

Methodologies and Data mining techniques for the analysis of Big Data based on Longitudinal Population and Epidemiological Registers

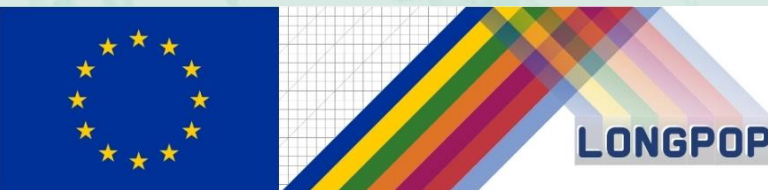


Caso de éxito Marie Skłodowska-Curie Innovative Training Networks (ITN)

Diego Ramiro Fariñas
Instituto de Economía, Geografía y Demografía

diego.ramiro@cchs.csic.es

21 noviembre 2017



This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676060

<http://longpop-itn.eu/>



Información básica de LONGPOP

- **4 años de duración (01/02/2016-31/01/2020) – Convocatoria ITN 2015**
- **7 países, 10 beneficiarios y 1 partner**
- **Contratación de 15 Early-stage Researchers (ESR)**
- **Presupuesto/ayuda total: 3.860.731,44 €**



Objetivo y temas principales

1. Registros longitudinales de población y epidemiológicos
2. Big Data
3. SIG

LONGPOP tiene como objetivo **crear una red** que reúna a diferentes equipos de investigación para compartir experiencias, construir investigación conjunta, y **FORMAR especialistas** en el campo y con ello aumentar el número de usuarios de las grandes, y posiblemente infrautilizadas, **bases de datos**.



Temas principales

- 1. Registros longitudinales de población y epidemiológicos**
2. Big Data
3. SIG



Temas principales



European Historical Population Samples Network

ABOUT

NEWS

LINKS

SEARCH

CONTACT

DATABASES
REPOSITORY
COLLABORATORY
E-JOURNAL

European Historical Population Samples Network

The European Historical Population Samples Network (EHPS-Net) brings together scholars to create a common format for databases containing non-aggregated information on persons, families and households. This format or Intermediate Data Structure (IDS) aims to form an integrated and joint interface between many European databases in order to transcend the currently regional and national studies.

EHPS-Net started in June 2011. Since some of our goals are still in the initial stage, some major items of this portal are still under development. This holds especially for the e-journal and the repository.

EHPS-Net is an ESF Research Networking Programme.



The European Science Foundation (ESF) provides a platform for its Member Organisations to advance science and explore new directions for research at the European level.

Established in 1974 as an independent non-governmental organisation, the ESF currently serves 78 Member Organisations across 30 countries.

[Read more about ESF](#)
[Visit ESF website](#)

[Historical Life Course Studies](#)
Browse articles in e-journal →

Latest news



Workshop working groups 5 and 6

On 17-18 October 2016 the working groups 5 and 6 held a meeting on 'Modern Data & Health' at the Academy of Sciences in Copenhagen, Denmark.



Fourth meeting of EHPS-Net

On 20 September 2016, the fourth meeting of the EHPS-Net took place in Leuven, Belgium in the Raadzaal of the Faculty of Social Sciences of KU Leuven.

This website is hosted and maintained by the International Institute of Social History

EHPS-Net#1 ONCDP Winter School Intermediate Data Structure



¿Qué información está buscando?

Buscar

Estadísticas Longitudinales de Supervivencia y Longevidad en Andalucía, 2002-2013 suministra los datos necesarios para la investigación de las funciones de supervivencia poblacional en el periodo 2002-2013, partiendo de la información recogida en el Censo de población y viviendas de 2001 y complementada con la proporcionada por los boletines estadísticos de defunción y las variaciones residenciales padronales, para el seguimiento de la población censada en 2001 durante los doce años posteriores. Esta publicación incluye tabulaciones con causa de muerte así como mapas de indicadores de mortalidad utilizando celdas de 250 m para mortalidad general y otras de 1 km para mortalidad por causa.



Marco 2007-2013

- Organismo: Instituto de Estadística y Cartografía de Andalucía.
- Ámbito: Comunidad Autónoma, Unidad Territorial, Provincia, Municipio y Secciones Estadísticas.

Actualizaciones:

- Última actualización: 17/03/2016
- Registro de actualizaciones

Metodología:

- Guía rápida
- Metodología
- Memoria técnica

Estadísticas Longitudinales de Supervivencia y Longevidad en Andalucía, 2002-2013

- Nota divulgativa

DATOS

- Tabulación detallada. Serie 2002-2013
 - Resultados básicos
 - Resultados ampliados
- Descarga de microdatos

MAPAS

- Visualizador de mapas*:
 - Razón de mortalidad general. Celdas de 250 metros
 - Razón de mortalidad por causa. Celdas de 1 kilómetro
 - Nota metodológica de las razones de mortalidad

* Nota: El visualizador de mapas ha sido testeado en los siguientes navegadores web: Mozilla Firefox 21 o superior, Internet explorer 9 o superior y Google Chrome 27 o superior.

