

**USOS Y APLICACIONES DE LA ESTADÍSTICA MUNICIPAL DE ACTIVIDAD EN LAS
OPERACIONES ESTADÍSTICAS DE EUSTAT**

Marta Mas Moreno m-mas@eustat.eus

Nerea Linaza Diarce nerea_linaza@eustat.eus

Instituto Vasco de Estadística

(EUSTAT)

0. Introducción

Durante los últimos diez años EUSTAT ha extendido el uso de fuentes administrativas en la elaboración de estadísticas oficiales. Esto ha posibilitado disponer y difundir información cada vez más detallada y exhaustiva en materia de población, actividad, educación o vivienda en periodos intercensales.

En el año 2008 se puso en marcha la operación estadística “Base de datos de estadísticas sociodemográficas”, cuyo objetivo es servir como repositorio para la recogida, grabación y tratamiento de la información basada en registros estadísticos y administrativos para la posterior explotación y difusión de las Estadísticas municipales de población, población activa, educación y vivienda.

Asimismo, el uso y objeto de estas operaciones no se limita únicamente a la difusión periódica de información a nivel municipal, sino que permite completar y mejorar la información recogida por encuestas tan relevantes como la Encuesta de población en relación con la actividad (en adelante PRA) o actualizar y depurar el Directorio de actividades económicas, así como el Directorio de viviendas.

Un buen ejemplo de esta nueva forma de trabajo es la Estadística municipal de población activa (en adelante EMPA) que se elabora con información que proviene de fuentes estadísticas y administrativas, entre las que se encuentran la afiliación a la Seguridad Social y el paro registrado. Esta información ha sido incorporada recientemente a la PRA. Además, la combinación de las distintas fuentes administrativas de la EMPA en la generación de variables como la situación profesional de la población ocupada, permite la mejora en la estimación de subgrupos de gran interés socioeconómico como miembros de cooperativas o el empresariado.

El objetivo de esta comunicación se centra principalmente en presentar éstas y otras utilidades de la información contenida en la EMPA y evaluar su impacto, tanto en el desarrollo de la propia operación, como en la mejora de otros procesos y operaciones estadísticas.

1. Uso e integración de las fuentes administrativas y estadísticas en la EMPA

El objetivo general de la EMPA se basa en la recogida, integración y explotación de la información individual detallada en materia de actividad, a nivel municipal y con una periodicidad anual. A tal fin se precisa información de origen administrativo y estadístico, a nivel de registro individual e identificado. Esta información proviene de diversas fuentes que deben ser depuradas e integradas de forma que se complementen unas con otras y proporcionen una información coherente y de calidad.

Las principales fuentes de información de la EMPA se pueden clasificar en:

