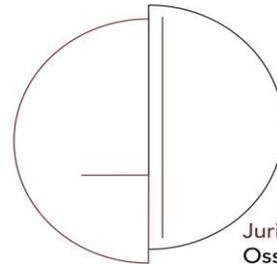


Convegno «*Ricerca, educazione e accesso al patrimonio culturale. Un confronto tra diritti fondamentali ed eccezioni al diritto d'autore*» - 6 maggio 2022, Consiglio Nazionale delle Ricerche - Aula Marconi (Piazzale Aldo Moro n. 7) Roma

L'eccezione di Text and Data Mining e le nozioni legali di "informazioni" e "dati" nella legislazione del mercato unico digitale



SAPIENZA
UNIVERSITÀ DI ROMA



Juridical Observatory on Digital Innovation
Osservatorio Giuridico sulla Innovazione Digitale

Prof. Salvatore Orlando - salvatore.orlando@uniroma1.it
Ordinario di Diritto Privato - Sapienza Università di Roma
Direttore dell'Osservatorio Giuridico sull'Innovazione Digitale (OGID)
<https://web.uniroma1.it/deap/ogid>

- 1) I punti fermi e gli equivoci linguistici**
- 2) Il sillogismo alla base del sistema dell'eccezione e il concetto di riproduzione**
- 3) I dati ed il processo di regolazione orientato alle finalità del loro uso**
- 4) L'idoneità del TDM ad essere un moltiplicatore di proprietà intellettuale**
- 5) Conclusioni**

1) I punti fermi e gli equivoci linguistici

2) Il sillogismo alla base del sistema dell'eccezione e il concetto di riproduzione

3) I dati ed il processo di regolazione orientato alle finalità del loro uso

4) L'idoneità del TDM ad essere un moltiplicatore di proprietà intellettuale

5) Conclusioni

Art. 7 Dir. 96/9/CE

2. For the purposes of this Chapter:

(a) **'extraction'** shall mean the permanent or temporary **transfer of all or a substantial part of the contents of a database to another medium** by any means or in any form;

[ITA: *il **trasferimento** permanente o temporaneo della **totalità o di una parte sostanziale del contenuto** di una banca dati **su un altro supporto***]

Article 2 Dir. (UE) 2019/790

Definitions

1. For the purposes of this Directive, the following definitions apply: [...]

(2) **'text and data mining'** means any **automated analytical technique aimed at analysing text and data in digital form in order to generate information** which includes but is not limited to patterns, trends and correlations;

[ITA: «*estrazione di testo e di dati*» (*text and data mining*): *qualsiasi tecnica di **analisi automatizzata** volta ad analizzare testi e **dati in formato digitale** avente lo scopo di **generare informazioni** inclusi, a titolo non esaustivo, modelli, tendenze e correlazioni*]

Articolo 2 Dir. (UE) 2019/790

Definizioni

[...]

2) «estrazione di testo e di dati» (*text and data mining*): qualsiasi tecnica di analisi automatizzata volta ad analizzare testi e dati in formato digitale avente lo scopo di generare informazioni inclusi, a titolo non esaustivo, modelli, tendenze e correlazioni;

Articolo 70-ter l.a.

[...]

2. Ai fini della presente legge per estrazione di testo e di dati si intende qualsiasi tecnica automatizzata volta ad analizzare grandi quantità di testi, suoni, immagini, dati o metadati in formato digitale con lo scopo di generare informazioni, inclusi modelli, tendenze e correlazioni[...]

Articolo 3 dir.(UE) 2019/790

Estrazione di testo e di dati per scopi di ricerca scientifica

1. Gli Stati membri dispongono un'eccezione ai diritti di cui all'articolo 5, lettera a), e all'articolo 7, paragrafo 1, della direttiva 96/9/CE [*banche dati*], all'articolo 2 della direttiva 2001/29/CE [*InfoSoc*], e all'articolo 15, paragrafo 1, della presente direttiva per le **riproduzioni e le estrazioni** effettuate da organismi di ricerca e istituti di tutela del patrimonio culturale ai fini dell'estrazione, per scopi di ricerca scientifica, di testo e di dati da opere o altri materiali cui essi hanno legalmente accesso.

