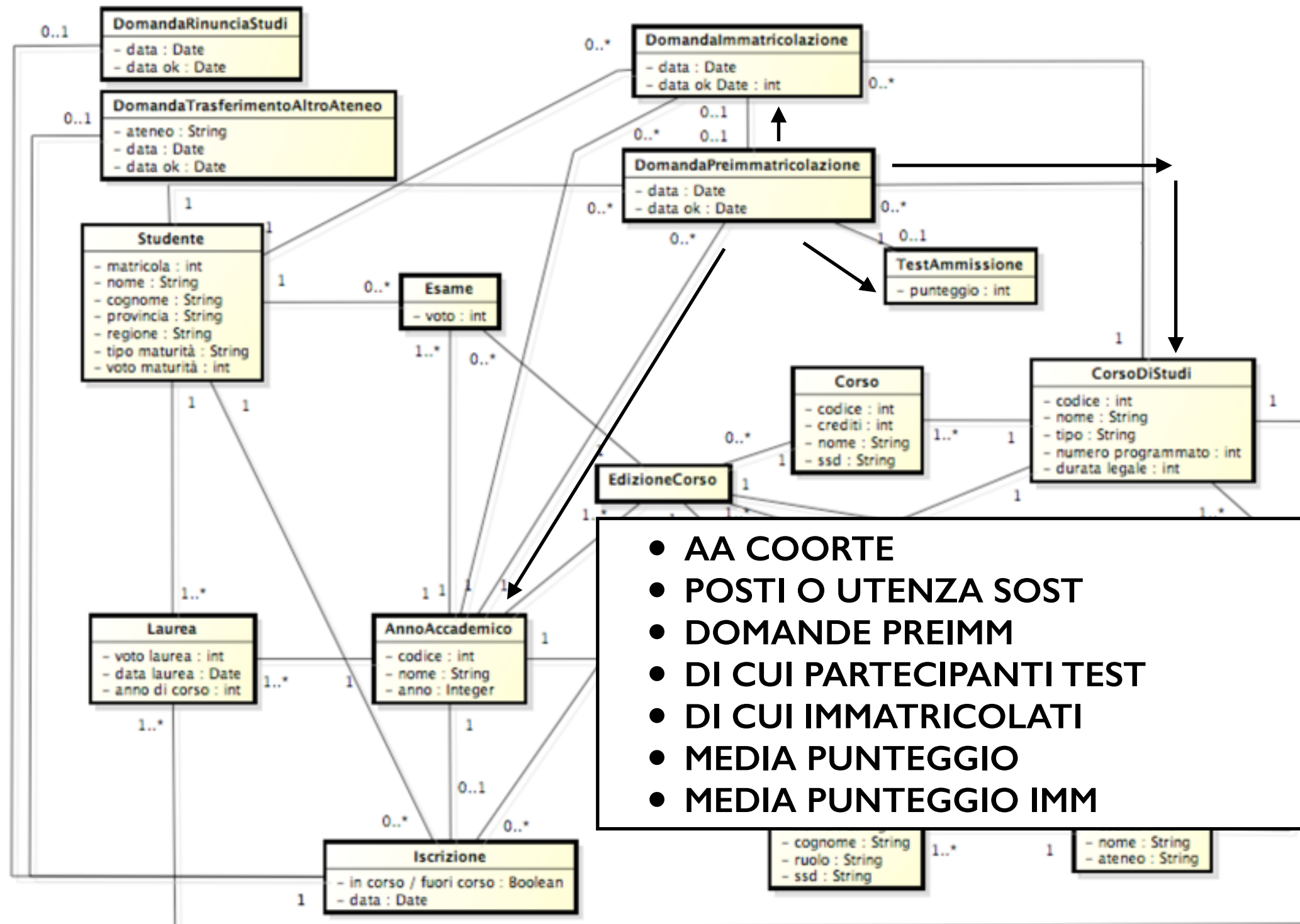
# Operational data store



# Validazione del carico di lavoro sull'ODS

- Tecnica simile alla scrittura dei *contratti delle operazioni* nella analisi orientata agli oggetti.
- Ogni report è un'operazione.
- Tecnica simile agli *schemi di navigazione* nella progettazione di basi di dati relazionali attraverso schemi E/R.
- Frecce che identificano i percorsi di join.