Actividades relacionadas:

- Censos de Población y Viviendas 2001
- Censos de Población y Viviendas 2011
- Movimiento Natural de la Población
- Estadísticas de Mortalidad por Causas de Andalucía
- Distribución Espacial de la Población en Andalucía



Temas principales

1. Registros longitudinales de población y epidemiológicos
2. Big Data
3. SIG





Temas principales

Big Data

1. Big data será utilizado en la producción estadística futura a partir de 2021
2. Cada vez más presente el big data y el data science en las ciencias sociales

SIG

1. Generalización del uso de información geocodificada (Grids, small areas, etc.)

INNOVACIÓN APLICADA A LA INVESTIGACIÓN SOCIAL



Estructura de LONGPOP

Table 1.1: Work Package (WP) List

WP No.	WP Title	Lead Beneficiary No.	Start Month	End month	Activity Type	Lead Participant Short Name	ESR involvement
1	Concepts and Techniques for Record Linkage and Data managing of longitudinal population, socioeconomic and epidemiological registers.	CSIC (IEGD-CCHS)	1	48	R	CSIC	All ESRs
2	Combining time and space - geographic context in analyses of longitudinal individual level data	LUNDS UNIVERSITET (ULUND)	1	48	R	ULUND	All ESRs
3	Data mining, extraction techniques and software with common methodologies and structures	KNAW (IISH)	1	48	R	IISH	All ESRs
4	Training	RADBOUD U.	6	42	TR	RADBOUD U.	All ESRs
5	Management	CSIC IEGD-CCHS	1	48	MGT	CSIC	0
6	Dissemination and outreach	CSIC IEGD-CCHS	1	48	DISS	CSIC	ESR2



Estructura de LONGPOP

3 WP de investigación - R

1. Todos los socios involucrados en una investigación común (responsabilidades compartidas)
2. *Deliverables* científicos (frecuentes)
3. Conectados con los proyectos individuales de investigación de los 15 ESR (*Individual Research Projects – IRP*)



Estructura de LONGPOP

Table 1.1: Work Package (WP) List

WP No.	WP Title	Lead Beneficiary No.	Start Month	End month	Activity Type	Lead Participant Short Name	ESR involvement
1	Concepts and Techniques for Record Linkage and Data managing of longitudinal population, socioeconomic and epidemiological registers.	CSIC (IEGD-CCHS)	1	48	R	CSIC	All ESRs
2	Combining time and space - geographic context in analyses of longitudinal individual level data	LUNDS UNIVERSITET (ULUND)	1	48	R	ULUND	All ESRs
3	Data mining, extraction techniques and software with common methodologies and structures	KNAW (IISH)	1	48	R	IISH	All ESRs
4	Training	RADBOUD U.	6	42	TR	RADBOUD U.	All ESRs
5	Management	CSIC IEGD-CCHS	1	48	MGT	CSIC	0
6	Dissemination and outreach	CSIC IEGD-CCHS	1	48	DISS	CSIC	ESR2



Estructura de LONGPOP

1 WP de formación - *TR*

1. Basado en las necesidades de formación de investigadores en los temas del proyecto
2. Carácter multidisciplinar → todos los beneficiarios implicados
3. Estructura poco flexible → se pueden realizar más cursos, pero hay que respetar el programa
 - ✓ **IMPORTANTE:** programar los cursos en los 2 primeros años de contrato de los ESR



LONGPOP: Training

	Main Training Events & Conferences	ECTS (if any)	Lead Institution	Project Month (estimated)
1	Through the 3-year training program, a Lecture Series on Demography, Health and Longitudinal Studies which will be transmitted via video streaming and will use the existing platform www.demografia.tv		CSIC	9-42
2	Big data and on time systems and its use for research		TELNET	9
3	One-year ESRI Master GIS Online	48	ESRI	9-21
4	ESRI Training Itinerary DESKTOP, 25-hour course		ESRI	9
5	Workshop: Data Linkage and Data management for longitudinal analysis		IISH	17
6	Two-week Training Course on Introduction to demographic analysis and longitudinal studies		RADBOUD U.	17
7	ESRI Training Itinerary ARCHITECTURE, 25-hour course		ESRI	21
8	Two-week Training Course on Historical Demographic Research Using Register-Type Data		ULUND	21
9	Training module on sequence analysis with TraMineR		U.GENEVE	32
10	Workshop: Methodologies and techniques for the analysis of past surveys and longitudinal registers		U.GENEVE	32
11	ESRI Training Itinerary DEVELOPMENT, 25-hour course		ESRI	33
12	Two-week training course on Agent-based modelling and Spatial Regression Models		CeSO	36

Table 1.1: Work Package (WP) List

WP No.	WP Title	Lead Beneficiary No.	Start Month	End month	Activity Type	Lead Participant Short Name	ESR involvement
1	Concepts and Techniques for Record Linkage and Data managing of longitudinal population, socioeconomic and epidemiological registers.	CSIC (IEGD-CCHS)	1	48	R	CSIC	All ESRs
2	Combining time and space - geographic context in analyses of longitudinal individual level data	LUNDS UNIVERSITET (ULUND)	1	48	R	ULUND	All ESRs
3	Data mining, extraction techniques and software with common methodologies and structures	KNAW (IISH)	1	48	R	IISH	All ESRs
4	Training	RADBOUD U.	6	42	TR	RADBOUD U.	All ESRs
5	Management	CSIC IEGD-CCHS	1	48	MGT	CSIC	0
6	Dissemination and outreach	CSIC IEGD-CCHS	1	48	DISS	CSIC	ESR2



