



Evaluación y simulación del impacto de las políticas públicas sobre las PYMES a través del uso de modelos basados en agentes

Tomás Mancha
Federico Pablo Martí
Juan Luis Santos
Antonio García Tabuenca
María Teresa del Val
María Teresa Gallo
Ricardo Buendía
Neftis Atallah
Paolo D'Arminio

INSTITUTO UNIVERSITARIO DE ANÁLISIS ECONÓMICO Y SOCIAL

DIRECTOR

Dr. D. Tomás Mancha Navarro

Catedrático de Economía Aplicada, Universidad de Alcalá

DIRECTOR FUNDADOR

Dr. D. Juan R. Cuadrado Roura

Catedrático de Economía Aplicada, Universidad de Alcalá

SUBDIRECTOR

Dr. D. Antonio García Tabuenca

Profesor Titular de Universidad, Universidad de Alcalá

AREAS DE INVESTIGACIÓN

ANÁLISIS TERRITORIAL Y URBANO

Dr. D. Rubén Garrido Yserte

Profesor Titular de Universidad Universidad de Alcalá

ECONOMÍA LABORAL

Dr. D. Carlos Iglesias Fernández

Profesor Contratado Doctor Universidad de Alcalá

ESTUDIOS SECTORIALES, FINANCIEROS Y PYME

Dr. D. Antonio García Tabuenca

Profesor Titular de Universidad Universidad de Alcalá

SERVICIOS E INNOVACIÓN

Dr. D. Juan R. Cuadrado Roura

Catedrático de Economía Aplicada Universidad de Alcalá

RESPONSABILIDAD SOCIAL CORPORATIVA

Dra. Dña. Elena Mañas Alcón

Profesora Titular del Dpto. de Economía Aplicada Universidad de Alcalá



DOCUMENTOS DE TRABAJO

La serie Documentos de Trabajo que edita el Instituto Universitario de Análisis Económico y Social (IAES), incluye avances y resultados de los trabajos de investigación realizados como parte de los programas y proyectos del Instituto y por colaboradores del mismo.

Los Documentos de Trabajo se encuentran disponibles en Internet:

http://www.iaes.es/publicaciones.htm

ISSN: 2172-7856

ÚLTIMOS DOCUMENTOS PUBLICADOS

WP-06/11 EL CUARTO AÑO DE CRISIS FINANCIERA Y ECONÓMICA: VISIÓN INTERNACIONAL Y ESPAÑOLA

Antonio Torrero Mañas

WP-07/11 CO-MOVEMENTS IN TERMS OF TRADE VOLATILITY IN LANDABUNDANT COUNTRIES

Alberto M. Díaz Cafferata y Maria Virginia Mattheus

WP-08/11 EL IMPACTO DE LA MONEDA ÚNICA EN ESPAÑA, DESDE UNA PERSPECTIVA SECTORIAL

Juan Federico Von Zeschau

WP- 01/12 ES TAN 'NUEVA' LA 'NUEVA GEOGRAFÍA ECONÓMICA'? SUS APORTACIONES, SUS LÍMITES Y SU RELACIÓN CON LAS POLÍTICAS Juan Ramón Cuadrado Roura

WP-02/12 LA REFORMA DE LAS PENSIONES EN ESPAÑA: UNA APROXIMACIÓN ECONÓMICA

José M. Domínguez Martínez

WP-03/12 CONTRIBUCIONES DEL DESARROLLO FINANCIERO AL CRECIMIENTO ECONÓMICO: ASPECTOS TEÓRICOS Y EMPÍRICOS DE LARGO PLAZO

Flavio E. Buchieri, Alejandro Pereyra y Tomás Mancha





Plaza de la Victoria, 2. 28802. Alcalá de Henares. Madrid - Telf. (34)918855225 Fax (34)918855211 Email: iaes@iaes.es. WEB: www.iaes.es



EVALUACIÓN Y SIMULACIÓN DEL IMPACTO DE LAS POLÍTICAS PÚBLICAS SOBRE LAS PYMES A TRAVÉS DEL USO DE MODELOS BASADOS EN AGENTES 1

RESUMEN

La utilización de sistemas fundados en ABM (Agent-based Models) para modelizar la realidad socio-económica permite realizar predicciones y simulaciones basadas en los atributos y el comportamiento de cada uno de los agentes individuales. El sistema MOSIPS implementará un entorno multiagente a partir de datos de multitud de fuentes públicas y abiertas, datos que construirán el Data Warehouse a partir del cual se alimentará al Motor de Simulaciones y Predicciones (SFE). Mediante la utilización de tecnologías de visualización espacio-temporales se desarrollará un interfaz de usuario avanzado para facilitar la toma de decisiones. Además, con el fin de fomentar la participación activa de los ciudadanos en el diseño de las políticas públicas, se desarrollarán módulos que permitan la integración de MOSIPS en las redes sociales y la recogida de la retroalimentación que los usuarios proporcionen.

Palabras clave: Modelos basados en agentes, Predicción, Simulación, Evaluación de políticas.

ABSTRACT

The use of systems with ABM (Agent-based Models) foundations in order to modelling the real economy allows making useful forecast and simulations based on the attributes and behaviour of each particular agent. The MOSIPS system implements the multiagent environments from multiple both open and specific data sources. These data will build the Data Warehouse and feed the Simulations and Predictions Engine (SFE). The spatio-temporal visualization technologies will be used in order to develop an advanced user interface to support the decision-making. Furthermore, to encourage active participation of citizens in the public policy design, a module will be developed to allow the MOSIPS integration with social networks and to gather the user feedback. Key words: Agent-based models, Simulation and Forecasting, Policy evaluation.