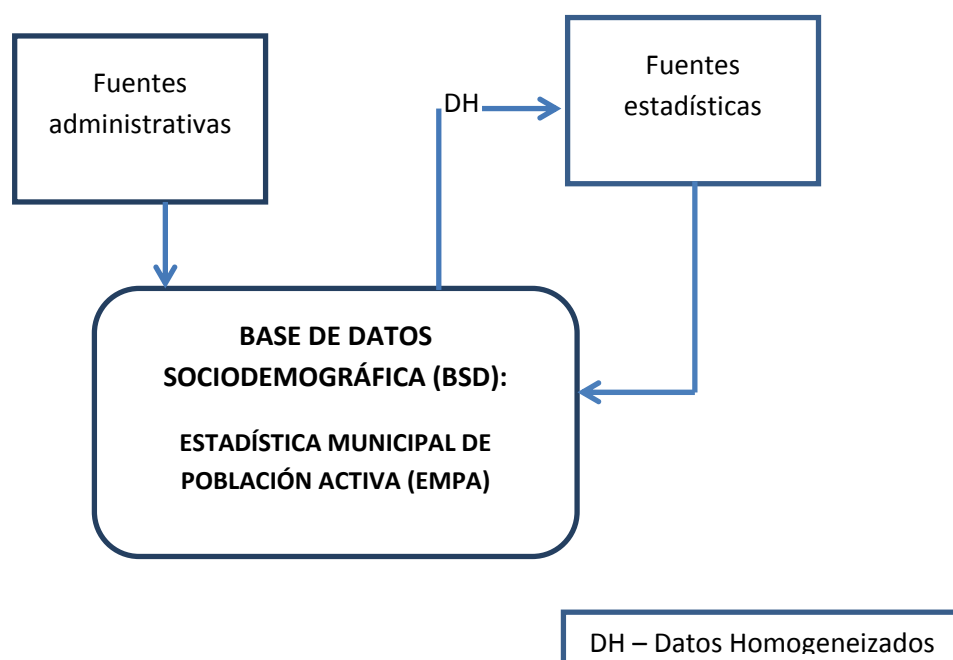
Fuentes administrativas

- Fichero de personas físicas, jurídicas y demás entidades que están obligadas a retener o a ingresar a cuenta del Impuesto sobre la Renta de las Personas Físicas (IRPF) reflejado en el modelo anual de retenciones (modelo 190) y que proporcionan las Haciendas Forales de los tres Territorios Históricos.
- Ficheros de afiliación y cuentas de cotización de la Seguridad Social enviados por Tesorería General de la Seguridad Social (TGSS).
- Ficheros de demandantes de empleo y contratos realizados facilitados por el Servicio Vasco de Empleo (LANBIDE)
- Fichero de titulares de pensiones contributivas que proporciona el Instituto Nacional de la Seguridad Social (INSS).
- Fichero de mutualistas vascos que constan en la Mutualidad General de Funcionarios Civiles del Estado (MUFACE).

Fuentes estadísticas:

- Encuesta de Población en Relación con la Actividad (PRA)
- Censo del Mercado de Trabajo (CMT)
- Censos de población y viviendas y/o Estadística de población y viviendas.
- Directorio de Actividades Económicas (DIRAE)
- Estadística Municipal de Educación (EME)
- Estadística Municipal de Habitantes (EMH)

Además, esta integración no se produce en un único sentido, ya que el trabajo de homogeneización e integración de fuentes realizado dentro de los procedimientos de ejecución de la EMPA revierte, a su vez, en la mejora de la información en las fuentes de origen, sobre todo en las fuentes estadísticas, tal y como refleja este diagrama:



La gestión de esta cantidad de fuentes requiere de unos procedimientos de ejecución claros y coordinados que permitan, por un lado, la coherencia temporal de la información (algunos ficheros son trimestrales, otros sin embargo, anuales) y, por otro, el establecimiento de un orden de prioridad a la hora de determinar qué fuente es más precisa y fiable para medir las variables o características de estudio. Así pues, para establecer los valores de variable principal objetivo de la EMPA, la relación con la actividad de la población que reside en la C.A. de Euskadi, se utilizan las siguientes fuentes y orden de prioridad, según el atributo a asignar:

- I. *PRA* – Se graba de forma directa los atributos recogidos por esta encuesta para la relación con la actividad en cada una de las unidades seleccionada en la muestra.
- II. *CMT* - Se graba de forma directa los atributos recogidos por esta encuesta para la relación con la actividad en cada una de las unidades seleccionada en la muestra.
- III. *MUFACE*- Se recoge la situación relacionada con la ocupación y diferentes situaciones de inactividad (jubilación, incapacidad,...) del funcionariado.
- IV. *Ficheros de la TGSS (cuentas y afiliación)* - Se trata de la principal fuente para recoger la población ocupada y diferentes situaciones profesionales (empresariado, tipologías de autónomos y asalariados, ayudas familiares, etc.).
- V. *Fichero del INSS* - De esta fuente se extrae principalmente información sobre personas jubiladas y pensionistas.
- VI. *Ficheros de LANBIDE - Demandas*. Constituye la fuente principal para el establecimiento de la población parada. En el tratamiento de esta fuente se tiene en cuenta la metodología utilizada en la cuantificación del paro registrado con el objeto de adecuarla a la definición de paro OIT y a las utilizadas en la PRA y otras encuestas de fuerzas del trabajo de referencia (EPA, EU-LFS, etc.).
- VII. *Ficheros de perceptores de rentas (Modelo 190)* de las Diputaciones Forales – Permite contrastar, depurar y completar aquella información sobre ocupación y paro proporcionada por otras fuentes principales.
- VIII. *Ficheros de LANBIDE - Contratos*. Actualmente se utiliza como fuente secundaria para el contraste de situaciones relacionadas con la ocupación. La obtención de información más precisa, sobretudo en cuanto a la fecha de término de las contrataciones, permitirá aprovechar al máximo el potencial de esta fuente y delimitar de forma más fiable la situación respecto a la actividad a una fecha concreta.
- IX. *Estadística municipal de educación (EME)* – Se señalan como población estudiante aquellas personas que figuran como tal en esta estadística.
- X. *Estadística municipal de habitantes (EMH)* – Es la base que define la población empadronada en la C.A. de Euskadi en un año determinado. Contiene las variables de identificación de cada persona. En concreto, las variables de edad y territorio contenidas en esta estadística permiten delimitar a la población mayor de 16 años y depurar la ubicación y lugar de residencia de las personas.

