

## Informe metodológico estandarizado de la actividad Métodos automáticos de enlace de registros

Producto de difusión: aLink-Herramienta de fusión de ficheros

### Código Subconcepto Valores Metadatos

#### Contacto

R1	Organización de contacto	Instituto de Estadística y Cartografía de Andalucía
R2	Unidad de organización de contacto	Servicio de Difusión y Publicaciones
R3	Nombre de contacto	--
R4	Función de la persona de contacto	Atención a la demanda de información
R5	Dirección postal de contacto	Pabellón de Nueva Zelanda. C/ Leonardo Da Vinci, nº 21. Isla de La Cartuja. 41071 - Sevilla
R6	Dirección de correo electrónico de contacto	<a href="mailto:informacion.ieca@juntadeandalucia.es">informacion.ieca@juntadeandalucia.es</a>
R7	Número de teléfono de contacto	Atención al público: 900 101 407 Centralita: 955 033 800
R8	Número de fax de contacto	955 033 816/17

#### Actualización de metadatos

R9	Última certificación de metadatos	16/12/2021
R10	Última difusión de metadatos	27/05/2022
R11	Última actualización de metadatos	27/05/2022

#### Presentación estadística

R12	Descripción de los datos	La finalidad general de esta actividad es generar una metodología adecuada que permita implementar métodos automáticos para el enlace de registros o fusión de ficheros. Además, se proporciona una herramienta informática orientada a este fin
R13	Sistema de clasificación	--
R14	Cobertura por sectores	--
R15	Definiciones y conceptos	- Un enlace o coincidencia es un par de registros que hace referencia a la misma entidad o unidad poblacional. - Un no enlace o no coincidencia es aquel par de registros que hace referencia a entidades distintas. - Un posible enlace es aquel par de registros del que no se tiene la seguridad de que sea un enlace o un no enlace.
R16	Unidad estadística	Cada uno de los ficheros objeto de enlace o búsqueda de duplicados
R17	Población estadística	Cualquier fichero de datos con información procedente de fuentes tan distintas como censos, encuestas o fuentes administrativas
R18	Ámbito geográfico	Andalucía y otras
R19	Cobertura temporal	La actividad se incorporó por primera vez a la programación estadística oficial la actividad Métodos Automáticos de Enlace de Registros en el Programa Estadístico Anual 2008, en el marco del Plan Estadístico de Andalucía 2007-2012.
R20	Período base	--

#### Unidad de medida

R21	Unidad de medida	--
-----	------------------	----

#### Período de referencia

R22	Período de referencia	-
-----	-----------------------	---

#### Mandato institucional

R23	Actos jurídicos y otros acuerdos	Ley 4/1989, de 12 diciembre, de Estadística de la Comunidad Autónoma de Andalucía Ley 3/2013, de 24 de julio, por la que se aprueba el Plan Estadístico y Cartográfico de Andalucía 2013-2017 y Ley 6/2017, de 27 de diciembre, por la que se modifica la Ley 3/2013, de 24 de julio, por la que se aprueba el Plan Estadístico y Cartográfico de Andalucía 2013-2017 para ampliar su vigencia hasta el 2020 Decreto 25/2020, de 24 de febrero, por el que se aprueba el Programa Estadístico y Cartográfico de la Comunidad Autónoma de Andalucía para el año 2020
R24	Intercambio de datos	--

#### Confidencialidad

R25	Política de confidencialidad	Ley 4/1989, de 12 diciembre, de Estadística de la Comunidad Autónoma de Andalucía Ley 12/1989, de 9 de mayo, de la Función Estadística Pública Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales
R26	Tratamiento de datos confidenciales	Esta actividad no está sujeta al secreto estadístico ya que se trata de una actividad de soporte cuyo objetivo es fusionar o enlazar ficheros con registros referentes a una misma unidad común. El que utilice la herramienta será responsable en todo momento de la información con la que trabaja y del uso que se haga de la misma

#### Política de difusión

R27	Fecha de difusión de resultados	--
R28	Fecha prevista de difusión de resultados	--
R29	Acceso a los resultados	El resultado de la actividad es la aplicación informática denominada aLink: Herramienta de Fusión de Ficheros. Se puede acceder a ella a través de la siguiente url: <a href="http://www.juntadeandalucia.es/institutoestadisticaycartografia/ieagen/otrosServidores/software/index.htm#uno">http://www.juntadeandalucia.es/institutoestadisticaycartografia/ieagen/otrosServidores/software/index.htm#uno</a>

#### Frecuencia de la difusión

R30	Frecuencia de la difusión	Irregular
-----	---------------------------	-----------

#### Formato de difusión

R31	Comunicados	No
R32	Publicaciones	aLink: Herramienta de Fusión de Ficheros
R33	Base de datos online	--
R34	Acceso a microdatos	--
R35	Otros	--