[...]

Articolo 70-ter I.a.

1. Sono consentite le **riproduzioni** compiute da organismi di ricerca e da istituti di tutela del patrimonio culturale, per scopi di ricerca scientifica, ai fini dell'estrazione di testo e di dati da opere o da altri materiali disponibili in reti o banche di dati cui essi hanno lecitamente accesso, nonché la comunicazione al pubblico degli esiti della ricerca ove espressi in nuove opere originali.

Articolo 4 dir.(UE) 2019/790

Eccezioni o limitazioni ai fini dell'estrazione di testo e di dati

1. Gli Stati membri dispongono un'eccezione o una limitazione ai diritti di cui all'articolo 5, lettera a), e all'articolo 7, paragrafo 1, della direttiva 96/9/CE, all'articolo 2 della direttiva 2001/29/CE, all'articolo 4, paragrafo 1, lettere a) e b), della direttiva 2009/24/CE [*programmi per elaboratore*] e all'articolo 15, paragrafo 1, della presente direttiva per **le riproduzioni e le estrazioni** effettuate da opere o altri materiali cui si abbia legalmente accesso ai fini dell'estrazione di testo e di dati.

Articolo 70-quater I.a.

1. Fermo restando quanto previsto dall'articolo 70-ter, sono consentite **le riproduzioni e le estrazioni** da opere o da altri materiali contenuti in reti o in banche di dati cui si ha legittimamente accesso ai fini dell'estrazione di testo e di dati. L'estrazione di testo e di dati è consentita quando l'utilizzo delle opere e degli altri materiali non è stato espressamente riservato dai titolari del diritto d'autore e dei diritti connessi nonché dai titolari delle banche dati.

Articolo 3 dir.(UE) 2019/790

Estrazione di testo e di dati per scopi di ricerca scientifica

1. Gli Stati membri dispongono un'eccezione ai diritti di cui all'articolo 5, lettera a) **RIPRODUZIONE**, e all'articolo 7, paragrafo 1, della direttiva 96/9/CE [*banche dati*] **ESTRAZIONE**, all'articolo 2 della direttiva 2001/29/CE [*InfoSoc*] **RIPRODUZIONE**, e all'articolo 15, paragrafo 1, della presente direttiva [*pubblicazioni giornalistiche-utilizzo online*] **RIPRODUZIONE per le riproduzioni e le estrazioni** effettuate da organismi di ricerca e istituti di tutela del patrimonio culturale ai fini dell'estrazione, per scopi di ricerca scientifica, di testo e di dati da opere o altri materiali cui essi hanno legalmente accesso.
[...]

Articolo 4 dir.(UE) 2019/790

Eccezioni o limitazioni ai fini dell'estrazione di testo e di dati

1. Gli Stati membri dispongono un'eccezione o una limitazione ai diritti di cui all'articolo 5, lettera a) **RIPRODUZIONE**, e all'articolo 7, paragrafo 1, della direttiva 96/9/CE [*banche dati*] **ESTRAZIONE**, all'articolo 2 della direttiva 2001/29/CE [*InfoSoc*] **RIPRODUZIONE**, all'articolo 4, paragrafo 1, lettere a) e b), della direttiva 2009/24/CE [*software*] **RIPRODUZIONE** e all'articolo 15, paragrafo 1, della presente direttiva [*pubblicazioni giornalistiche-utilizzo online*] **RIPRODUZIONE per le riproduzioni e le estrazioni** effettuate da opere o altri materiali cui si abbia legalmente accesso ai fini dell'estrazione di testo e di dati.

[CAPITOLO III

DIRITTO «SUI GENERIS»]

Articolo 7 dir. 96/9/CE

Oggetto della tutela

1. Gli Stati membri attribuiscono al costituente di una banca di dati il diritto di vietare operazioni di **estrazione*** e/o reimpiego **della totalità o di una parte sostanziale del contenuto** della stessa, valutata in termini qualitativi o quantitativi, qualora il conseguimento, la verifica o la presentazione di tale contenuto attestino un investimento rilevante sotto il profilo qualitativo o quantitativo.