La siguiente tabla muestra, a modo de resumen, las principales fuentes utilizadas y los atributos de la relación con la actividad para los que proporcionan información:

FUENTES		Atributos de la relación con la actividad						
		Población activa		Población inactiva				
		Ocupación	Paro	Jubilación	Incapacidad, dependencia	Estudiantes	Labores del hogar	Otros
Fuentes estadísticas	PRA	✓	✓	✓	✓	✓	✓	✓
	CMT	✓	✓	✓	✓	✓	✓	✓
	EME					✓		
Fuentes administrativas	MUFACE	✓		✓	✓			
	TGSS	✓						
	INSS			✓	✓			✓
	RETEN	✓	✓	✓	✓			✓
	LANBIDE Demandas		✓			✓		
	LANBIDE Contratos	✓						

Esta enumeración de fuentes estadísticas y administrativas debe tomarse como una simplificación de los procedimientos llevados a cabo en la imputación de los atributos de la relación con la actividad. La imputación de esta característica dentro de la EMPA es fruto de una combinación de procesos determinísticos y aleatorios que utiliza y contrasta variables presentes en las diferentes fuentes de forma que, con toda la información disponible, la variable resultante refleje de la forma más precisa y contrastada la situación respecto a la actividad para la población objetivo y el periodo de referencia concretos.

2. Aproximación a variables de interés: La situación profesional

Tal y como se ha mencionado anteriormente, en algunos casos es necesario el uso y contraste de la información proporcionada por varias fuentes para derivar características de interés. Es el caso de la situación profesional, que define la forma que toma la relación laboral de la población ocupada con la empresa donde trabaja. Para la población parada que haya trabajado anteriormente y para la población jubilada, también es posible la definición de esta

característica que hará referencia a la forma que toma la relación laboral anterior al paro o a la jubilación.

La reciente crisis económica y las consecuencias derivadas de ésta en el tejido económico y empresarial ha acrecentado el interés por conocer la evolución de ciertos colectivos medidos por esta variable. Estos colectivos son:

1. Grupo empresarial: Trabaja por cuenta propia y emplea personal.
2. Grupo autónomos: Trabaja por cuenta propia y no emplea personal.
3. Miembros de cooperativas: Pertenece a una cooperativa.
4. Grupo asalariado fijo: Trabaja por cuenta ajena a sueldo, jornal, comisión u otra clase cualquiera de remuneración, con carácter fijo.
5. Grupo asalariado eventual: Trabaja por cuenta ajena a sueldo, jornal, comisión u otra clase cualquiera de remuneración, con carácter eventual o interino.
6. Ayuda familiar: Realiza trabajos sin remuneración reglamentada en la empresa o negocio de un familiar.
7. Otra situación.

La generación de esta característica dentro de la EMPA combina la información de varias fuentes. Por un lado, la del fichero de afiliación a la Seguridad Social de la que se obtiene una primera aproximación a las categorías de la variable y, por otro lado, el DIRAE como fuente interna para comprobar y completar el colectivo de autónomos y empresarios. Además, se utiliza el fichero de retenciones basadas en el modelo 190 de las Diputaciones Forales para afinar y depurar el colectivo de cooperativistas utilizando la información auxiliar que proporciona dicha fuente.

Por lo tanto, el proceso de generación de la variable situación profesional se llevará a cabo en varias fases según la fuente utilizada en cada caso:

Generación de la situación profesional – Fichero de afiliación y cuentas de la Seguridad Social

En esta fase se utiliza como fuente principal los ficheros de la Seguridad Social, más concretamente, los distintos regímenes de afiliación, el tipo de relación laboral con la empresa, los colectivos especiales y los tipos de contrato presentes en los ficheros de cuentas y trabajadores afiliados. En esta primera aproximación los grupos o colectivos de interés quedarán clasificados de la siguiente manera:

1. Grupo empresarial:
 - o Personas incluidas en los colectivos de armadores y consejeros-administradores de Sociedades Mercantiles.