# RC.I Domande di preimmatricolazione e immatricolazione



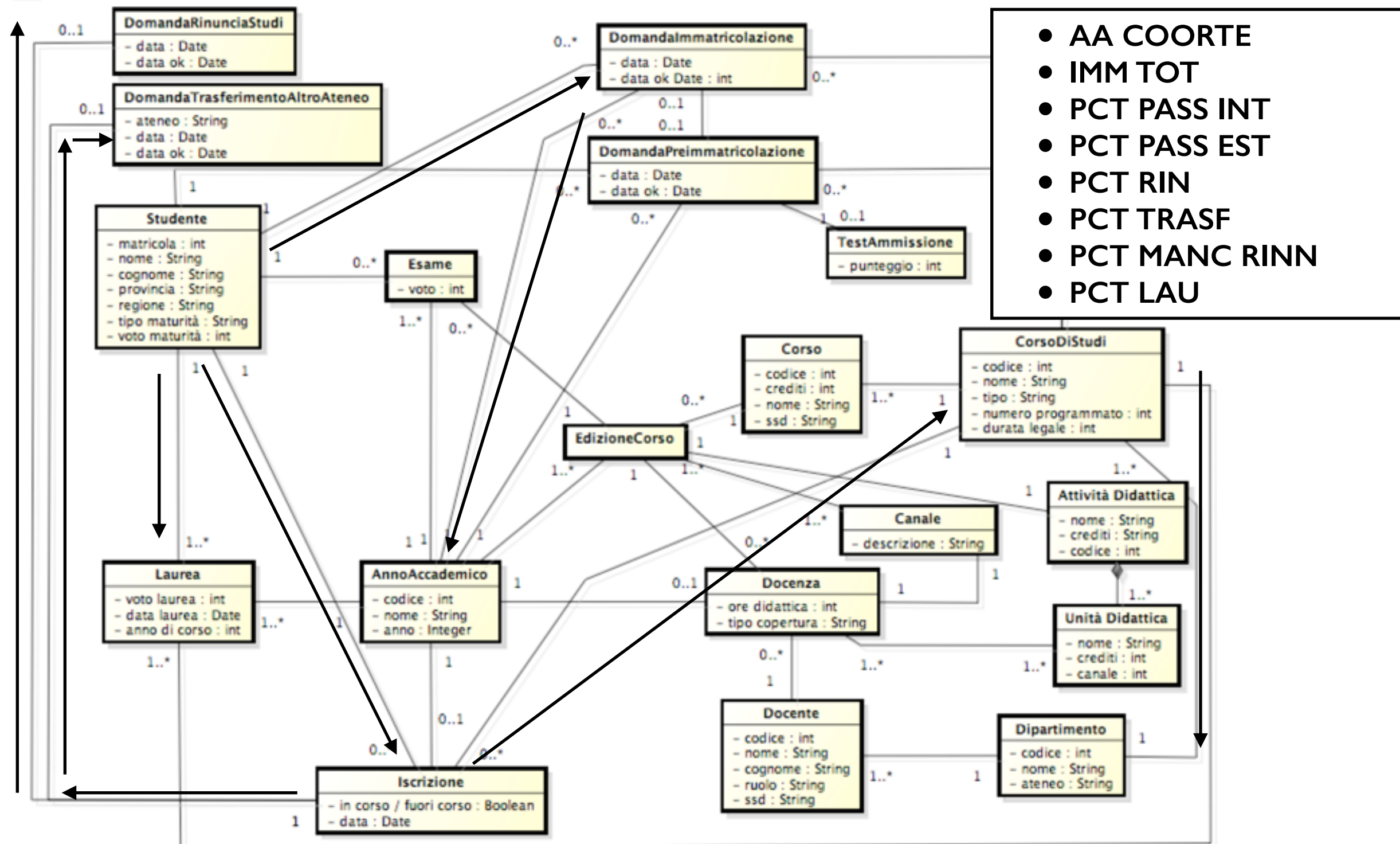


# Attrattività del Cds

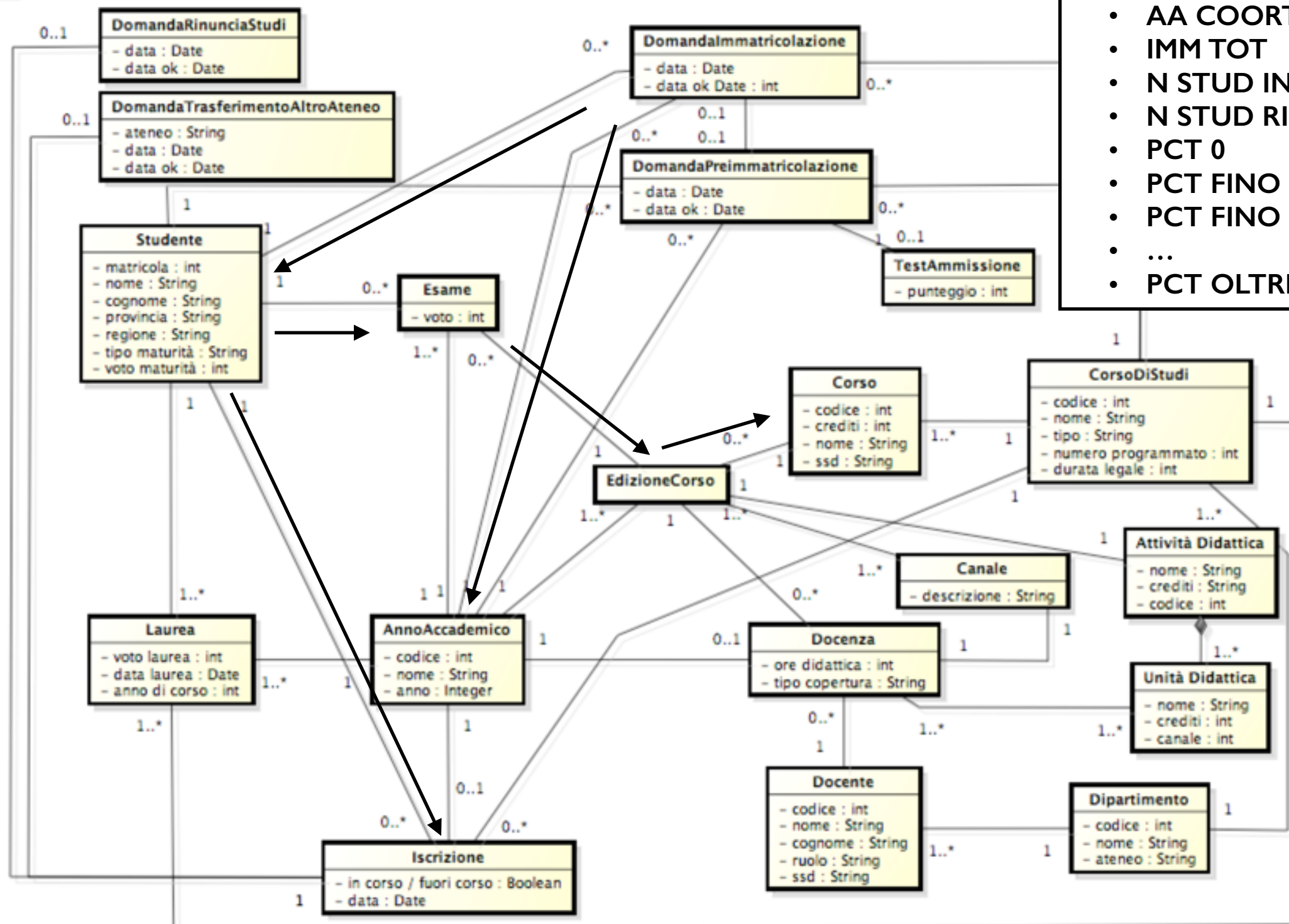


# RC.3

## Motivazioni dell'abbandono

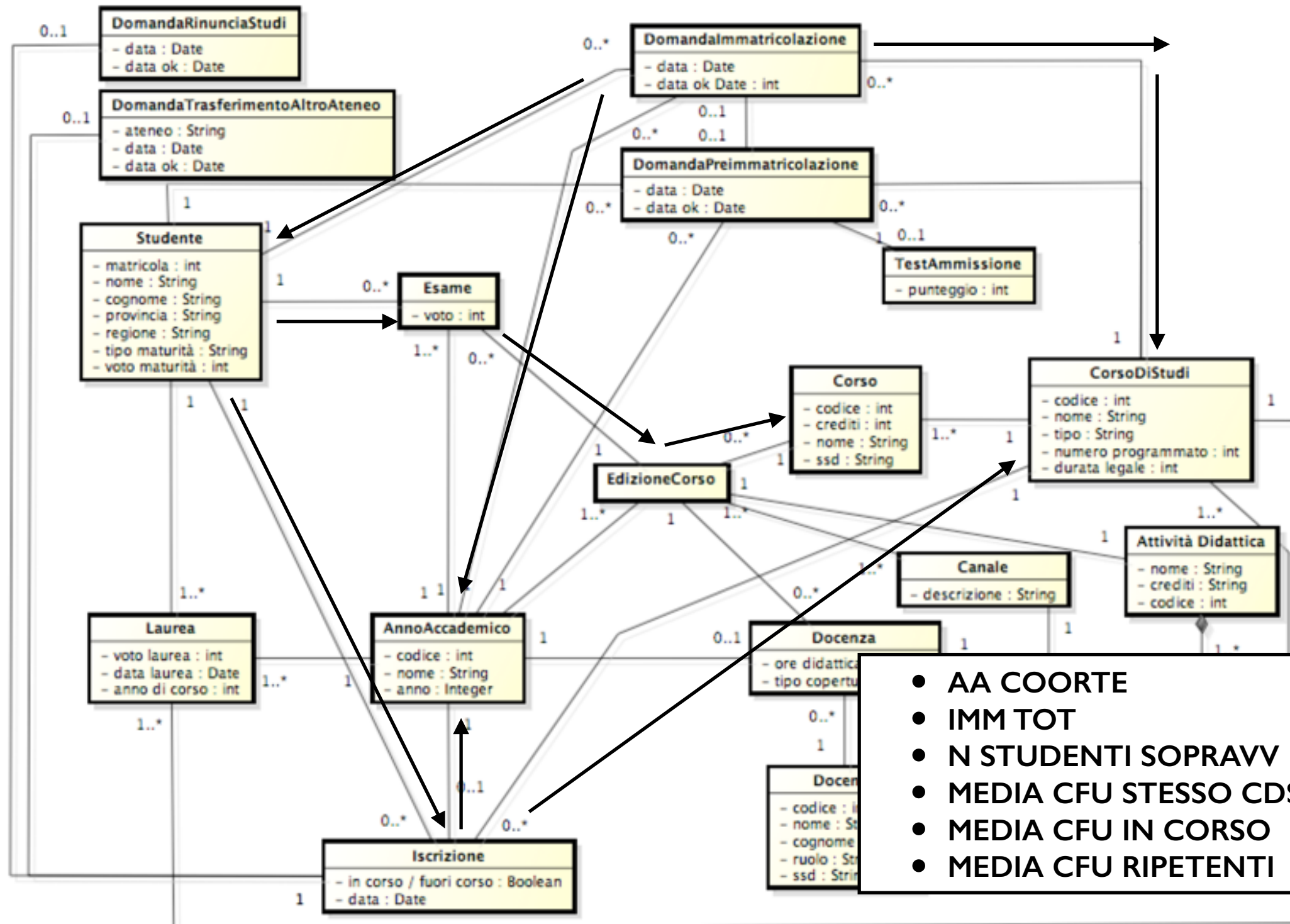


# RC.4.I Esiti del percorso formativo (range CFU)



- AA COORTE
- IMM TOT