* = il **trasferimento permanente o temporaneo della totalità o di una parte sostanziale del contenuto di una banca dati su un altro supporto**

Prima regola inesistente che rischia di essere creata attraverso l'eccezione, se non si interpreta bene la nuova disciplina:

che esista un generale diritto esclusivo di 'estrazione di dati' da opere e altri materiali protetti

Articolo 3 dir.(UE) 2019/790

Estrazione di testo e di dati per scopi di ricerca scientifica

1. Gli Stati membri dispongono un'eccezione ai diritti di cui all'articolo 5, lettera a), e all'articolo 7, paragrafo 1, della direttiva 96/9/CE [*banche dati*], all'articolo 2 della direttiva 2001/29/CE [*InfoSoc*], e all'articolo 15, paragrafo 1, della presente direttiva [*pubblicazioni giornalistiche-utilizzo online*] per le **riproduzioni e le estrazioni** effettuate da organismi di ricerca e istituti di tutela del patrimonio culturale ai fini dell'estrazione, per scopi di ricerca scientifica, di testo e di dati da opere o altri materiali cui essi hanno legalmente accesso.

Articolo 70-ter I.a.

1. Sono consentite le **riproduzioni** compiute da organismi di ricerca e da istituti di tutela del patrimonio culturale, per scopi di ricerca scientifica, ai fini dell'estrazione di testo e di dati da opere o da altri materiali disponibili in reti o banche di dati cui essi hanno lecitamente accesso, **nonché la comunicazione al pubblico degli esiti della ricerca ove espressi in nuove opere originali.**

Seconda regola inesistente che rischia di essere creata attraverso l'eccezione, se non si interpreta bene la nuova disciplina:

che esista un generale divieto di comunicazione al pubblico degli esiti delle ricerche compiute in esito all'applicazione di tecniche di TDM (e che tale divieto non si applichi soltanto agli esiti delle ricerche espressi in nuove opere originali).

- 1) I punti fermi e gli equivoci linguistici
- 2) Il sillogismo alla base del sistema dell'eccezione e il concetto di riproduzione**
- 3) I dati ed il processo di regolazione orientato alle finalità del loro uso
- 4) L'idoneità del TDM ad essere un moltiplicatore di proprietà intellettuale
- 5) Conclusioni

Il sillogismo riflesso negli artt. 3 e 4 Direttiva (UE) 2019/790 e 70-ter e 70-quater l.a.

premessa maggiore

La riproduzione e l'estrazione da opere e altri materiali sono **attività riservate ex lege**, ossia formano oggetto di diritti esclusivi (dunque per svolgere tali attività o esiste una eccezione di legge oppure c'è bisogno di una licenza)

premessa minore

Le **attività tecniche** poste in essere ai fini delle operazioni di TDM sono attività di riproduzione o di estrazione **riservate ex lege**

conclusione

Per consentire le attività tecniche poste in essere ai fini delle operazioni di TDM c'è bisogno o di una eccezione di legge o di una licenza.

Il sillogismo

riflesso negli artt. 3 e 4 Direttiva (UE) 2019/790 e 70-ter e 70-quater l.a.

premessa minore

SI RITIENE CHE

le **attività tecniche** poste in essere ai fini delle operazioni di TDM **siano** attività di riproduzione o di estrazione **riservate ex lege**

«... the argument according to which, the introduction of the TDM exceptions in the DSM dir. implies that the technical acts needed for mining some content are falling under the reproduction right can be challenged. [...] Even if the exception can be set aside by an opt-out, this does not entail that all technical copies made in the TDM process are to be considered as reproductions: the opt-out is meant to reserve the right, where the right is applicable, i.e. where a “reproduction” can be found (in particular determining whether the work is recognizable and thus exploited in the course of the TDM process or afterwards)[...]»*

Along this reasoning, a developer or user of an AI solution could argue that the technical copies made of protected cultural content in the course of a TDM process (while training an AI solution) are not sufficient for finding an infringement of the reproduction right ...»