Estructura de LONGPOP

1 WP de gestión - *MGT*

1. Crear estructura como si fuese la creación de un instituto: Supervisory Board, Advisory Board, WP leaders, Project Manager, reuniones y monitoraje
2. *Deliverables* de gestión (informes)
3. Atención a la parte de ética
 - ✓ **IMPORTANTE:** nivel de diseminación de los deliverables de gestión (ej: PU = público; CO = confidencial, etc.)



Estructura de LONGPOP

Table 1.1: Work Package (WP) List

WP No.	WP Title	Lead Beneficiary No.	Start Month	End month	Activity Type	Lead Participant Short Name	ESR involvement
1	Concepts and Techniques for Record Linkage and Data managing of longitudinal population, socioeconomic and epidemiological registers.	CSIC (IEGD-CCHS)	1	48	R	CSIC	All ESRs
2	Combining time and space - geographic context in analyses of longitudinal individual level data	LUNDS UNIVERSITET (ULUND)	1	48	R	ULUND	All ESRs
3	Data mining, extraction techniques and software with common methodologies and structures	KNAW (IISH)	1	48	R	IISH	All ESRs
4	Training	RADBOUD U.	6	42	TR	RADBOUD U.	All ESRs
5	Management	CSIC IEGD-CCHS	1	48	MGT	CSIC	0
6	Dissemination and outreach	CSIC IEGD-CCHS	1	48	DISS	CSIC	ESR2



Estructura de LONGPOP

1 WP de divulgación y diseminación - *DISS*

1. Crear estructura como si fuese la creación de un instituto: Supervisory Board, Advisory Board, WP leaders, Project Manager, reuniones y monitoraje
2. *Deliverables* de divulgación y diseminación (publicaciones, página web, redes sociales)



Estructura de LONGPOP

15 proyectos de investigación individuales

1. Cada proyecto tiene una estructura: objetivos, resultados y estancias (*secondments*)
2. Objetivos individuales tienen que estar alineados con objetivo general y WP de investigación
3. No exceder en la carga de trabajo individual (doctorado a la vez que tareas del proyecto)



Gestión del proyecto

- Tener en cuenta el tiempo real de contratación de 15 ESR (publicidad, visados, etc.)
- Contratación de Project Manager a través de mecanismos CSIC (FC o bolsa)
- Repartición posterior del presupuesto entre los beneficiarios (*B: Institutional Costs*)
 - Reflejada en el acuerdo de consorcio (CA)
 - Cantidades a justificar son las del Acuerdo de Subvención (GA)



LONGPOP - <http://longpop-itn.eu/>

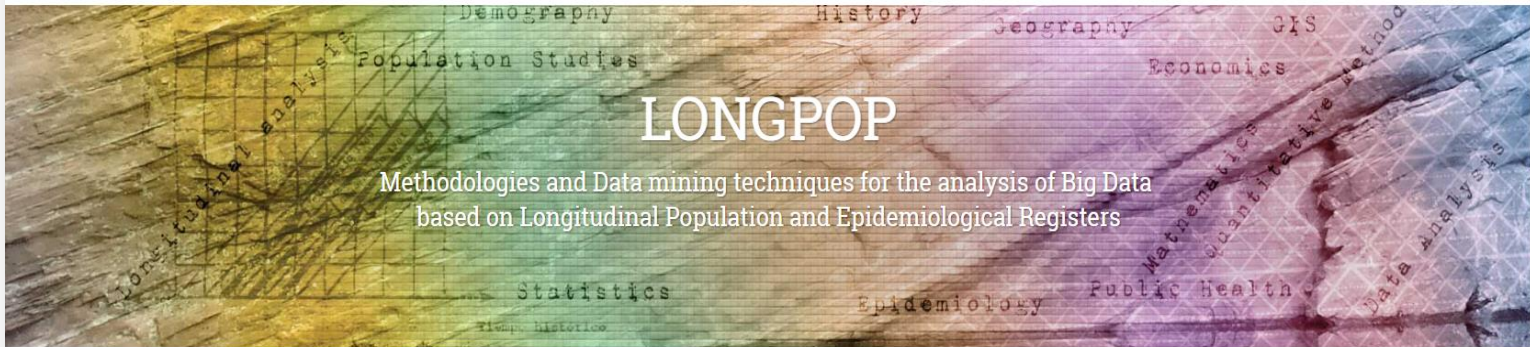
Twitter LinkedIn Facebook

Contact us: +34 916022403 Share



This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676168

HOME PROJECT NETWORK RECRUITMENT RESEARCH TRAINING DISSEMINATION EVENTS Q **Early-Stage Researchers**



Coordinator



Beneficiaries



Partner