- N STUD INCORSO
- N STUD RIP
- PCT 0
- PCT FINO 10
- PCT FINO 20
- ...
- PCT OLTRE 100

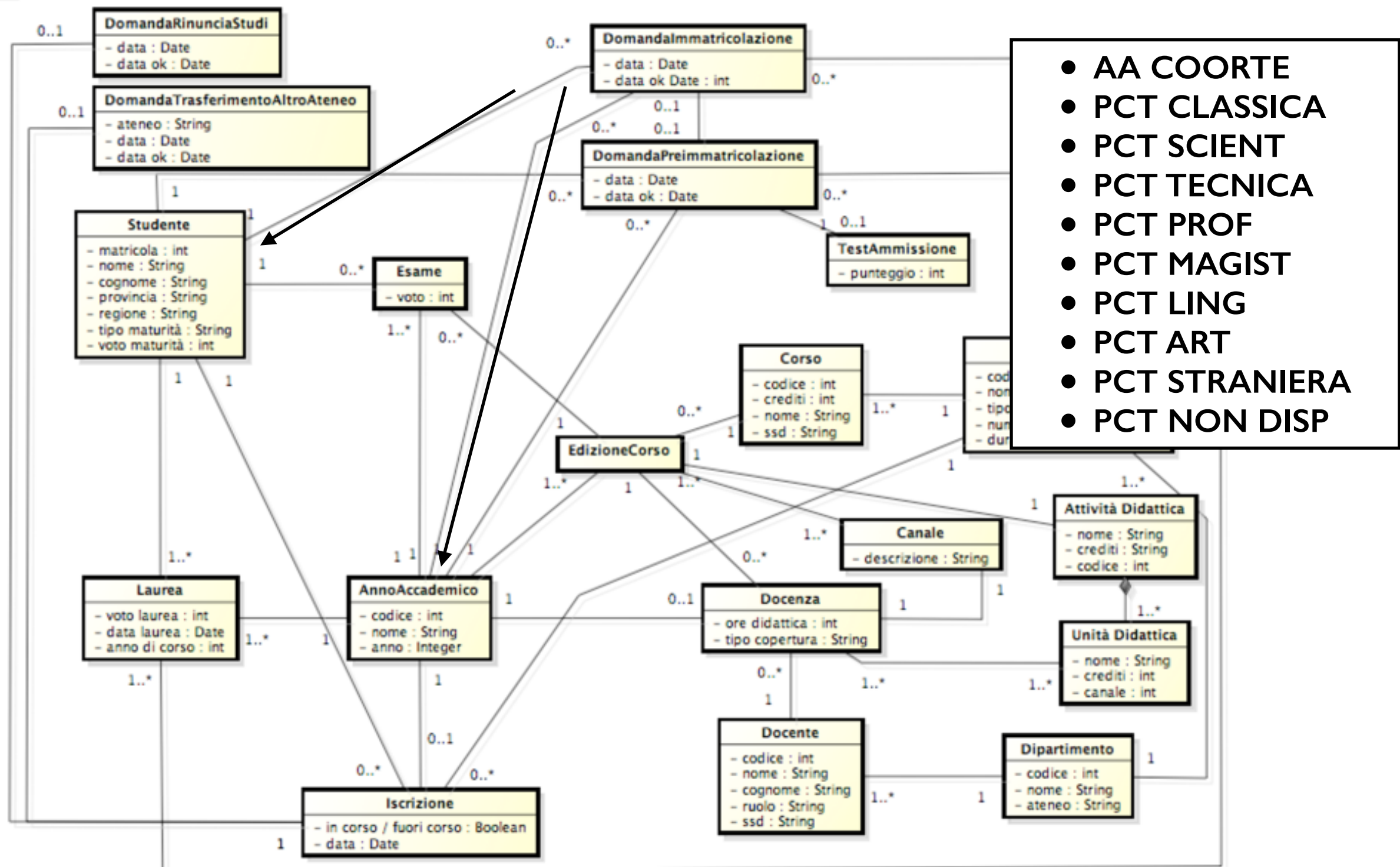
## RC.4.2 Esiti del percorso formativo (medie)



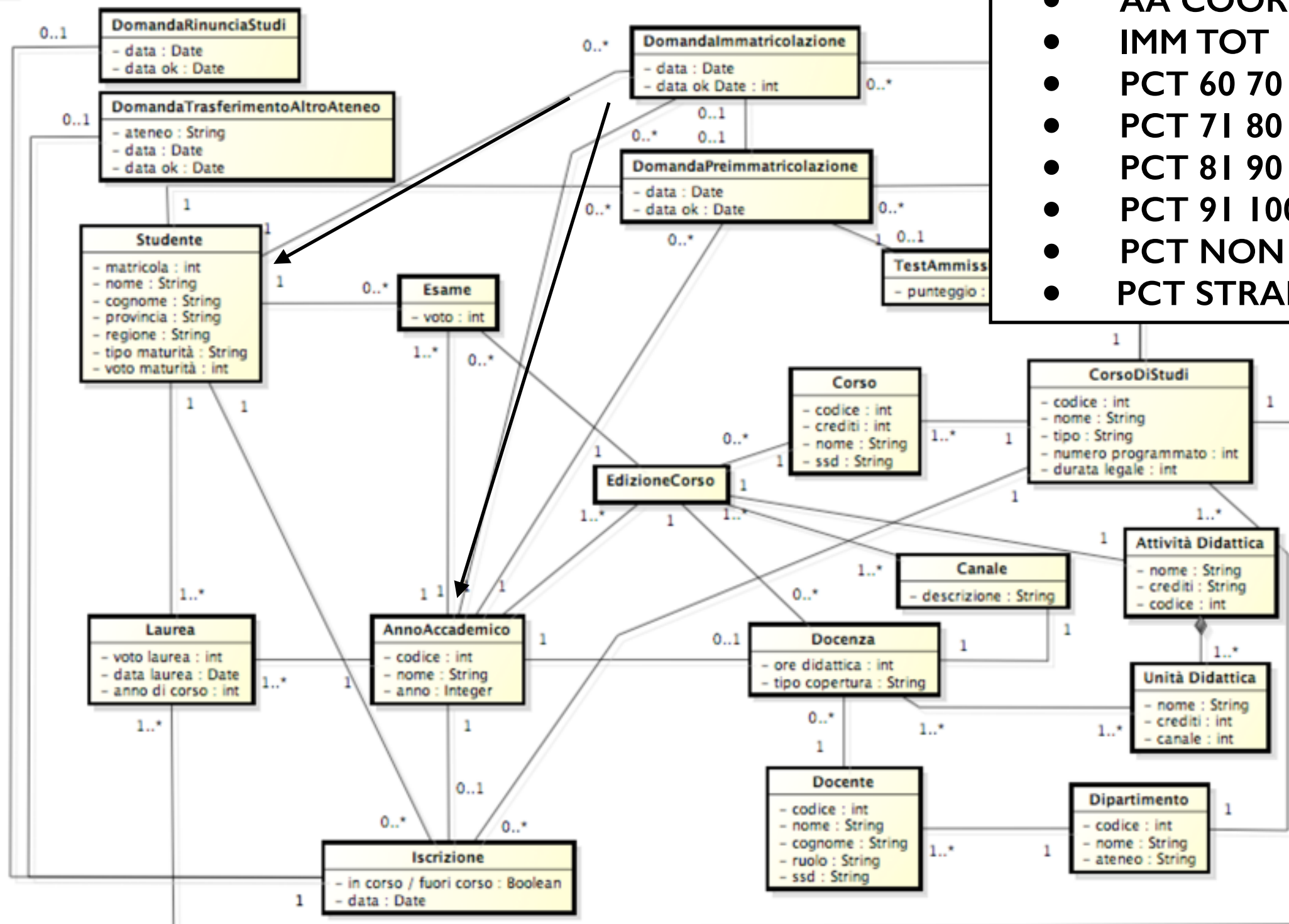
- AA COORTE
- IMM TOT
- N STUDENTI SOPRAVV
- MEDIA CFU STESSO CDS II ANNO
- MEDIA CFU IN CORSO
- MEDIA CFU RIPETENTI