Commissione Europea, Study on copyright and new technologies: copyright data management and artificial intelligence, Luxembourg (Publications Office of the European Union), febbraio 2022, p. 215

*** CGUE – sentenza Pelham 29.7.2019 C-476/17, Pelham GmbH et aliis / Ralf Hütter e Florian Schneider-Esleben**

Terza regola inesistente che rischia di essere creata attraverso l'eccezione, se non si interpreta bene la nuova disciplina:

che qualunque attività tecnica di 'riproduzione' strumentale al TDM è assoggettata al regime delle eccezioni e limitazioni in quanto attività riservata *ex lege*

Cfr. Considerando (9) dir. (UE) 2019/790



Artificial Intelligence and IP: copyright and patents

A consultation on AI and IP co...

www.gov.uk

<https://www.gov.uk/government>

We are now seeking to understand the issues in more detail and are consulting on three specific areas:

1. Copyright protection for computer-generated works without a human author. These are currently protected in the UK for 50 years. But should they be protected at all and if so, how should they be protected?
2. Licensing or exceptions to copyright for text and data mining, **which is often significant in AI use and development.**
3. Patent protection for AI-devised inventions. Should we protect them, and if so, how should they be protected?

- 1) I punti fermi e gli equivoci linguistici
- 2) Il sillogismo alla base del sistema dell'eccezione e il concetto di riproduzione
- 3) I dati ed il processo di regolazione orientato alle finalità del loro uso**
- 4) L'idoneità del TDM ad essere un moltiplicatore di proprietà intellettuale
- 5) Conclusioni

Table 2: Overview of the EU digital acts and proposals by policy domains

ICT services, infrastructure, networks	Trust and security	Consumer protection and competition	Online services and e-commerce	Data protection and governance	Copyrights and audiovisuals	e-Government
Regulation on High Performance Joint Undertaking (EU) 2021/1173 ¹	EU Cybersecurity Act Regulation (EU) 2019/881 ²	Regulation on platform-to-business relations (P2B Regulation) (EU) 2019/1150 ³	EU VAT for e-commerce package (EU VOEC) ⁴	Open Data Directive (EU) 2019/1024 ⁵	Directive on Copyrights and related rights in the DSM Directive EU (2019) 790 ⁶	E-invoicing Directive (EU) 2014/55 ⁷
European electronic communications Code Directive (EU) 2018/1972 ⁸	Network and Information Security (NIS) Directive EU 2016/1148 ⁹ [under revision]	Roaming Regulation (EU) 2017/920 ¹⁰ [under revision]	Payment Services Directive (EU) 2015/2366 ¹¹	General Data Protection Regulation (GDPR) (EU) 2016/679 ¹²	Satellite and Cable II Directive (EU) 2019/789 ¹³	Electronic identification rules - eIDAS Regulation (EU) No 910/2014 ¹⁴ [under revision]
Broadband Cost Reduction Directive (EU) 2014/61 ¹¹⁵ [under revision]	ePrivacy Directive 2002/58/EC	Measures concerning open internet access Regulation (EU) 2015/2120 ¹⁷	e-Commerce Directive 2000/31/EC ¹⁸	Regulation on the free flow of non-personal data (EU) 2018/1807 ¹⁹	Portability Regulation (EU) 2017/1128 ²⁰	Digitalisation of visa procedures ²¹
Regulation establishing the European Cybersecurity Competence Centre and Network 2021/887 ²²	Legislative proposal for an Artificial Intelligence Act COM 2021/206 ¹⁶	Proposal for a Digital Markets Act COM/2020/842 ²⁴	Geo-Blocking Regulation (EU) 2018/302 ²⁵	Directive on certain aspects concerning contracts for the supply of digital content and digital services 2019/770 ²⁶	Radio Equipment Directive 2014/53/EU ²⁷	
Proposal on a pilot regime for market infrastructures based on distributed ledger technology ²⁸	Proposal for a Regulation on a high common level of cybersecurity Directive COM(2020) 823 ²³	Proposal for a Regulation on Markets in Crypto-assets COM/2020/593 ³⁰	Proposal for a Digital Services Act COM/2020/825 ² ³¹	Data Governance Act (DGA) COM/2020/767 ³²	Database Directive 96/9/EC ³³ [under revision]	
Legislative proposal on building an EU space-based global secure communication system ³⁴	Proposal for Regulation on Privacy and Electronic Communications ²⁹ [revision of ePrivacy Directive]	Common charges for mobile phones and similar devices ³⁶	Legislative proposal on a Single Market Emergency Instrument ³⁷	European Data Act ³⁸		
European Chips Act ³⁹	Proposal on a Regulation on Machinery products COM/2021/202 ³⁵	New design requirements and consumer rights for electronics ⁴¹		European Health Data Space ⁴²		
Declaration of the Industrial Alliance for Processors and Semiconductor Technologies ⁴⁴	Proposal for a Regulation on digital operational resilience for the financial sector ⁴⁰	Multimodal digital mobility services ⁴⁶		European Alliance for Industrial Data, Edge and Cloud ⁴³		
EU-US Trade and Technology Council (TTC) ⁴⁷	European Cyber Resilience Act ⁴⁵					