AUTORES:

TOMÁS MANCHA. IAES. tomas.mancha@uah.es
FEDERICO PABLO MARTÍ. IAES. federico.pablo@uah.es
JUAN LUIS SANTOS. IAES. juanluis.santos@iaes.es
ANTONIO GARCÍA TABUENCA. IAES. antonio.gtabuenca@uah.es
MARÍA TERESA DEL VAL. IDOE. mteresa.val@uah.es
MARÍA TERESA GALLO. IAES. maria.gallo@uah.es
RICARDO BUENDÍA. ANOVA IT. ricardo.buendia@anovagroup.es
NEFTIS ATALLAH. ANOVA IT. neftis.atallah@anovagroup.es
PAOLO D'ARMINIO. ANOVA IT. paolo.darminio@anovagroup.es

¹ Los resultados recogidos en este documento provienen del proyecto MOSIPS (Modeling and Simulation of the Impact of Public Policies on SMEs), cofinanciado por el VII Programa Marco de Investigación de la Comisión Europea bajo el contrato 288833 y en el participa un consorcio formado por los siguientes socios: ANOVA IT Consulting (España, coordinador), Instituto Universitario de Análisis Económico y Social de la Universidad de Alcalá (España), Research Studio Austria Forschungsgesellschaft (Austria), Universidad de Reading (Reino Unido), Opera21 (Italia), Universidad de Konstanz (Alemania), European Institute of Interdisciplinary Research (Francia), Ayuntamiento de Madrid (España) y Comune di Verona (Italia)



ÍNDICE

1.	Introducción	6
2.	Modelización de la realidad socio-económica en el sistema MOSIPS	7
	2.1. Modelización de los agentes	.14
	3.1. Capa de recogida de datos3.2. Capa de simulación de políticas3.3. Capa de visualización3.4. Capa de interacción de usuarios	. 18 . 19
4.	Conclusiones	.22
5.	Referencias bibliográficas	.23



1. Introducción

n 2008, la Comisión Europea aprobó la "Small Business Act (SBA)" para los países miembros de la Unión Europea (Comisión Europea, 2008). Esta iniciativa establece un marco estratégico para las políticas orientadas a las PYMEs con el objetivo de potenciar su crecimiento y su actividad innovadora. Además pone énfasis en la necesidad de reconocer el papel esencial de la pequeña y mediana empresa en la economía europea, mostrando cómo la actividad económica de las PYMEs repercute, por la propia interacción entre ellas, en el aumento de la productividad, la innovación, la creación de empleo y la integración social y regional en Europa.

Existen diferentes enfoques que permiten simular el impacto socioeconómico de las intervenciones públicas en la sociedad, pero ninguna solución satisface plenamente las necesidades de predicción de los gobiernos europeos debido a la enorme complejidad de los modelos de simulación utilizados; la falta de modelos de comportamiento humano suficientemente avanzados; y en muchos casos, la ausencia de datos suficientemente fiables o desagregados.

En la actualidad, la mayoría de los modelos utilizados para el tratamiento de los fenómenos económicos agregados se basan en modelos de equilibrio general dinámico estocástico DSGE (Dynamic Stochastic General Equilibrium), modelos que aceptan la hipótesis de que existe un equilibrio preestablecido hacia el que convergen los mercados y permiten fluctuaciones de cierta dimensión en torno a ese equilibrio. Este hecho hace que los modelos DGSE no sean los más adecuados en las situaciones de crisis económica en las que no se alcanza ese estado de equilibrio en el corto ni medio plazo (Doyne, 2009). Además, por su concepción, con estos modelos no pueden observarse con un suficiente nivel de desagregación los efectos de las políticas públicas sobre el espacio y sobre los distintos tipos de personas y empresas, de acuerdo, por ejemplo, a su nivel de renta.

Ante este escenario, surge la necesidad de encontrar una nueva metodología para simular y predecir el impacto de las políticas públicas sobre las PYMEs.

Los modelos ABM (Agent-Based Models) aparecen como una opción prometedora capaz de hacer predicciones basadas en una modelización de la economía real al nivel más bajo posible, es decir, considerando el impacto sobre cada uno de los agentes en función de sus características. Los modelos ABM se basan en no asumir que la economía puede alcanzar un equilibrio establecido, permitiendo describir los comportamientos no lineales de los agentes económicos de una forma más cercana a la realidad (Mulligan, 1999).



En el marco del proyecto MOSIPS (Modeling and Simulation of the Impact of Public Policies on SMEs), el consorcio desarrollará un simulador de proyecciones socio-económicas, basado en sistemas multiagente y combinando diferentes tecnologías, como herramienta de soporte a la decisión para el diseño de políticas públicas. Dicha herramienta permitirá simular y visualizar ex-ante el impacto socio-económico que diferentes diseños políticos pueden tener sobre las PYMEs localizadas en un territorio determinado. Adicionalmente, a través de su integración con las principales redes sociales, se incorporarán opiniones y comentarios procedentes de la ciudadanía, contribuyendo así a una mayor transparencia de la actividad política y haciendo posible la participación del ciudadano en los procesos de decisión.