# RC.5.I Tipo maturità

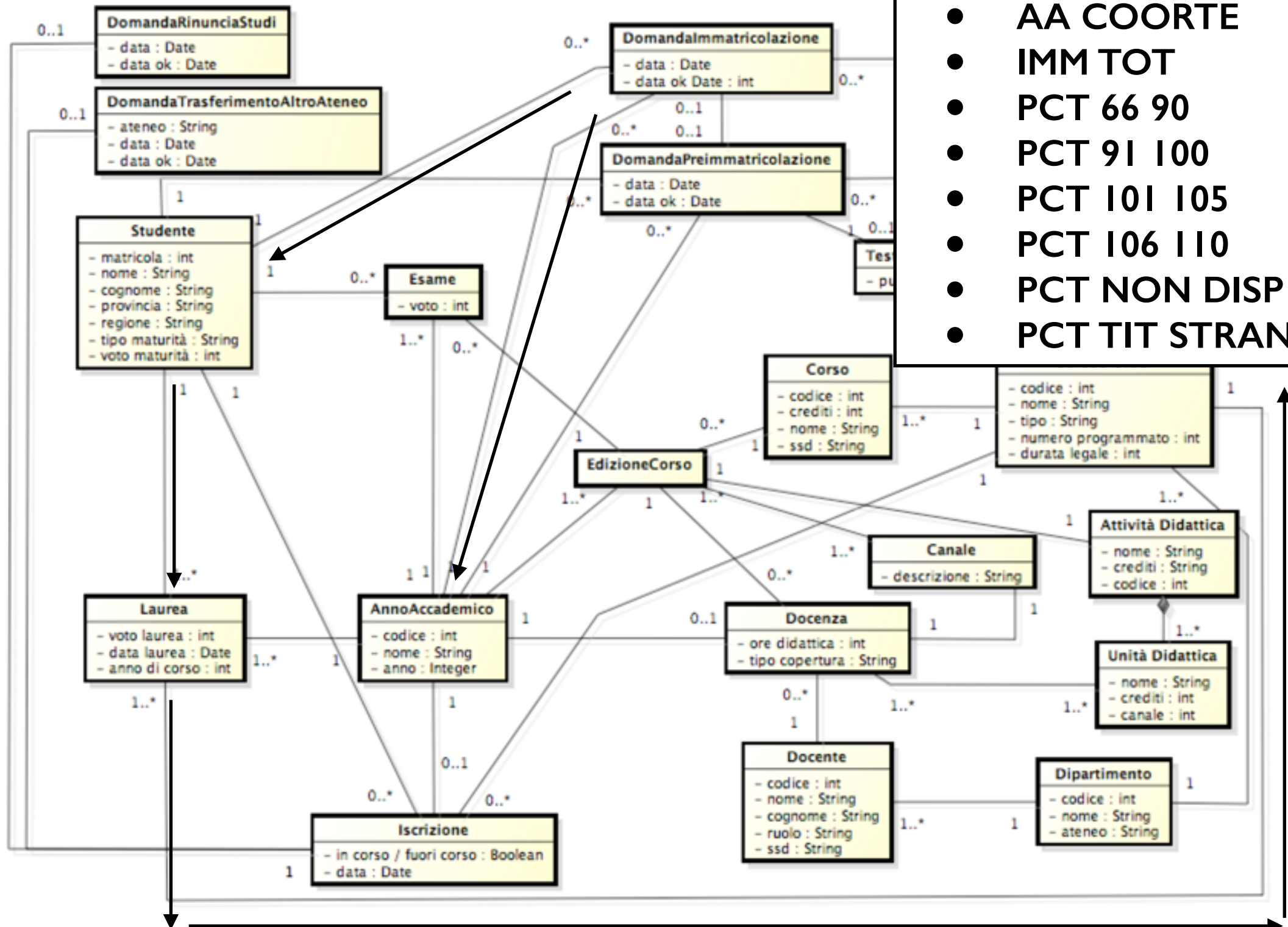


## RC.5.2 Voto maturità



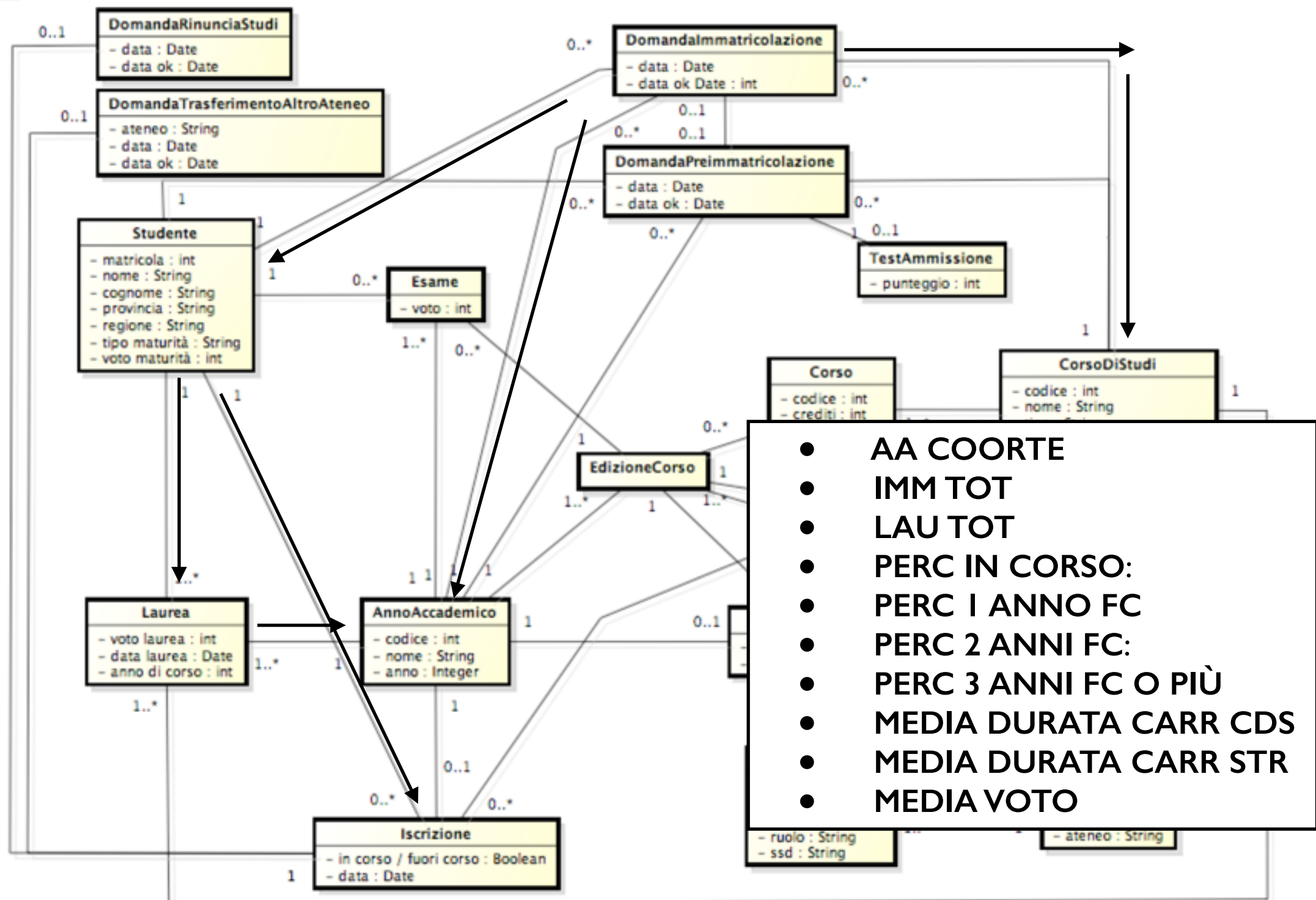
- AA COORTE
- IMM TOT
- PCT 60 70
- PCT 71 80
- PCT 81 90
- PCT 91 100
- PCT NON DISP
- PCT STRANIERA