- Personas incluidas en el grupo de autónomos que constan como miembros del órgano de administración en la relación laboral con otras entidades o autónomos.
 - Personas asalariadas que constan como personas físicas en el fichero de cuentas de cotización y con nº de personas que trabajan en la empresa mayor que 1.
2. Grupo autónomos: Personas afiliadas en los siguientes Regímenes:
 - Régimen Especial de Trabajadores Autónomos (incluye el Régimen Especial Agrario).
 - Régimen Especial del Mar Cuenta Propia.
 3. Miembros de cooperativas:
 - Personas que constan en la relación laboral como socios trabajadores de cooperativas (trabajo asociado y explotación de la tierra).
 - Personas con cuenta de cotización asociada al colectivo especial de cooperativas.
 4. Grupo asalariado fijo:
 - Personas con tipo de contrato indefinido.
 - Personas que constan en la relación laboral como funcionariado, personal estatutario o clero.
 5. Grupo asalariado eventual:
 - Personas con tipo de contrato de aprendizaje, temporal o de duración determinada.
 - Personas que constan en la relación laboral como funcionariado interino, altos cargos y alcaldes, personal estatutario temporal, personas reclusas con actividad laboral, personas que trabajan en el extranjero y programas de formación.
 6. Ayuda familiar:
 - Personas del grupo de autónomos que constan en la relación laboral con otras entidades o autónomos como familiar de socio de sociedad mercantil capitalista.
 - Personas del grupo de autónomos que constan como familiar colaborador del titular (bonificado o no) en la relación laboral con otras entidades o autónomos.

Generación de la situación profesional – Directorio de Actividades Económicas

En un segundo paso se actualizará la asignación de valores de la variable situación profesional con la información que aparece en el DIRAE para el mismo periodo de referencia. Esta actualización pretende clasificar en el grupo de autónomos o en el grupo empresarial a aquellas personas físicas cuya actividad económica esté recogida en esta fuente. Por lo tanto, para aquellos registros de afiliados de la Seguridad Social que estén presentes en DIRAE se actualizará el valor de la situación profesional según los siguientes criterios:

- Si el registro es coincidente (se trata de la misma persona física) y la información sobre el número de personas empleadas contenida en el DIRAE es 1, se considerará a dicha persona dentro del colectivo de autónomos.
- Si el registro es coincidente (se trata de la misma persona física) y la información sobre el número de personas empleadas contenida en el DIRAE es mayor que 1, se considerará a dicha persona dentro del colectivo de empresarios.

Generación de la situación profesional – Fichero de retenciones de las Diputaciones Forales

De cara a completar y depurar el grupo de cooperativistas se utilizará la información contenida en el fichero anual de retenciones practicadas según modelo 190 proporcionado por las Diputaciones Forales. El colectivo de cooperativistas está, en su gran mayoría, afiliado al Régimen Especial de Trabajadores Autónomos por lo que, al carecer de una cuenta de cotización asociada en los ficheros de cuentas de la Seguridad Social, no es posible establecer un vínculo directo con la empresa para la que trabajan. Por ello, y únicamente para el colectivo de autónomos que declaran ser miembros socios de cooperativa o empresa colectiva, se localizará en el fichero de retenciones el NIF de la empresa a la que aparece vinculado y si éste es atribuible a una Sociedad Cooperativa se corregirá su estado en la variable situación profesional, pasando de autónomo a ser clasificado como cooperativista.

3. La Unidad de Población Básica (UPB): El enlace único con el resto de encuestas

Todos los tratamientos descritos anteriormente, así como el enlace de fuentes necesario para llevarlos a cabo, no serían posibles sin la existencia de una clave única presente en cada una de las fuentes y ficheros implicados. La asignación de esta clave única es un paso imprescindible y sistemático incluido en el tratamiento de todas y cada una de las fuentes que utiliza la EMPA. Esta asignación se lleva a cabo mediante procedimientos de fusión determinística y probabilística con el Registro de Población de EUSTAT. Dentro de este registro se asigna y mantiene la denominada Unidad de Población Básica (UPB) que constituye la clave universal de población para el enlace de encuestas y estadísticas de población en EUSTAT. La asignación de esta UPB debe cumplir en todos los casos con las siguientes premisas:

- La clave de UPB es **única** para cada individuo y no existen duplicados. Esto significa que una misma UPB no puede estar asignada a más de un individuo y éste aparece una única vez en el fichero o fuente.
- La clave de UPB es **coherente a lo largo del tiempo**, es decir, un mismo individuo en diferentes referencias temporales de la operación estadística tiene siempre la misma clave de UPB. Se implementan procedimientos que permiten mantener actualizada la

información sobre las claves de UPB en todos los periodos cuando se haya producido algún cambio, ya sea alta, baja o modificación.