El presente documento de trabajo se estructura en cuatro secciones. Tras esta introducción, en la segunda sección se describe a nivel funcional la modelización de los agentes y de las políticas en las que se basará el simulador MOSIPS. En la sección tres se discute brevemente la arquitectura del sistema, detallando los componentes a implementar y las tecnologías a ser utilizadas. Finalmente, en la última parte se presentan las principales conclusiones.

2. MODELIZACIÓN DE LA REALIDAD SOCIO-ECONÓMICA EN EL SISTEMA MOSIPS

I sistema MOSIPS afronta el ambicioso reto de crear un modelo económico basado en agentes capaz de hacer predicciones sobre la economía real de una región, desarrollando un simulador de tipo ABM (Agent Based Model). La combinación de modelos teóricos con fuentes de datos públicas y tecnologías de inteligencia artificial permitirá la creación de un software capaz de someter a análisis los efectos de las diferentes las medidas de políticas públicas, y de este modo ayudar a tomar decisiones que permitan alcanzar un mayor nivel de bienestar.

Para el desarrollo práctico de este planteamiento, el proyecto MOSIPS propone un avance en el estado del arte en el ámbito de los sistemas de simulación multiagente para la modelización de diferentes escenarios y diseños socio-económicos.

El simulador MOSIPS permite modelizar la economía como un sistema complejo que no ha de converger necesariamente hacia un estado de equilibrio, permitiendo incorporar los comportamientos no lineales de los agentes económicos, acercando el modelo a la realidad. Así, se creará un "universo virtual" en el que las diferentes partes interesadas pueden interactuar de manera realista. En este sistema se plantea la utilización de la psicología aplicada y la economía del comportamiento para la definición del conjunto de normas conductuales de los agentes individuales del modelo, y de la teoría microeconómica y de otras



ciencias sociales aplicadas para elaborar las relaciones y reglas de interacción ere ellos. Se utilizan también metodologías dirigidas a la extracción, tratamiento y clasificación de los datos socio-económicos necesarios para complementar la modelización del entorno, incluyendo datos públicos publicados bajo iniciativas de Open Data que hagan posible la réplica del sistema en otros escenarios.

2.1. Modelización de los agentes

En MOSIPS, cada uno de los agentes actúa según su estado actual, el estado del entorno que lo rodea y las normas que rigen su comportamiento. El sistema monitoriza en el tiempo el estado y las consecuencias de las interacciones entre todos los agentes y el entorno macroeconómico, permitiendo la simulación de una especie de "economía artificial", pero estrechamente relacionado con la realidad que ha permitido generarla, bajo diferentes diseños socio-económicos y la consiguiente exploración cuantitativa de las consecuencias de una actuación política.

En el sistema MOSIPS se definen dos macrocategorías de agentes, descritas cada una de ellas por una fuente principal de información y otras fuentes complementarias que permitirán añadir información específica fundamental para el correcto funcionamiento del modelo:

Familias - Individuos:

- Son los propietarios y los empleados de las PYMEs. Además, son los agentes que consumen bienes y servicios.
- Fuente principal de información: Censo de población. Esta fuente deberá actualizarse mediante el uso de indicadores estadísticos y los microdatos deberán ser inferidos al total de la población.
- Fuentes complementarias: Encuesta de población activa (EPA), Encuestas de movilidad, Encuesta de presupuestos familiares y Panel de declarantes de IRPF.

Empresas – Establecimientos:

- Son los agentes que producen bienes y servicios. Además, emplean a ciudadanos, pagándoles salarios como contraprestación a su trabajo.
- > Fuente principal: Directorio Central de Empresas (DIRCE).
- ➤ Fuentes complementarias: Base de Datos Amadeus, Base de Datos BACH (Bank for the Accounts of Companies Harmonised), Panel de Innovación Tecnológica, Encuesta de Acceso a la Financiación.

Dada la complejidad de la modelización propuesta, la única solución plausible para integrar los datos de las fuentes principales y de las



fuentes complementarias es el uso de técnicas de fusión estadística (statistical matching) con el fin de estandarizar la información. El papel de las fuentes complementarias es básicamente el de permitir completar los datos de la fuente principal aumentando así la precisión del modelo en un área específica. Por ejemplo, la información recogida en la Encuesta de Población Activa permite ampliar la información del Censo con otra sobre variables de empleo; y las Encuestas de movilidad hacen al sistema capaz de conocer los desplazamientos de los agentes en el territorio.

Por lo tanto, de acuerdo con el modelo propuesto habrá dos procesos de fusión estadística, uno para los individuos-familias y otro para las empresas-establecimientos. No obstante, este proceso necesita un paso adicional: ambas bases de datos deben estar vinculadas, cada empresa debe tener identificados sus empleados en la base de datos de individuos y los dueños de las empresas deben tener identificadas las empresas de su propiedad en la base datos de empresas.

Para conseguir un buen nivel de exactitud en este proceso, la localización de los individuos y de las empresas juega un papel clave, por lo que es necesaria la utilización de técnicas de escalaje (downscaling techniques). Estas técnicas, ampliamente utilizadas para los modelos climáticos, permitirán obtener información territorial de alta resolución a partir de diferentes fuentes de información. MOSIPS propone utilizar como fuente principal CORINE (Coordination of Information on the Environment) que permite conocer la densidad de población o el uso del suelo. Para conseguir un mayor nivel de precisión de la población y de las PYMEs en un área determinada, dicha información deberá ser complementada con bases de datos locales que aporten información sobre la localización de las empresas, zonas despobladas, desigualdades de ingresos a lo largo de las ciudades, etc.