# RC.6 Votazione riportata nel titolo di primo livello



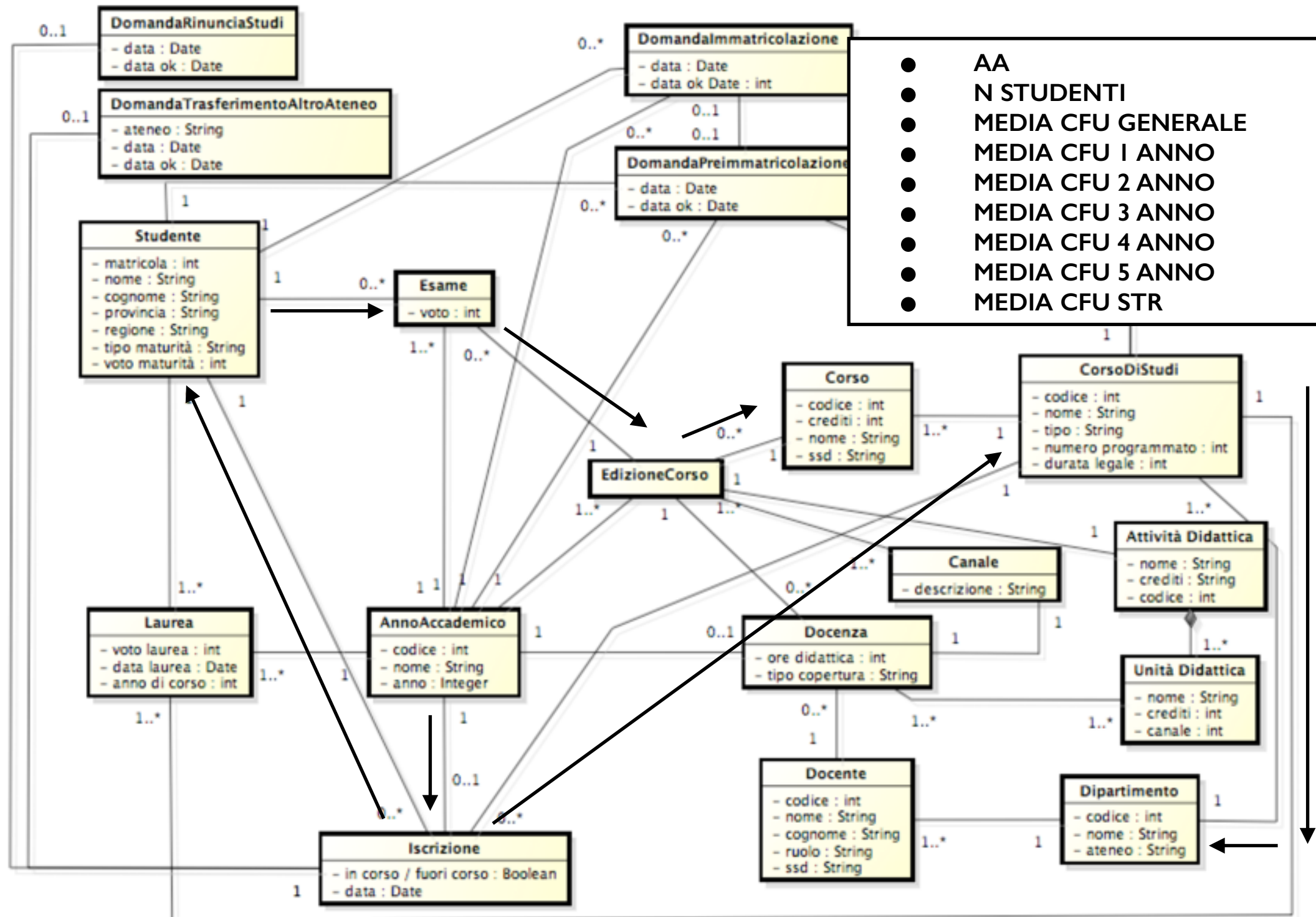
- AA COORTE
- IMM TOT
- PCT 66 90
- PCT 91 100
- PCT 101 105
- PCT 106 110
- PCT NON DISP
- PCT TIT STRANIERO