- La información de la clave de UPB es **completa**. Esto implica que los procedimientos de fusión garantizan el máximo porcentaje de UPB's asignadas de forma correcta, descartando para la fusión aquellos registros no válidos, bien sea por no contener suficiente información para la fusión o bien por no cumplir ésta con los umbrales de calidad requeridos. Además, el Registro de Población abarca, exclusivamente, a la población residente en la Comunidad Autónoma de Euskadi, lo que permite descartar de forma directa a la población no residente que se encuentra presente en algunas de las fuentes de origen.

4. Principales aplicaciones de las fuentes administrativas de la EMPA en la PRA

La utilización de registros administrativos sobre afiliación y paro registrado en la PRA se inserta de lleno en la metodología de la operación desde el año 2015 y afecta a todas las fases de esta operación estadística, desde la recogida de información hasta los procedimientos de explotación y difusión.

El objetivo es establecer una relación entre la muestra de la PRA, los ficheros de afiliación a la Seguridad Social y los de paro registrado de LANBIDE que permita explicar las diferencias existentes entre la información proporcionada por la encuesta y los respectivos registros administrativos en lo que se refiere a personas ocupadas y paradas de acuerdo con las definiciones aplicables a la Organización Internacional del Trabajo (OIT).

Según la fase a las que afectan en cada momento, las principales aplicaciones se pueden resumir en:

- *Recogida de información*

El cuestionario Web de la encuesta enlaza con los ficheros de afiliación y de paro registrado y presenta a la persona encuestada al inicio de cada trimestre su situación con respecto a dichas variables. Esto permite realizar chequeos de las respuestas proporcionadas por dichas personas en el mismo momento de la encuestación de manera que se puede verificar la coherencia de las respuestas o, si es el caso, establecer las razones de las diferencias existentes.

- *Diseño muestral*

Se utiliza la última información disponible de afiliación a la Seguridad Social y paro registrado en LANBIDE como variables de equilibrio de la muestra. El equilibrado de muestras garantiza una mejor representatividad de las características utilizadas como

variables de equilibrio y mantiene las proporciones de la población original en la muestra, teniendo en cuenta las probabilidades de inclusión del diseño.

- *Elevación*

La disponibilidad a nivel poblacional de la información sobre afiliación y paro registrado permite el cálculo de elevadores que se ajustan a los marginales definidos por estas variables para la población.

- *Generación de características*

La variable Situación Profesional generada en la EMPA a partir de la información de Seguridad Social es utilizada dentro de los procedimientos de la PRA para corregir y depurar atributos de esta característica sumándole así el valor añadido que supone el uso y combinación de diferentes fuentes administrativas y estadísticas.

- *Difusión*

Esta integración de fuentes administrativas en los procedimientos de ejecución de la PRA permite la difusión de las denominadas **tablas de conciliación** con registros administrativos. Estas tablas de análisis contrastan la principal información sobre mercado de trabajo ajustada a los criterios OIT proporcionada por la PRA, con la información de afiliación y paro registrado, siempre teniendo en cuenta las diferencias existentes entre los conceptos medidos por cada fuente y el ámbito poblacional que abarcan cada una de ellas.

Cabe destacar de nuevo la importancia de la clave única de UPB que permite realizar el enlace de las diferentes fuentes administrativas con las unidades muestrales de la PRA de forma directa y sencilla. La presencia de este identificador único tanto en las fuentes de origen como en las de destino será una de las principales claves para el desarrollo de diferentes usos de la información administrativa en ésta y otras operaciones estadísticas.