De acuerdo con el modelo descrito previamente y las fuentes de datos disponibles, la a continuación se resumen los requisitos de datos para el sistema MOSIPS.

2.1.1. Creación de la base de datos para individuos y familias

El desarrollo de la base de datos para los individuos utiliza como punto de partida el Censo de Población. Como el correspondiente al año 2011 todavía se encuentra en proceso de validación y aún no está disponible, se recogen los datos del último censo disponible, correspondiente a 2001. El mismo, por razones de secreto estadístico, solo ofrece información exhaustiva (microdatos) para el 5% de la población; lo que obliga a la creación del resto de la población mediante un procedimiento de elevación de la información suministrada.



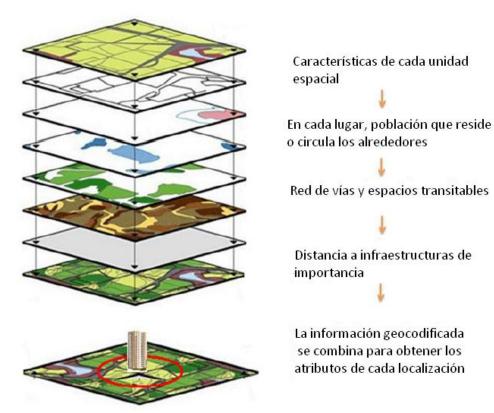


FIGURA 1. Esquema de análisis para identificar el entorno de las localizaciones.

Fuente: Mejía-Dorantes (2011).

Además, el Censo ofrece información correspondiente sólo al último trimestre de 2001, pero para hacer operativo el modelo –en un sentido dinámico- es necesario conseguir un fotograma de la población para cada uno de los trimestres comprendidos entre el último de 2007 y el más cercano a la fecha actual; todo ello con la finalidad de conseguir una especie de película de la evolución de la población, y así poder lograr tanto parametrizar como validar el modelo.

La implementación de este proceso exige, entre otros, los siguientes pasos: actualización de las edades, identificación de los decesos, los nacimientos, las migraciones, la creación de parejas y los ceses de relaciones maritales. Además, cada miembro de la población puede haber alterado su nivel de estudios y cambiado su municipio de residencia.



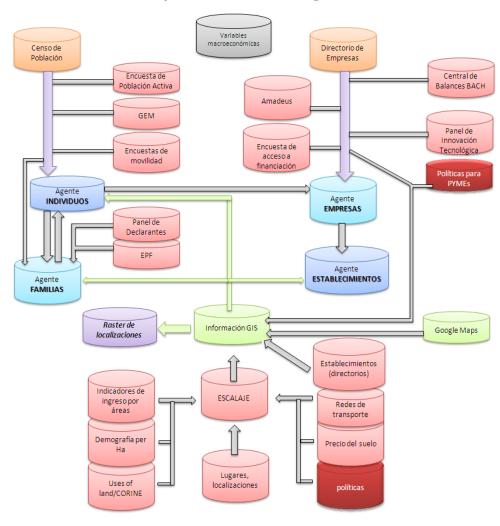


FIGURA 2. **Bases de Datos para el modelado de agentes en MOSIPS**

Fuente: Elaboración propia

La fuente principal de información, Censo de población, se complementa con las siguientes fuentes complementarias:

1. La Encuesta de Población Activa: De periodicidad trimestral, nos permite conocer la situación laboral de cada uno de los individuos mayores de 16 años de acuerdo a sus características. El proceso de fusión estadística de esta fuente con el Censo de población requiere un trabajo de gran complejidad y precisión, ya que llevarla a cabo correctamente implica que no pueda hacerse trimestre a trimestre pues se incurriría en diversos tipos de errores. Por ejemplo, un funcionario podría pasar a ser desempleado de larga duración en el periodo siguiente; o bien que un inactivo en un periodo refiera llevar varios años



trabajando en el periodo siguiente. Para evitar estos errores, el proceso de fusión debe llevarse a cabo secuencialmente, en un proceso cada vez más complejo debido a la proliferación de variables de fusión. Otra posibilidad sería recurrir a la elaboración de un modelo de microsimulación que tenga en cuenta el presente de cada individuo para construir su pasado de acuerdo a una senda de posibles estados en el pasado.

- 2. El <u>Global Entrepreneurship Monitor</u> aporta variables relacionadas con el emprendimiento, foco sobre el que actúan las políticas a medio y largo plazo, con el objetivo de fomentar la creación de nuevas empresas y el avance tecnológico.
- 3. Las <u>Encuestas de movilidad</u> nos permiten conocer los patrones de desplazamiento de los individuos, ya que no solo se consume en el área cercana al lugar de residencia, sino que el lugar de trabajo es también importante, así como otros viajes que se llevan a cabo con una periodicidad determinada. Sirva de ejemplo la profusión de establecimientos comerciales en el centro de grandes ciudades, cuyos niveles de demanda no podría explicarse teniendo en cuenta los residentes en ese área. En la figura 3 se observa la importancia de los patrones de movilidad de una familia en la que uno de sus miembros trabaja en una localidad cercana a su hogar. Observamos que el consumo no se produce en los mismos establecimientos. Si esta familia perdiera su fuente de ingresos no solo se vería afectada el área en la que uno de sus miembros trabaja o la localidad de residencia, también la zona intermedia vería reducida la demanda de bienes. Observar estos efectos solo es posible a través de la correcta construcción de los modelos basados en agentes.
- 4. <u>Panel de Declarantes del IRPF</u>: Proporciona información sobre los ingresos de los individuos.
- 5. Encuesta de Presupuestos Familiares: Facilita información sobre cómo los individuos distribuyen su renta entre consumo y ahorro, permitiendo conocer cómo afectaría una medida que afecta a los patrones de consumo a las diferentes empresas.