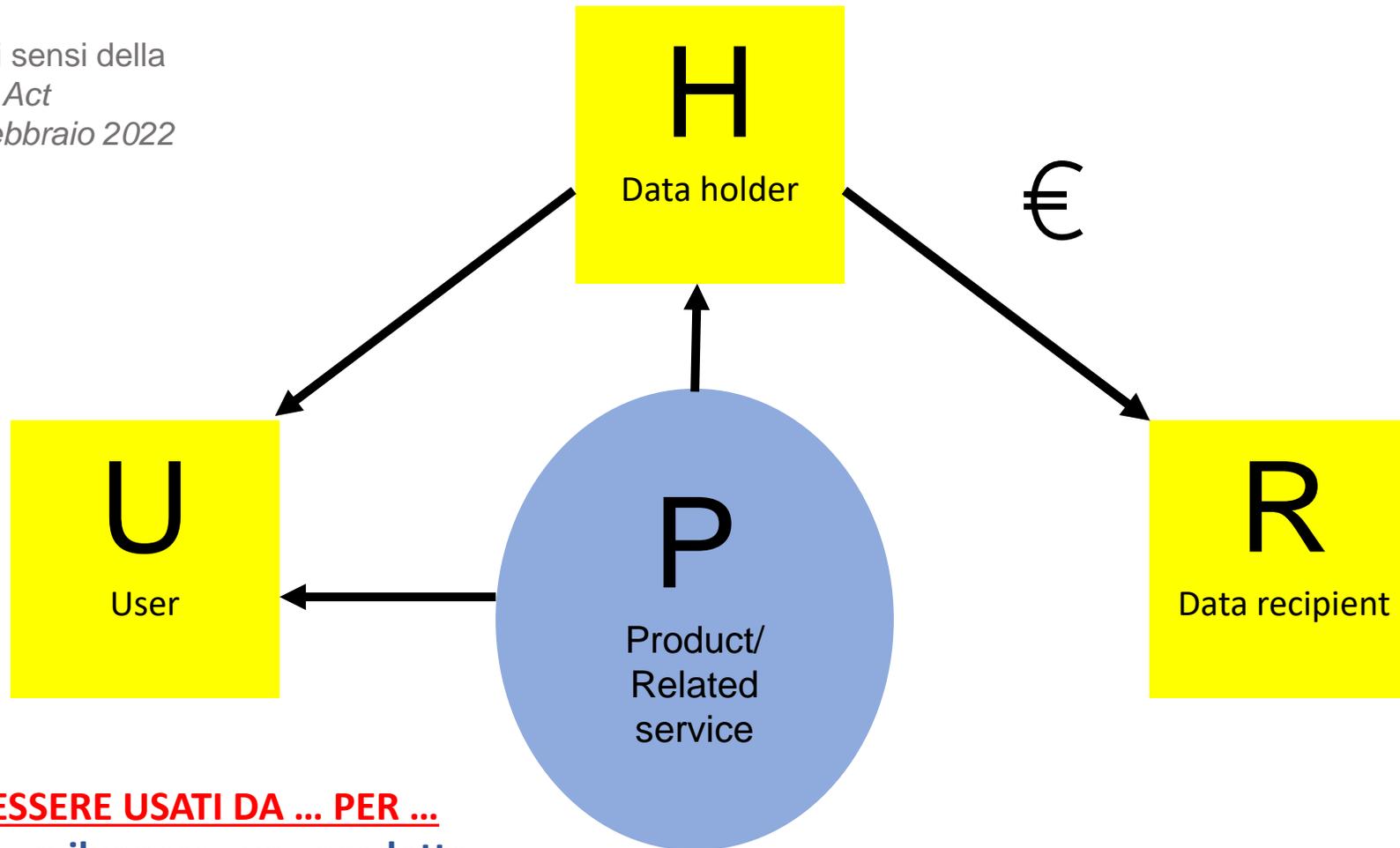
C. CODAGNONE, G. LIVA, T. RODRIGUEZ DE LAS HERAS BALLELL, *Identification and assessment of existing and draft EU legislation in the digital field*, studio richiesto dalla speciale commissione AIDA (Artificial Intelligence in a Digital Age) del Parlamento Europeo [https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU\(2022\)703345](https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2022)703345)

