

# Análisis exploratorio de patentes relacionadas a productos de Instacrops

Catalina Perez-Garcia

Agosto 2021



INSTASOIL - Monitoreo de Suelo

## 1. Introducción

Una patente es un conjunto de derechos exclusivos concedidos por un Estado al inventor de un nuevo producto o tecnología, susceptibles de ser explotados comercialmente por un período limitado de tiempo, a cambio de la divulgación de la invención. En un mundo globalizado, donde el éxito de desarrollos tecnológicos depende significativamente de la colaboración de los distintos actores, analizar los datos de las patentes tiene un potencial aún inexplorado. Las herramientas para realizar distintos análisis de la información disponible, podría permitir mejorar la toma de decisiones a todo tipo de organizaciones. El número de patentes aprobadas en el mundo crece rápidamente cada año, por lo que se requieren análisis de patentes que extraigan efectivamente la información valiosa contenida en ellas.

En particular para Instacrops, el análisis de las patentes relacionadas existentes en el mundo entrega información valiosa sobre la competencia existente para los productos productivos o en desarrollo de la empresa. Esto es fundamental para evaluar el mérito innovativo de las soluciones creadas para desarrollar estrategias de patentabilidad. Por otro lado, será posible evaluar si existen espacios para nuevos desarrollos vinculados a la tecnología vinculada.

## 2. Base de datos inicial

PatentScope es la base de datos pública de WIPO, la Organización Mundial de Propiedad Intelectual por sus siglas en inglés. Incluye la cobertura de las solicitudes del Tratado de Cooperación en materia de Patentes (administradas por WIPO) y una amplia gama de otros países, incluida la Oficina Europea de Patentes, la USPTO y Japón. Actualmente la base cuenta con más de 97 millones de documentos de patentes. Para este análisis de patentes se utilizarán distintas bases descargadas desde PatentScope.

Para todos los análisis, se filtran las patentes para dejar solo aquellas desde el año 2000 hasta la fecha según el año de publicación. Se limpian las bases descargadas borrando missing values, cambiando formatos en las columnas según corresponda, eliminando duplicados, entre otras cosas, para dejar la base de datos lista para realizar análisis.

En la siguiente tabla se ve la línea de negocio, el concepto, la búsqueda realizada en PatentScope y la cantidad de filas antes y después de la limpieza realizada.

Cuadro 1: Descripción de la base utilizada

Concepto	Monitoreo de Suelo
Línea de negocio	InstaSoil
Query	FP:(soil) AND ((crop) OR (farm)) AND ((catheter) OR (probe)) AND (monitoring)
N de patentes pre-procesamiento	457
N de patentes post-procesamiento	214

Esta base inicial sirve para realizar una primera iteración para obtener aprendizajes necesarios para cumplir con el objetivo de encontrar las patentes más relacionadas al servicio a estudiar. En la

siguiente tabla se muestra un ejemplo con la información de contenida en cada una las columnas de la base de datos.

Cuadro 2: Ejemplo de una fila en la base de patentes usada

Columnas	Ejemplo
Application Id	US302854360
Application Number	16853294
Application Date	20.04.2020
Publication Number	20200264154
Publication Date	20.08.2020
Country	US
Title	SYSTEM FOR MONITORING CROPS AND SOIL CONDITIONS
Abstract	<p id="p-0001" num="0000>According to an aspect, a system for monitoring crops and soil conditions below a crop canopy includes a retractable boom assembly adapted to be coupled to an unmanned aerial vehicle. Further according to this aspect, the boom assembly includes an actuator and an elongate probe is coupled to the retractable boom assembly. Still further, the system includes a controller for maneuvering the elongate probe below the crop canopy while the boom assembly is extended by the actuator.</p>
I P C	G01N 33/24; G01N 33/02; B64D 47/00; B64C 39/02; A01B 79/00; G01N 33/00
Applicants	Aker Technologies, Inc.
Inventors	Orlando Saez; Tim Golly; Todd Golly
Priorities Data	NA
National Phase Entries	NA

### 3. Análisis - Primera Iteración

El análisis de las bases se compone de un ciclo de tres etapas. La primera es la estructuración de la data, la ejecución de AP Clustering y definición de los clusters obtenidos.

En esta primera iteración se utilizan las configuraciones que se muestran en la siguiente tabla.

Cuadro 3: M étodología utilizada

Aspecto	Configuraciones
Columna utilizada	IPC nivel A01B

Cuadro 3: M etodolog a utilizada (*continued*)

Aspecto	Configuraciones
Limpieza de texto	No aplica
Utilizaci�n de TF-IDF para ponderar TDMatrix	No
Distancia	Euclidianas
Escalamiento multidimensional	Si - 2D

Al implementar AP Clustering se obtienen los siguientes resultados:

