



**AGENT-BASED
SUPPORT TOOL FOR
THE DEVELOPMENT
OF AGRICULTURE POLICIES**

D2.2. Módulo de extracción de *Big Data*



Deliverable Number	D2.2
Lead Beneficiary	AUTH
Authors	AUTH, IDENER
Work package	WP2
Delivery Date	31/08/2022 (M36)
Dissemination Level	Público

www.agricore-project.eu



The Agricore project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No. 816078





Document Information

Project title	Agent-based support tool for the development of agriculture policies
Project acronym	AGRICORE
Project call	H2020-RUR-04-2018-2019
Grant number	816078
Project duration	1.09.2019-31.8.2023 (48 months)
Deliverable authors	Michail Tsagris (AUTH), Vangelis Tzouvelekas (AUTH), IDENER Team
Deliverable reviewers	IDENER Team

Version History

Version	Descripción	Organización	Fecha
0.1	Propuesta	AUTH	10 Mayo 2022
0.2	Aprobación	IDENER	15 Mayo 2022
0.3	Inclusión de contenido (Primer borrador)	AUTH	19 Julio 2022
0.4	Revisión y comentarios	IDENER	07 Agosto 2022
0.5	Implementación de las correcciones	IDE	20 Agosto 2022
1.0	Exportación final y formato (Versión Final)	AUTH	31 Agosto d2022

Resumen ejecutivo

AGRICORE es un proyecto de investigación financiado por la Comisión Europea bajo la convocatoria RUR-04-2018, parte del programa H2020, que propone una forma innovadora de aplicar la modelización basada en agentes para mejorar la capacidad de los responsables políticos de evaluar el impacto de las medidas relacionadas con la agricultura dentro y fuera del marco de la Política Agrícola Común (PAC).

Este entregable presenta el módulo de extracción de datos AGRICORE (DEM), que permite localizar, extraer y almacenar todo tipo de datos necesarios para el uso de las diferentes herramientas y módulos de la plataforma AGRICORE. También presenta los principales tipos de información que son necesarios para la inicialización, calibración y uso de los diferentes módulos AGRICORE, así como las principales fuentes de datos de las que se puede extraer esta información.

Entre las funcionalidades del módulo de extracción de datos se encuentra también el análisis estadístico de las variables contenidas en los conjuntos de datos importados, así como la detección de relaciones de correlación entre ellas. Este entregable presenta la secuencia de procesamiento que se realiza sobre cada una de las variables individualmente, así como sobre todas ellas para detectar interdependencias.

Los resultados de estas operaciones generan metadatos estadísticos y direcciones prohibidas entre las variables. Ambos resultados son argumentos que también se pasan como entrada al módulo de fusión de datos, presentado en el entregable D2.3.

Conclusiones

Este entregable presenta las funcionalidades del módulo de extracción de datos de AGRICORE, incluyendo la obtención de datos desde ubicaciones locales o en la nube, la transformación básica y la exploración de los conjuntos de datos obtenidos, y su carga en el Almacén de Datos. Las conexiones del MDE con ARDIT y con el propio DWH, que se explican en el apartado 4, son fundamentales para el correcto funcionamiento del proceso de búsqueda y obtención de conjuntos de datos.

También se ha presentado el flujo genérico de conjuntos de datos desde su localización a través de ARDIT hasta su eventual proceso de fusión para producir conjuntos de datos derivados enriquecidos. Para ello, se ha utilizado como ejemplo un indicador climatológico (SPEI12) cuya resolución espacial debe cambiarse ({Lat,Lon} a NUTS3) para adaptarse al formato requerido por el Módulo Biofísico de AGRICORE.

Este entregable también presenta los tipos de información que son necesarios para inicializar y configurar los módulos que componen la herramienta AGRICORE. También se describen brevemente las principales fuentes de datos de las que normalmente se extraen los datos necesarios.

Además de ser un elemento necesario para el funcionamiento del resto de los módulos, el MDE tiene una importancia fundamental en el proceso de generación de poblaciones sintéticas, ya que es (junto con ARDIT) la herramienta que permite obtener los datos brutos a partir de los cuales el DFM generará el objeto matemático (Red Bayesiana) que permite asignar valores numéricos a los atributos de cada agente.

Los siguientes pasos en esta tarea T2.2 son la integración del código del MDE en el despliegue del DWH, y las pruebas funcionales sobre el proceso de generación de la población sintética, una vez que se haya completado el módulo del SPG (futuro entregable D2.4).