#### Accesibilidad de la documentación

R36	Documentación sobre metodología	La memoria técnica está disponible en Internet en el apartado Metodología
R37	Documentación sobre calidad	El IECA para garantizar la calidad en esta actividad aporta la memoria técnica, la ficha de indicadores de calidad y el informe metodológico estandarizado de la actividad

#### Gestión de la calidad

R38	Garantía de calidad	En cada proceso (tanto en la normalización como en el proceso de enlaces) se realiza un análisis exhaustivo y manual de cada uno de ellos. Para garantizar que los enlaces son correctos se estudian también los enlaces duplicados para obtener sólo aquel enlace de registro de mayor exactitud
R39	Evaluación de calidad	Se ha elaborado una ficha de calidad y está disponible en Internet en el apartado Metodología

#### Relevancia

R40	Usuarios	Administraciones públicas, empresas privadas, universidades y personas interesadas en procesos de normalización y geocodificación de direcciones postales, así como de normalización de nombres de personas
R41	Satisfacción del usuario	--
R42	Exhaustividad	El éxito de los resultados en la fusión de ficheros va a depender de como esté recogida la información de las variables de los ficheros. Cuanto mejor sea dicha información, los procesos de enlaces y los resultados que se obtengan serán más óptimos

#### Exactitud y fiabilidad

R43	Exactitud global	La información con la que se trabajará en esta actividad procede de fuentes administrativas en la mayoría de los casos. Por tanto, la información extraída de estas fuentes es fiable
R44	Errores de muestreo	--
R45	Errores ajenos al muestreo	Los recursos personales disponibles en cada momento y el tamaño de los ficheros pueden ralentizar los procesos de fusión de ficheros. El uso de la herramienta requiere un conocimiento amplio de la herramienta por lo que es imprescindible contar con personal cualificado para soportar suministro técnico

#### Oportunidad y puntualidad

R46	Oportunidad	No procede
R47	Puntualidad	No procede

#### Comparabilidad

R48	Comparabilidad geográfica	A nivel internacional, destaca la aplicación Febri desarrollada por la Universidad Nacional de Australia. A nivel europeo, el Instituto de Estadística Italiano, teniendo como apoyo metodológico un grupo de profesionales de distintos ámbitos (estadística e informática), ha desarrollado una herramienta de enlace de registros denominada RELAIS (REcord Linkage at IStat). En España, la aplicación de métodos automáticos de fusión de registros comenzó en el Instituto Vasco de Estadística (EUSTAT) en los años 90, desarrollado con lenguaje SAS
R49	Comparabilidad temporal	--

#### Coherencia

R50	Coherencia - cruce de sectores	--
R51	Coherencia - interna	La coherencia interna se garantiza a través de diversas validaciones realizadas antes de la publicación de los datos

#### Costes y carga

R52	Costes y carga	72.000,00 €
-----	----------------	-------------

#### Revisión de los datos

R53	Revisión de datos - política	Al no existir datos provisionales no se realiza una revisión de los datos
R54	Revisión de datos - práctica	--

Tratamiento estadístico

R55	Datos de origen	Fuente administrativa
R56	Frecuencia de la recogida de datos	-
R57	Recopilación de datos	Aunque el tipo de recogida de esta actividad se considera fuente administrativa, no podemos dar la información exhaustiva de las fuentes o registros administrativos que se van a utilizar en esta actividad ya que pueden ser cualquiera de los que se recogen en el Inventario de fuentes administrativas de Andalucía. En definitiva, esta actividad trabajará con las fuentes administrativas o cualquier otro tipo de fichero que se requieran dependiendo del proceso que se lleve a cabo y de las necesidades que surjan en cada momento.
R58	Validación de datos	--
R59	Compilación de los datos	Previamente, se debe realizar de modo manual una revisión de los ficheros para analizar el diseño de registro y detectar campos que ambos registros tienen en común. A continuación es imprescindible, y se debe hacer de forma obligatoria, un Tratamiento previo con la herramienta aLink a todos los ficheros de datos con los que se va a trabajar. Y por último se realizan cada una de las fases del proceso de fusión de ficheros: fase de normalización, fase de agrupación de registros, fase de comparación de pares de registros y fase de clasificación
R60	Ajuste	--

Observaciones

R61	Observaciones	El manual de la herramienta aLink en el siguiente enlace: <a href="http://www.juntadeandalucia.es/institutoestadisticaycartografia/ieagen/otrosServidores/software/alink/manual_aLink.pdf">http://www.juntadeandalucia.es/institutoestadisticaycartografia/ieagen/otrosServidores/software/alink/manual_aLink.pdf</a>
-----	---------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Nota.- Datos referidos a Informe de Evaluación del Programa Estadístico y Cartográfico de la Comunidad Autónoma de Andalucía para el año 2020, aprobado por la Comisión Interdepartamental Estadística y Cartográfica el 16/12/2021