# RC.7 Dati sui laureati del Cds

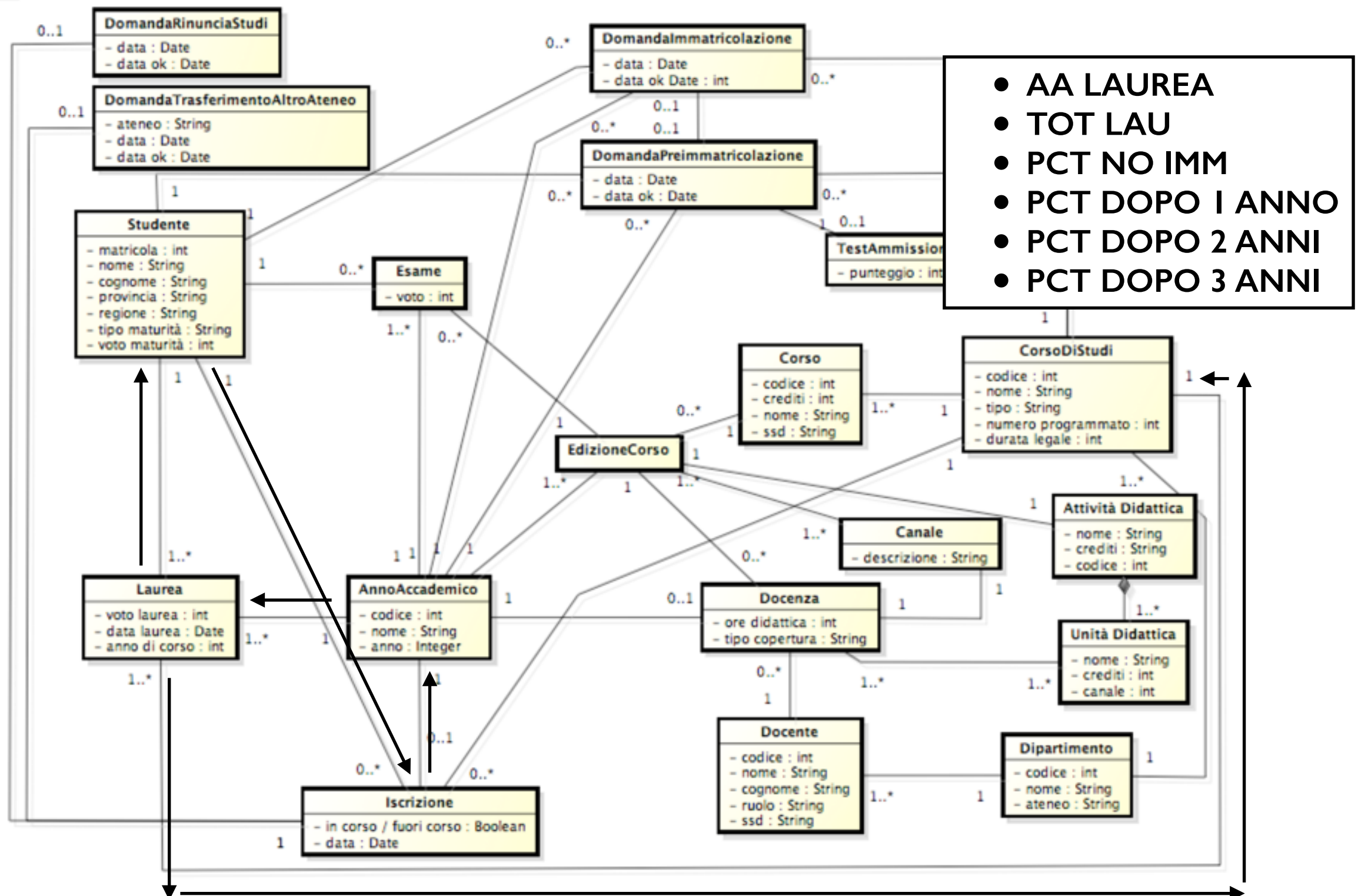




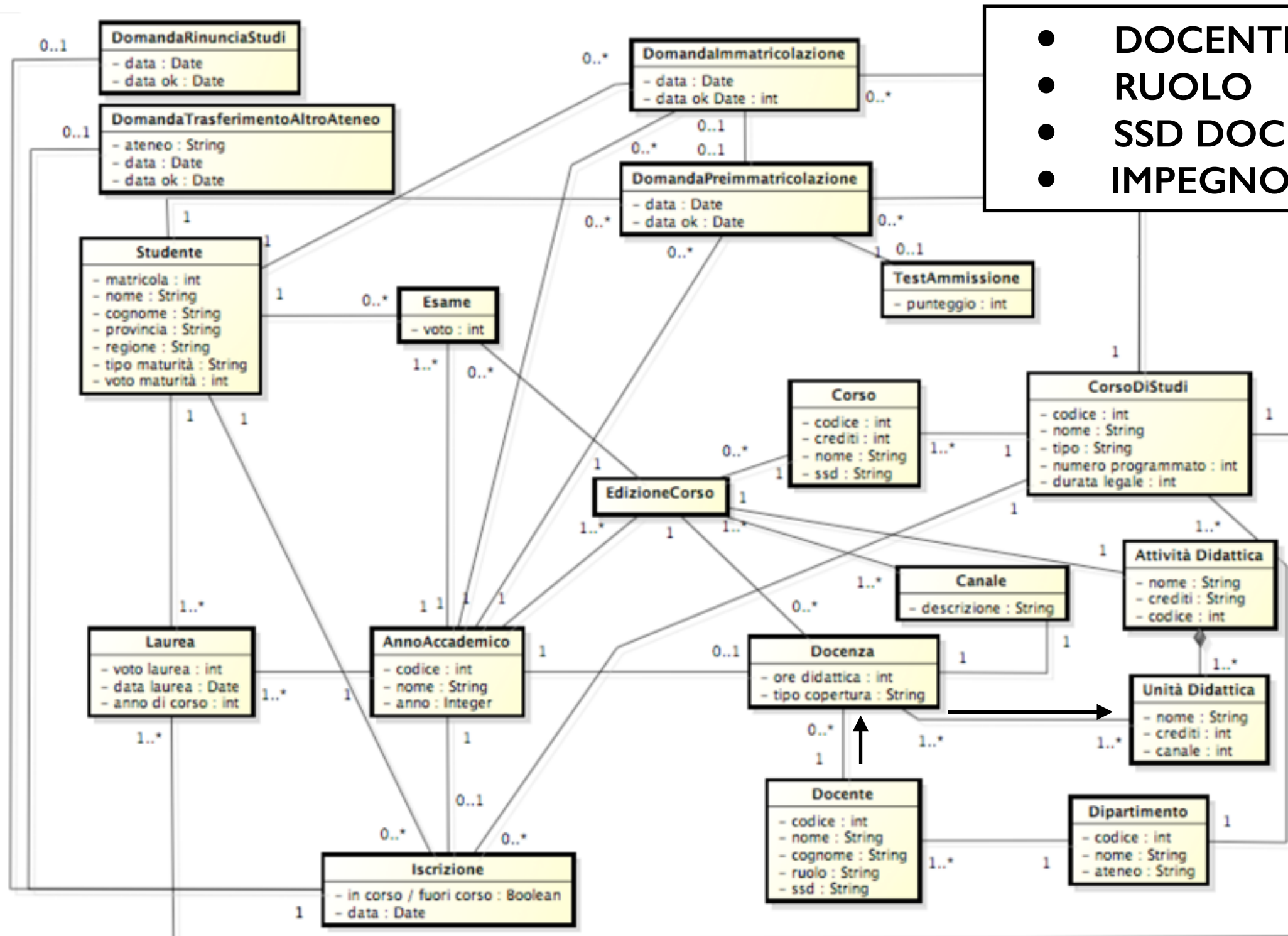
# RC.8 Numero medio di CFU conseguiti dagli iscritti per anno accademico