GDPR (2016) «dati personali»: qualsiasi informazione riguardante una persona fisica identificata o identificabile

Reg. *Free Flow of non-personal data* (2018) «dati»: i dati diversi dai dati personali definiti dal *GDPR*

Dir. *Open Data* (2019) manca definizione di dati; in alcune definizioni di dati («dati dinamici», «dati della ricerca», «serie di dati di elevato valore») i dati sono chiamati «documenti»: a) qualsiasi contenuto, a prescindere dal suo supporto (su supporto cartaceo o elettronico, registrazione sonora, visiva o audiovisiva); o b) qualsiasi parte di tale contenuto

Proposta Reg. *Data Governance Act* (2020) Proposta Reg. *Data Act* (2022)
«dati»: qualsiasi rappresentazione digitale di atti, fatti o informazioni e qualsiasi raccolta di tali atti, fatti o informazioni, anche sotto forma di registrazione sonora, visiva o audiovisiva



I DATI NON DEVONO ESSERE USATI DA ... PER ...

- da U e da R per sviluppare un prodotto concorrente con P
- da H per pregiudicare la posizione commerciale di U o R

→ = data flow

- 1) I punti fermi e gli equivoci linguistici
- 2) Il sillogismo alla base del sistema dell'eccezione e il concetto di riproduzione
- 3) I dati ed il processo di regolazione orientato alle finalità del loro uso
- 4) L'idoneità del TDM ad essere un moltiplicatore di proprietà intellettuale**
- 5) Conclusioni



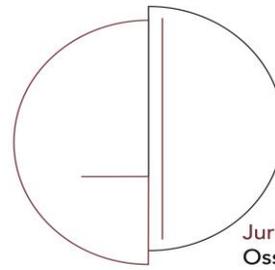
<https://www.eurocomunicazione.com/2021/12/05/quayola-rilegge-larte-classica-grazie-al-digitale/>

- 1) I punti fermi e gli equivoci linguistici
- 2) Il sillogismo alla base del sistema dell'eccezione e il concetto di riproduzione
- 3) I dati ed il processo di regolazione orientato alle finalità del loro uso
- 4) L'idoneità del TDM ad essere un moltiplicatore di proprietà intellettuale
- 5) Conclusioni**

Conclusione:

La nuova disciplina rischia, se non correttamente e rigorosamente inquadrata, di creare ermeneuticamente e surrettiziamente un nuovo sistema di IPRs sui 'dati', che la stessa struttura logica delle eccezioni e delle limitazioni relative a diritti che hanno ad oggetto beni immateriali diversi dai 'dati' impone senz'altro di escludere.

Grazie per l'attenzione



Juridical Observatory on Digital Innovation
Osservatorio Giuridico sulla Innovazione Digitale

salvatore.orlando@uniroma1.it

Visita OGID: <https://web.uniroma1.it/deap/ogid>



SAPIENZA
UNIVERSITÀ DI ROMA