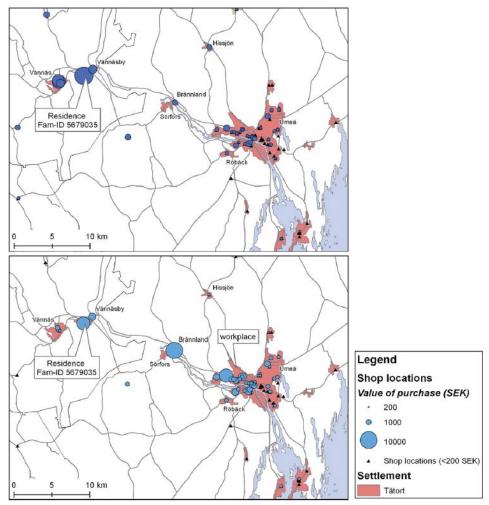


FIGURA 3.

Distribución del gasto de una familia antes y después de localización del lugar de trabajo.

Fuente: Tomado de Schenk et.al (2007)

2.1.2. Creación de la base de datos para empresas y establecimientos

La construcción de la base de datos de empresas requiere un proceso similar, aunque con algunas particularidades.

- 1. En primer lugar, el <u>Directorio Central de Empresas</u> incluye muy pocas variables, y no ofrece microdatos. Sin embargo, es la fuente que nos aporta el tamaño de la población de empresas, dividida en distintos tamaños y sectores.
- 2. Dada la escasa información desagregada de esta base, es imprescindible completarla con la base de datos <u>Amadeus</u>, que recoge información individualizada de una parte importante de



las empresas europeas. Uniendo cada empresa de la población con una de las empresas incluidas en esta base obtenemos una gran cantidad de indicadores.

- 3. Esta base de datos se complementa con otros indicadores obtenidos de la Base de Datos BACH (<u>Bank for the Accounts of Companies Harmonised</u>). Esta recoge los balances y la cuenta de ingresos y gastos de empresas y establecimientos, permitiéndonos comprobar la precisión de los datos de la población generada, y ayudando a corregir posibles desviaciones respecto a esta otra fuente.
- 4. El <u>Panel de Innovación Tecnológica</u> monitoriza la actividad de innovación de las empresas desde 2004. El panel estudia el comportamiento innovador de las compañías clasificadas en submuestras de acuerdo al número de trabajadores y se observan las actividades en el seno de la empresa, las externalizadas y la incorporación de innovaciones sin coste.
- 5. Por último, la Encuesta de Acceso a la Financiación provee información de los problemas y restricciones con los que las compañías se enfrentan a la hora de conseguir fondos ajenos. Se analizan las distintas alternativas de financiación, en el sector industrial, de servicios y construcción, y también se incluyen expectativas sobre el futuro de este factor crucial en el crecimiento y la creación de empresas, en especial las de pequeño tamaño y/o reciente constitución.

2.2. Modelización de las políticas

Otros actores que deben incluirse en el modelo son el sector financiero, el sector público y el comercio exterior. Las variables macroeconómicas más relevantes para el modelo se encuentran en los Institutos Nacionales de Estadística y están relacionadas con el PIB, los impuestos, las importaciones y exportaciones, el mercado de trabajo o el presupuesto público. Estos datos permitirán conocer la evolución del entorno económico desde la perspectiva de la PYME. También es importante conocer cómo el sector financiero afecta a las decisiones de PYMEs y ciudadanos, así que tendremos en cuenta el tipo de interés, la tasa de préstamos impagados y otros indicadores del entorno como la variación del mercado de valores. Además, para valorar las relaciones intersectoriales se usarán tablas insumo-producto en un nivel regional, el mínimo disponible actualmente.

Existe una ingente cantidad de políticas públicas que afectan directa o indirectamente a las PYMEs, hecho que imposibilita el desarrollo de un modelo que permita definir individualmente cada política. Por este motivo, será necesario clasificarlas en categorías que traten de cubrir el mayor número de posibilidades. Basándonos en los principios y políticas definidos en la "Small Business Act (SBA)", mencionada anteriormente, se han identificado los siguientes ámbitos políticos que serán



implementados en la modelización de MOSIPS –ver figura 4-: entorno macroeconómico, innovación, mercado laboral, actividad emprendedora, gestión interna, financiación, infraestructuras, relaciones entre empresas, relaciones con la administración y medio ambiente.

FIGURA 4. Ámbitos políticos implementados en el modelo MOSIPS



Fuente: Elaboración propia

El propósito de esta categorización es incluir todas las posibles áreas en las que las políticas afectan a las PYMEs. Además, estas categorías se adaptan fácilmente a los principios de la "Small Business Act", reforzando el grado de generalidad del proyecto. Asimismo, permite tratar sistemáticamente la inabarcable cantidad de instrumentos que existen a disposición de los policy-makers y que tienen efectos sobre las PYMEs.