# RC.9 Propensione dei laureati triennali ad immatricolarsi ad un corso di laurea magistrale

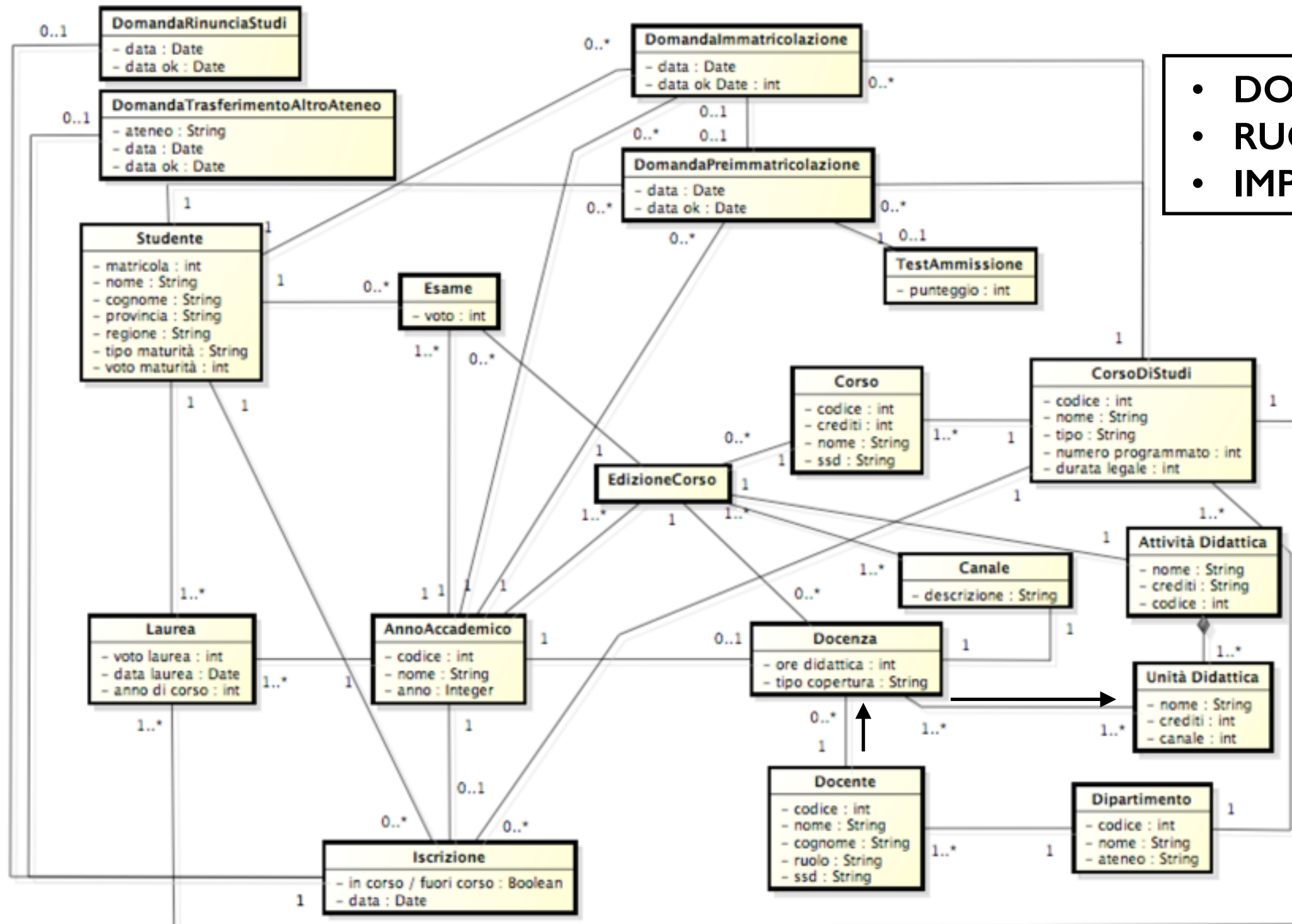


## RI.1.1 Docenti



- **DOCENTE**
- **RUOLO**
- **SSD DOCENTE**
- **IMPEGNO**

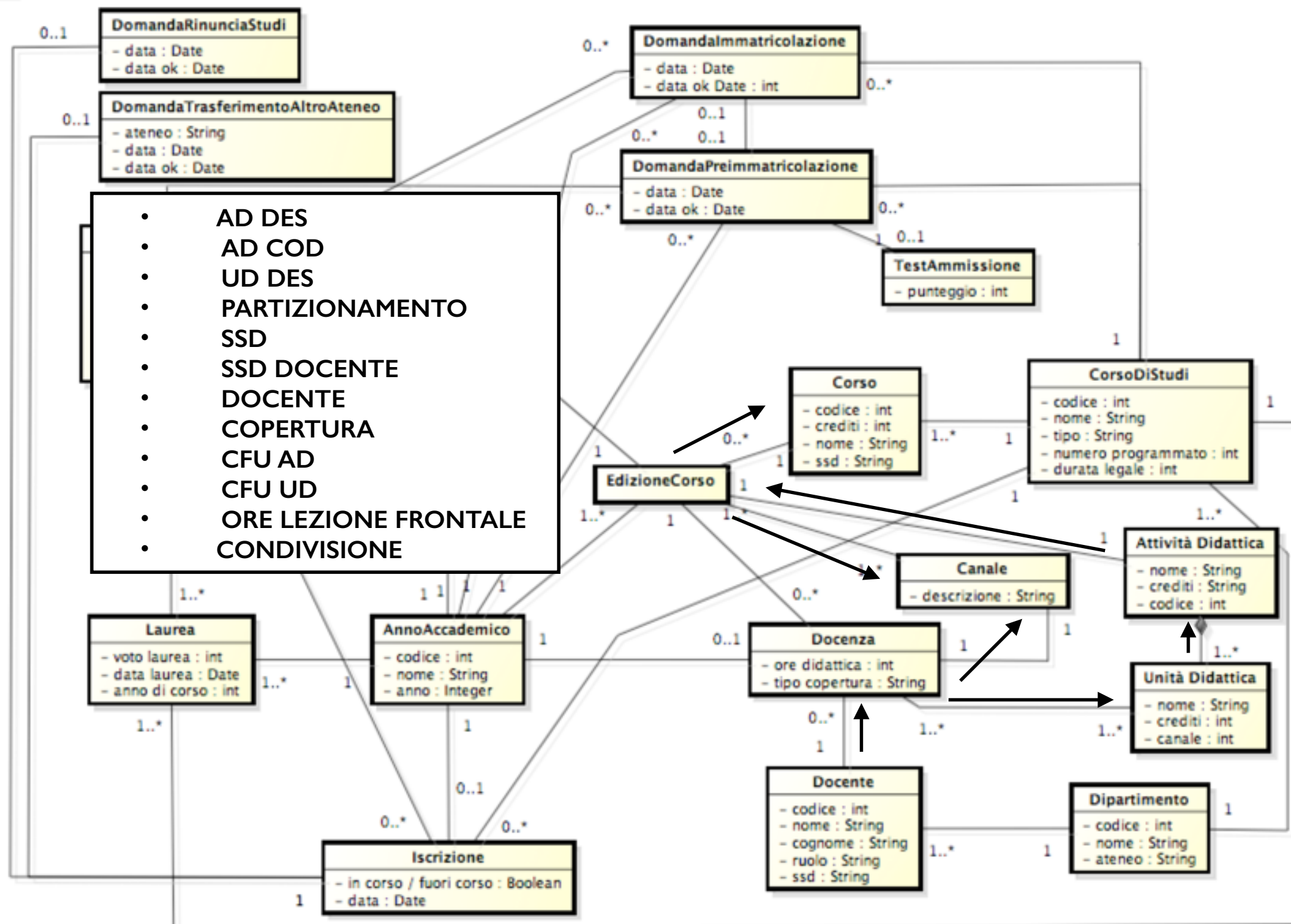
## RI.1.2 Docenti esterni



- **DOCENTE**
- **RUOLO**
- **IMPEGNO**



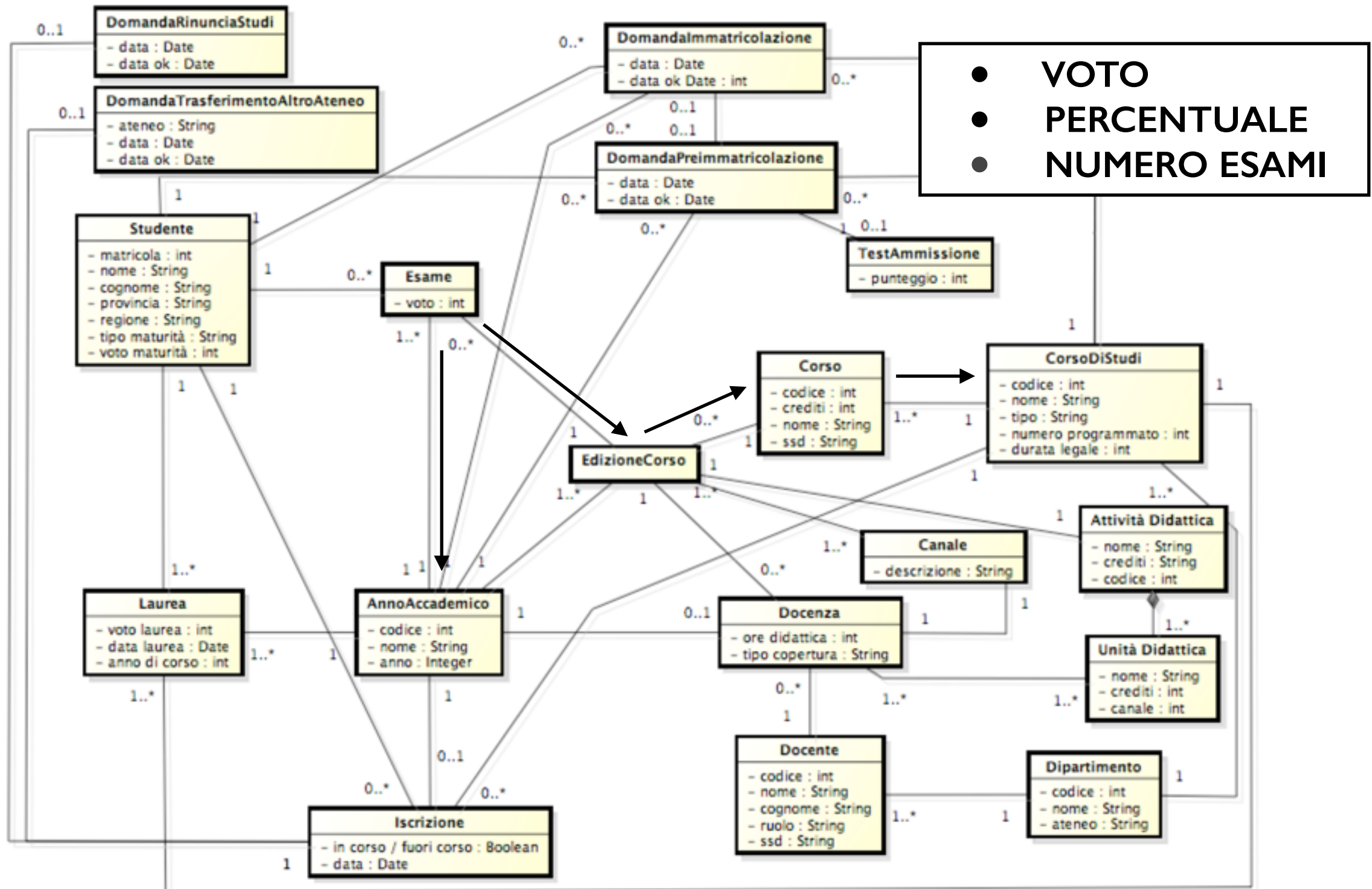
## RI.2 Insegnamenti e tipologie di copertura



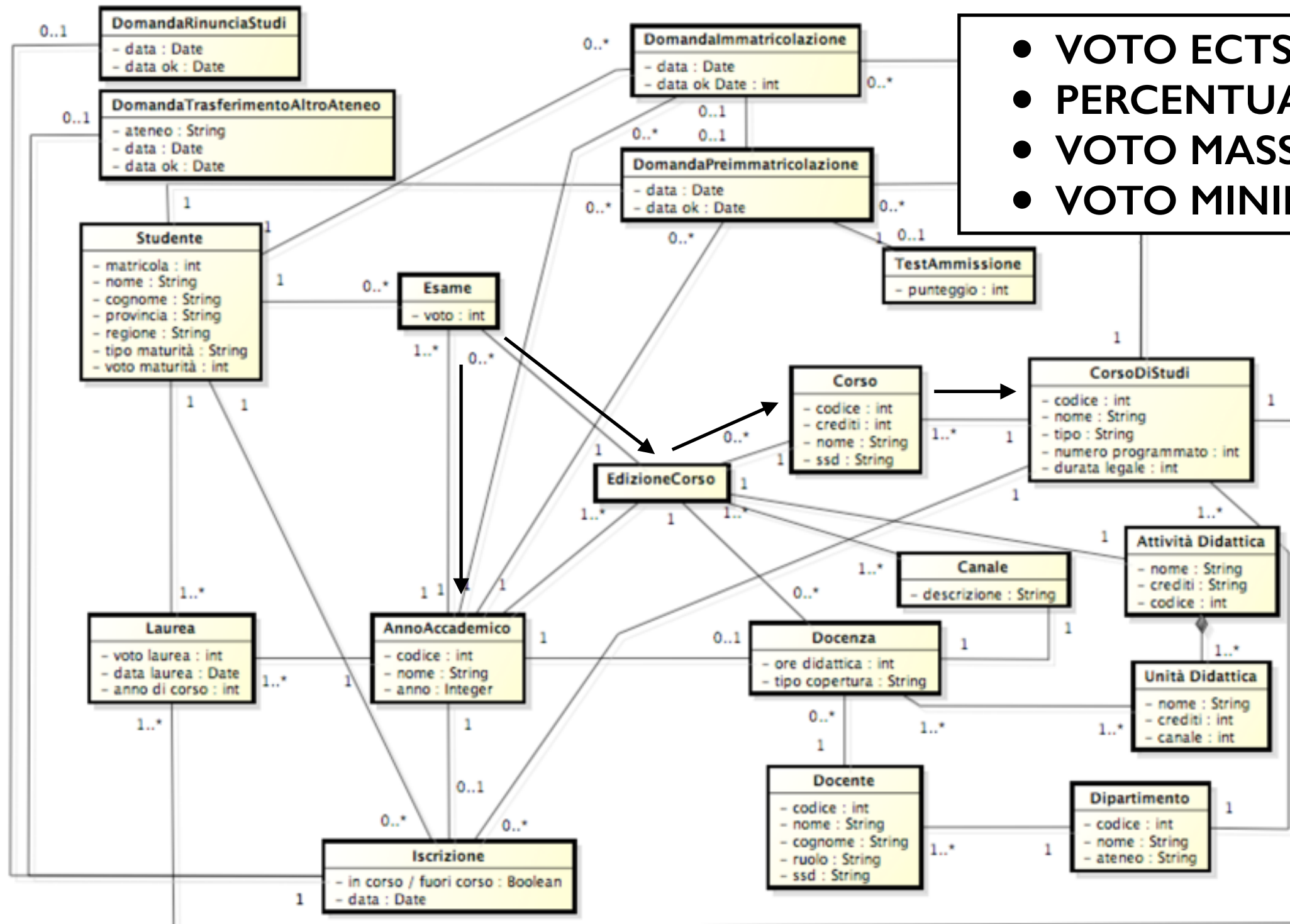
# RI.3 Esami di profitto relativi agli insegnamenti del Cds



## RI.4.1 Statistiche della distribuzione dei voti



## RI.4.2 Statistiche ETCS della distribuzione dei voti



- VOTO ECTS
- PERCENTUALE
- VOTO MASSIMO
- VOTO MINIMO



# Output della specifica dei requisiti

- Definizione di massima dei fatti dei data mart.
- Individuazione dei processi di business.
- Definizione di massima delle dimensioni.
- Definizione di massima della granularità di ciascun fatto.
- Scelta delle misure.

# Scelta dei processi

- **Data mart carriere studenti**
  - Immatricolazioni e test di ammissione
  - Iscrizioni
  - Trasferimenti e rinunce
  - Esami
  - Lauree
- **Data mart docenze e insegnamenti**
  - Docenze

# Matrice dell'architettura a bus del DW

	Tempo	Anno Accademico	Studente	Docente	Corso	Corso di Studi
Immatricolazioni	X	X	X			X
Test ammissione	X	X	X			X
Iscrizioni	X	X	X			X
Trasferimenti	X	X	X			X
Rinunce	X	X	X			X
Esami	X	X	X		X	X
Lauree	X	X	X			X
Docenze		X		X	X	X

# Crediti e riferimenti

- Matteo Golfarelli, Stefano Rizzi - Data Warehouse - Teoria e pratica della progettazione (McGraw Hill), 2006
- R. Kimball et al. - The Data Warehouse Lifecycle Toolkit Second Edition (Wiley & Sons), 2008
- <http://caccio.bimodeler.com/archives/337>