5. Otras aplicaciones

El anterior es un claro ejemplo de las aplicaciones y usos que se le pueden dar a las fuentes administrativas tratadas en la EMPA en otras operaciones estadísticas. Tal y como se ha visto en el caso de la PRA, el trabajo previo de homogeneización e integración de fuentes administrativas realizado dentro de los procedimientos de ejecución de la EMPA puede revertir en la calidad y mejora de la información elaborada por otras operaciones estadísticas. Es el caso de las operaciones y aplicaciones que se describen a continuación y que constituyen por un lado, demandas y necesidades de información actuales por parte de distintas operaciones estadísticas y, por otro lado, potenciales líneas de trabajo:

- *DIRAE*: Este directorio ya constituye actualmente una fuente de contraste en la generación de variables dentro de la EMPA. A su vez, la información administrativa contenida en EMPA puede ser de gran utilidad para los fines de esta operación. Concretamente, se está trabajando en las siguientes líneas:
 - o Intercambio de información sobre la afiliación a la Seguridad Social que ayude a delimitar de forma más exhaustiva el colectivo de autónomos y cooperativistas, de especial relevancia en la actividad económica de la Comunidad Autónoma.
 - o Inclusión en el DIRAE de la información de género de las personas físicas presente en los ficheros administrativos de la EMPA. Esta demanda pretende cumplir con el compromiso histórico de EUSTAT con las políticas de igualdad y la difusión de la información estadística por sexo.
 - o Mejora en la recogida del empleo de empresas y establecimientos. La información sobre empleo recogida en los ficheros de cuentas de cotización, así como la contabilización de trabajadores afiliados asociados a una misma empresa permitiría el contraste y depuración de la variable empleo recogida por el Directorio.

- *Directorio de viviendas*: Este directorio constituye el marco principal de extracción de muestras para multitud de encuestas a familias e individuos realizadas dentro de la estadística oficial. Este marco requiere de una serie de variables individuales y de vivienda, lo más completas y actualizadas posible, de cara a establecer tipologías o grupos de secciones que se utilizan posteriormente en la estratificación de algunas encuestas. Además, las características individuales se utilizan en los diseños muestrales equilibrados, en los que la información auxiliar puede mejorar la selección y representatividad de la muestra cuando se utilizan variables relacionadas con el objeto de estudio. Disponer de información actualizada sobre las principales variables del mercado de trabajo mejora la definición de las tipologías y añade posibilidades al equilibrado de muestras.

- *Estadística Municipal de Educación (EME)*: La integración en la EMPA de la información proporcionada por esta estadística permitirá la detección de itinerarios profesionales y el análisis de la inserción laboral según el nivel de estudios.

- *Estadística de Nacimientos*: Los cambios en la recogida de la información referida a los nacimientos (actualmente, la inscripción de los recién nacidos puede realizarse directamente desde los centros sanitarios) ha afectado a la calidad en la recogida de variables como la profesión de los progenitores. Existe una demanda específica sobre esta variable dirigida a los registros administrativos de la EMPA de cara a actualizar y completar esta falta de información.

- *Registro de población:* Actualmente este registro es la base para la fusión y asignación de las claves únicas de población o UPB's a los registros de la EMPA. En los propios procedimientos de fusión implementados para lograr esta asignación de claves se han detectado puntos de mejora en el almacenamiento de algunas variables del registro de población tales como el DNI. Es el caso del DNI para extranjeros que a menudo suele registrarse en el Padrón Municipal (fuente principal del Registro de población) con un cifrado provisional, hasta que se les asigna un número de DNI definitivo. Las fuentes administrativas de la EMPA contienen, en muchas ocasiones, la información del documento definitivo, por lo que se podrían actualizar los valores provisionales del Registro de Población.

6. Conclusiones y futuro.

Es evidente el gran potencial de los registros administrativos en la generación y producción de estadísticas oficiales. No obstante, la interpretación y gestión de estas fuentes no es siempre sencilla por lo que la existencia de operaciones como la EMPA que trata, depura y homogeneiza fuentes tan complejas como los datos de afiliación o paro registrado, facilita el acceso a estos datos por parte de otras operaciones estadísticas.

Además el vínculo rápido y directo con otros ficheros estadísticos de EUSTAT que proporciona la Unidad de Población Básica (UPB) abre la puerta a enlaces y aplicaciones de la EMPA en nuevos proyectos que están en marcha en estos momentos dentro del Instituto tales como la homogeneización de direcciones postales, el nuevo servicio de difusión de información geográfica LUR-Data o la descripción de la inserción laboral de las personas tituladas universitarias.