3. LA ARQUITECTURA DE MOSIPS

- a plataforma MOSIPS ha sido concebida a alto nivel como un conjunto de capas distribuidas y fuertemente desacopladas, tal—como se recoge en la figura 5 y se explica a continuación.
- Capa de recogida de datos: Identificación, recolección, transformación, procesamiento y homogenización de los datos públicos disponibles en un formato único y normalizado. Se implementará una API que permitirá a otros componentes del sistema realizar consultas específicas para la ejecución de simulaciones concretas.



- Capa de simulación de políticas: Incluye el motor de simulación y previsión (SFE, Simulation and Forecast Engine). Responsable de consultar los datos necesarios para alimentar un escenario particular y de ejecutar el proceso de simulación según los parámetros indicados por el usuario final. Además, almacenará los resultados de las simulaciones.
- Capa de exposición de Open Data: Hace disponible los datos normalizados y los resultados de las simulaciones como Open Data, siguiendo las recomendaciones y buenas prácticas de Open Government Data.
- Capa de interacción de usuarios: Solicita al usuario los parámetros para las simulaciones y permite visualizar los resultados del proceso bajo técnicas de Análisis Visual. Cuando sea necesario, almacenará localmente resultados de simulaciones para una visualización interactiva eficiente. Además, es responsable de la integración con las redes sociales.

Capa de recolección de datos

Subsistema de repositorio de datos

Subsistema de recolección de datos

Capa de simulación de políticas

Diseño del modelo de simulación

Ejecución del modelo de simulación

Resultados de la simulación

Feedback de usuarios

Capa de interacción de usuarios

Arquitectura

FIGURA 5. **Arquitectura de capas distribuidas del sistema MOSIPS**

Fuente: Elaboración propia

Cada una de las capas implementará los componentes e interfaces necesarios para la ejecución de sus funcionalidades.

En las siguientes secciones se procederá a describir las diferentes capas con un mayor detalle.



CUADRO 1. **Principales componentes de cada capa MOSIPS**

Capa de la arquitectura	Componentes	Funcionalidad
Recogida de datos	Repositorio de datos	Almacenar y pre-procesar los datos extraídos de fuentes externas
	Recogida de datos	Ofrecer los datos al resto del sistema en un formato normalizado
Simulación de políticas	Diseño del modelo de simulación	Crear los modelos de simulación a partir de datos extraídos del sistema
	Ejecución del modelo de simulación	Ejecutar y almacenar las simulaciones parametrizadas
Exposición de Open Data	Exposición de Datos	Exponer los datos del repositorio MOSIPS como Open Data
	Exposición del modelo y de resultados de previsiones	Exponer los resultados de las simulaciones como Open Data
Interacción con el usuario	Visualización	Visualizar los resultados de las simulaciones
	Interacción con el modelo	Crear el interfaz para la interacción del usuario con el sistema
	Apps de redes sociales	Recoger la retroalimentación de los usuarios a partir de las redes sociales

Fuente: Elaboración propia

3.1. Capa de recogida de datos

La extracción y el procesado de los datos de las diversas fuentes externas se lleva a cabo separadamente.

El subsistema de Repositorio de Datos, implementado mediante un Data Warehouse, se basará en 4 módulos tal y como propone el modelo teórico de agentes:

- Módulo de extracción: Permite acceder y recoger todos los datos de las diferentes fuentes heterogéneas. Se implementarán distintos conectores para cada una de las fuentes de datos, típicamente basadas en formatos txt o xls, existiendo un conector específico para la integración de Open Data en cuyo caso se utilizará una API SPARQL.
- 2. Módulo de transformación e integración: Posibilita realizar diferentes operaciones sobre los datos como limpieza de datos, filtrado y estructuración, fusión estadística, técnicas de escalaje, en función de las necesidades del modelo teórico.
- 3. Módulo de Data Warehouse: Permite modelar y almacenar todos los datos integrados siguiendo un esquema homogéneo.



 Módulo de Data Marts: Posibilita modelar estructuras de datos específicas para las potenciales necesidades de datos. Cada Data Mart almacenará los datos que alimentan cada categoría de agentes (Individuos-Familias, Empresas-Establecimientos).

El subsistema de Recogida de Datos será un middleware que gestione las comunicaciones entre el Repositorio de Datos y la capa de Simulación de Políticas de MOSIPS. Un interfaz DIM DB será implementado para recuperar los datos del repositorio a partir de una solicitud de información recibida. El formateo de los datos según el protocolo Key-Value compartido entre las dos capas también será realizado por este subsistema.

En esta capa, se utilizarán las siguientes tecnologías:

- Para extracción/filtrado/limpieza/carga de datos: herramientas open source ETL como KETTLE.
- Para el Data Warehouse se utilizarán tecnologías robustas como Oracle DBMS, cuyo uso es gratuito para proyectos de investigación.
- Para Data Mining y Data Fusion/Integration se utilizarán herramientas como WEKA, R Environment (para fusión estadística).
- Para la integración de los diferentes módulos puede necesitarse la utilización de lenguajes de programación abiertos como Java.

3.2. Capa de simulación de políticas

El modelado de sistemas basado en agentes (ABM) es una potente técnica de simulación de modelos que permite modelar un sistema como un conjunto de entidades que toman decisiones de forma autónoma. Cada agente evalúa individualmente su situación y toma decisiones en base a un conjunto de reglas establecido. Los agentes pueden ejecutar distintos comportamientos definidos por el sistema, por ejemplo, producir, consumir o vender. Además, los agentes interaccionan entre sí pudiendo evolucionar hacia comportamientos imprevistos. Habitualmente, los sistemas ABM incorporan técnicas de aprendizaje como algoritmos evolutivos o de inteligencia artificial con el fin de permitir una evolución más realista.

Para empezar con el proceso de simulación, es necesario disponer de tres elementos:

- Modelos de simulación: Modelado de cada uno de los agentes individuales y de los grupos de agentes.
- > Datos de simulación: Inicio de los individuos para un escenario concreto definido por el usuario.



Parámetros de control controlados por el usuario: Los usuarios podrán cambiar los parámetros del modelo y ejecutar nuevas simulaciones.

FIGURA 6. **Proceso de simulación basado en agentes de MOSIPS**



Fuente: Elaboración propia

Para la comunicación entre los distintos componentes de la capa de simulación de políticas se desarrollarán diferentes APIs basadas en servicios REST para facilitar la integración con cualquier interfaz de usuario. Para la puesta en funcionamiento del motor de simulación se utilizarán tecnologías open sources basadas en Java como Repast, Netlogo o MASON. Las Bases de Datos que almacenen los resultados de las ejecuciones y el feedback de los usuarios serán BBDD relacionales y se utilizarán tecnologías como Linked Data para la exposición de los datos.

3.3. Capa de visualización

La capa de visualización estará formada por dos componentes, una para cada tipología de usuario: analistas y público general. Mientras que el público general podrá acceder a una versión simplificada del simulador, con un interfaz sencillo que permitirá acceder a subconjuntos de datos, los analistas accederán al sistema a través de un interfaz basado en técnicas avanzadas de visualización que permitirá realizar análisis e interpretaciones de grandes cantidades de datos espacio-temporales (Keim et al., 2008).

La capa de visualización estará formada por cuatro módulos:

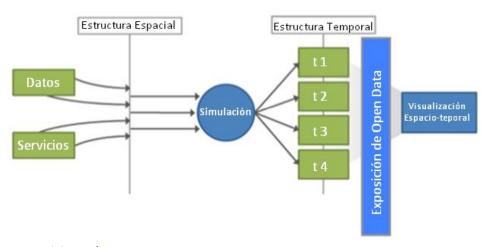
- > Módulo de integración de datos: responsable de integrar los datos del Data Warehouse en el sistema de visualización.
- Módulo de representación visual: responsable de elegir la asignación adecuada de los datos para su visualización.



- Módulo de representación: responsable de la conversión de la representación visual a una imagen renderizada.
- Módulo de interacción: responsable de recoger los parámetros de visualización del usuario y transferirlos al módulo de representación visual para adaptarla en espacio-tiempo.

FIGURA 7.

Integración de resultados espacio-temporales de la simulación en el módulo de visualización.



Fuente: Elaboración propia

Tal y como muestra la figura 7 en la capa de visualización se integran dos estructuras:

- Resultados espacio-temporales de la simulación. En el módulo de visualización se integrarán los resultados de simulaciones para distintos intervalos temporales. Para ello, en función de la granularidad de la resolución temporal de las diferentes instantáneas ejecutadas del modelo, se implementarán esquemas URI, dos estructuras (Uniform Resource Identifier) que incluyan los espacios temporales o se utilizarán técnicas de filtrado mediante consultas SPARQL.
- Información geográfica adicional. Para realizar el pre-renderizado de la imagen se utilizará información geográfica adicional, no necesariamente utilizada en el proceso de simulación, que enriquecerá la visualización incorporando información que facilite la comprensión del contexto de los resultados de las previsiones.

Para la implementación de la capa de visualización se utilizarán modelos como Google Maps para el interfaz de usuario orientado al público general, módulo integrable en las redes sociales, y se estudiará la viabilidad de utilizar tecnologías client-side como AJAX. No obstante, el interfaz de usuario de los usuarios avanzados, analistas, deberá incluir



componentes específicos que permitan configurar temporalmente los resultados de las simulaciones y realizar filtrados de variables.

3.4. Capa de interacción de usuarios

Esta capa servirá para recoger los inputs de los usuarios: valores de las variables paramétricas para ejecutar las simulaciones y la retroalimentación obtenida sobre los resultados. Para el público general, dicha interacción ser llevará a cabo a través de un módulo integrado en las redes sociales permitiendo a los usuarios retroalimentar el sistema aportando comentarios y discusiones a nivel individual o colectivo. Por lo tanto, la capa de interacción implementará las siguientes funciones:

- > Obtener los datos del Motor de Simulación y Predicción (SFE) para mostrar a los usuarios distintos entornos.
- Solicitar a la capa de Simulación de Políticas resultados para diferentes simulaciones.
- Enviar la retroalimentación obtenida de los usuarios a la capa de Simulación de Políticas.
- Enviar información a la capa de visualización sobre los modelos de simulación y la retroalimentación de los usuarios para su representación gráfica.

Para recoger la retroalimentación de una discusión estructurada o la interacción entre un grupo de usuarios en torno a unos resultados de previsión concretos, se utilizarán técnicas como Delphi, una metodología de investigación multidisciplinar para la realización de pronósticos y predicciones (Gordon, 1994; Delbecq et al., 1975; Brancheau et al., 1996), que implementa un modelo de participación convergente para la evaluación de los pronósticos.

El método Delphi es una técnica que permite realizar predicciones basadas en el consenso de las opiniones de los expertos. Se ajustan muy bien a los modelos de discusión informal como los desarrollados en el entorno de una red social, combinado con una evaluación de los resultados de la ejecución de la simulación. Estos métodos suelen realizar varias iteraciones del modelo hasta llegar al consenso y la consecución de las políticas óptimas para el grupo de evaluadores. Para el modelado de los workflows se utilizará la notación BPMN (Business Process Modeling Notation) y el lenguaje utilizado sería WS-BPEL (Web Services – Business Process Execution Language).

Respecto al funcionamiento del sistema en las redes sociales Se deberán implementar dos componentes:

Interfaz de Información del Modelo, interfaz independiente de las redes sociales, que permitirá comunicarse con la capa de



- Simulación de Políticas. Como en principio no va a ser un módulo distribuido, no requerirá un interfaz orientado a servicios.
- Servicio de Interacción, dependiente de la red social, responsable de la gestión de los inputs de usuario y de llamar al componente de visualización cuando sea necesario.

4. CONCLUSIONES

I principal objetivo del sistema MOSIPS aquí expuesto es diseñar y desarrollar una herramienta de soporte a la decisión, que permita mejorar significativamente la calidad de las políticas implementadas por las autoridades públicas a distintos niveles (local, regional, nacional,...). Adicionalmente, facilitará información de interés a los ciudadanos sobre las consecuencias y el impacto de políticas que puedan ser establecidas en su localidad, fomentando así su compromiso y participación activa en el proceso de formulación de dichas políticas.

MOSIPS se basa en sistemas multiagente para modelizar de la realidad socio-económica, pudiendo definir los atributos y el comportamiento de cada uno de los agentes del sistema: Individuos/Familias y Empresas/Establecimientos. Además, se implementarán distintas políticas parametrizables, basadas en la "Small Business Act (SBA)", las cuales permitirán a los analistas simular el impacto de estas políticas, principalmente en referencia a las PYMEs.

Por su parte, los modelos ABM permiten corregir algunos de los errores de los sistemas utilizados en la actualidad, ya que describen los comportamientos no lineales de los agentes económicos, basados en las redes de contactos entre ellos, y no solo en sus atributos. Las predicciones realizadas se fundamentan en la modelización de la economía integrada por agentes individuales, cada uno con atributos y comportamientos particulares.

En definitiva, a partir de este tipo de modelización se espera alcanzar unos resultados más consistentes con la realidad, dado que los fundamentos de todos los agentes que intervienen en el funcionamiento del sistema se han diseñado teniendo en cuenta su heterogeneidad, incorporándola tanto en sus atributos como en la caracterización de sus comportamientos.

La metodología aquí presentada se contrapone a la que se venía utilizando, y aún se utiliza, que adopta asunciones como la del agente representativo, que permite obtener resultados de carácter macroeconómico, y que a pesar de utilizar fundamentación microeconómica, pretende alcanzar situaciones de equilibrio que contrastan fuertemente con la realidad observada.



5. REFERENCIAS BIBLIOGRÁFICAS

- Brancheau, J.C.; Janz, B.D. and Wetherbe, J.C. (1996). Key issues in information systems management: 1994-95 SIM Delphi results. MIS Quarterly, vol. 20 (2), pp. 225- 242
- Delbecq, A.L.; Van de Ven, A.H. and Gustafson, D.H. (1975). Group techniques for program planning: A guide to nominal group and Delphi processes. Glenview, Ill.: Scott, Foresman.
- DOYNE FARMER, D.F.J. (2009). The economy needs agent-based modeling. Nature, vol. 460, no 6, pp 685-686.
- EUROPEAN COMMISSION (2008). Small Business Act for Europe., COM (2008) 394 final. http://ec.europa.eu/enterprise/policies/sme/small-business-act
- GORDON, T.J. (1994). The Delphi Method. Futures Research Methodology v3.0. The Millennium Project.
- KEIM, D.; ANDRIENKO, G.; FEKETE, J.D. ET AL. (2008). Visual Analytics: Definition, Process and Challenges. Information Visualization. LNCS, vol. 4950, pp. 154 175. Springer, Heidelberg.
- MEJÍA-DORANTES, L. (2011) Transportation infrastructure impacts on house prices and firms' location. PhD. Thesis.
- MULLIGAN, P. (1999). Differentiating Service Tasks for IT application. An exploratory analysis in financial services. International Journal of Service Industry Management, Vol. 10 Iss: 2, pp.190 212
- SCHENK, T., LÖFFLER, G., RAUH, J. (2007) Agent-based simulation of consumer behaviour in grocery shopping on a regional level. Journal of Bussiness Research, 60, pp.894-903



IAES – Instituto Universitario de Análisis Económico y Social de la Universidad de Alcalá de Henares, Plaza de la Victoria, 2, 28802, Alcalá de Henares (Madrid), España.

IDOE – Instituto de Dirección y Organización de Empresas de la Universidad de Alcalá de Henares, Plaza de la Victoria, 2, 28802, Alcalá de Henares (Madrid), España.

ANOVA IT Consulting, Avda. Punto Mobi, 4 - Parque Científico Tecnológico de la Universidad de Alcalá, 28805, Alcalá de Henares (Madrid), España.