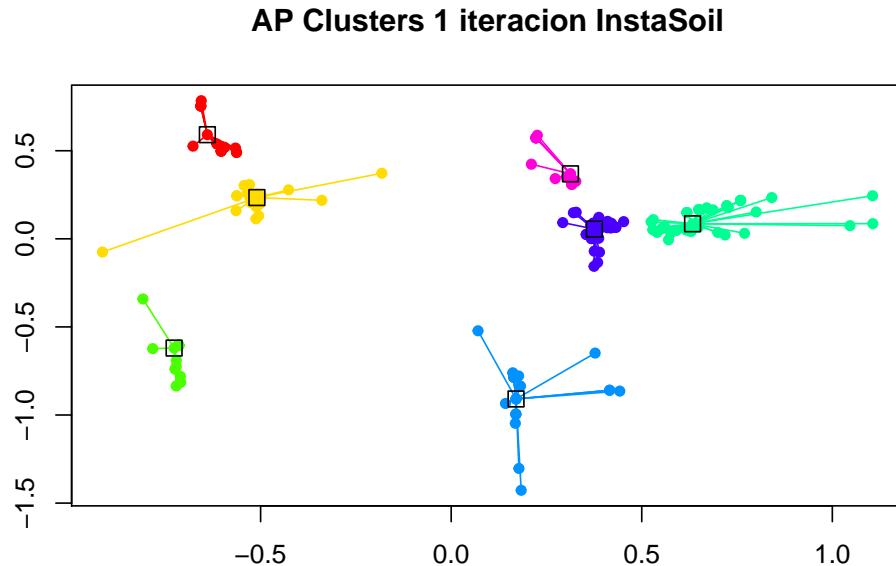


Figura 1: Resultados con AP Clustering

Cuadro 4: M etricas cluster base InstaSoil

Tem�tica	# de Clusters	N de patentes	Net Similarity	Input Preference	# de Iteracciones
Monitoreo de Suelo	7	214	-9.42	-0.829	139

Y al implementar la metodolog a para la definici n de cada uno de los cluster se obtiene lo siguiente

Cuadro 5: Definición de las clasificación de cada cluster

Cluster	Tamaño	Clasificaciones
1	15	g01n, a01c, b60k, b60w, a01b
2	55	g01n, g01d, g08c, b81b, b82b, g01h, g01l, h01l, h04r
3	22	a01g, g01n, e02b, f21s, g01t, h02j
4	42	c12n, c12p, c05f, a01n, a01p
5	23	a01g, c05g, b05b, g05b, h04w
6	41	b09c, b65d, b67d, g01m, g01w
7	16	a01c, g01c, g01s, b01j, b25j, h01j

Al revisar el resultado anterior, se realiza una nueva descarga, esta vez usando la clasificación internacional para generar el dataset. Además se añade un paso en el que se eliminan patentes con clasificaciones no relacionadas a la linea de producto a estudiar, las cuales son descritas en la siguiente tabla.

Cuadro 6: Categorías eliminadas de la base

IPC no relacionadas	Column2
A01B79	Methods for working soil (essentially requiring the use of particular machines, see the relevant groups for the machines)
A01B79	Methods for working soil (essentially requiring the use of particular machines, see the relevant groups for the machines)
A01C	PLANTING; SOWING; FERTILISING (combined with general working of soil A01B0049040000;parts, details or accessories of agricultural machines or implements, in general A01B0051000000-A01B0075000000)
A01D	HARVESTING; MOWING
A01F	THRESHING (combines A01D0041000000); BALING OF STRAW, HAY OR THE LIKE; STATIONARY APPARATUS OR HAND TOOLS FOR FORMING OR BINDING STRAW, HAY OR THE LIKE INTO BUNDLES; CUTTING OF STRAW, HAY OR THE LIKE; STORING AGRICULTURAL OR HORTICULTURAL PRODUCE (arrangements for making or setting stacks in connection with harvesting A01D0085000000)
A01G7	Botany in general
A01G9	Cultivation in receptacles, forcing-frames or greenhouses (of mushrooms A01G0018000000; soilless cultivation A01G0031000000); Edging for beds, lawn or the like
A01K	ANIMAL HUSBANDRY; AVICULTURE; APICULTURE; PISCICULTURE; FISHING; REARING OR BREEDING ANIMALS, NOT OTHERWISE PROVIDED FOR; NEW BREEDS OF ANIMALS

Cuadro 6: Categorías eliminadas de la base (*continued*)

IPC no relacionadas	Column2
A01M	CATCHING, TRAPPING OR SCARING OF ANIMALS (appliances for catching swarms or drone-catching A01K005700000; fishing A01K006900000-A01K009700000; biocides, pest repellants or attractants A01N); APPARATUS FOR THE DESTRUCTION OF NOXIOUS ANIMALS OR NOXIOUS PLANTS
B09C	RECLAMATION OF CONTAMINATED SOIL (gatherers for removing stones or the like from the soil A01B004300000; sterilising soil by steam A01G001100000; removing undesirable matter, e.g. rubbish, from the land E01H001500000)
B62D	MOTOR VEHICLES; TRAILERS (steering, or guiding on a desired track, of agricultural machines or implements A01B006900000; wheels, castors, axles, increasing wheel adhesion B60B; vehicle tyres, tyre inflation or tyre changing B60C; connections between vehicles of a train or the like B60D; vehicles for use on rail and road, amphibious or convertible vehicles B60F; suspension arrangements B60G; heating, cooling, ventilating or other air treating devices B60H; windows, windscreens, non-fixed roofs, doors or similar devices, protective coverings for vehicles not in use B60J; propulsion plant arrangements, auxiliary drives, transmissions, controls, instrumentation or dashboards B60K; electric equipment or propulsion of electrically-propelled vehicles B60L; power supply for electrically-propelled vehicles B60M; passenger accommodation not otherwise provided for B60N; adaptations for load transportation or to carry special loads or objects B60P; arrangement of signalling or lighting devices, the mounting or supporting thereof or circuits therefor, for vehicles in general B60Q; vehicles, vehicle fittings or vehicle parts, not otherwise provided for B60R; servicing, cleaning, repairing, supporting, lifting, or manoeuvring, not otherwise provided for B60S; brake arrangements, brake control systems or parts thereof B60T; air-cushion vehicles B60V; motorcycles, accessories therefor B62J, B62K; testing of vehicles G01M)
B64C	AEROPLANES; HELICOPTERS (air-cushion vehicles B60V)
B64D	EQUIPMENT FOR FITTING IN OR TO AIRCRAFT; FLYING SUITS; PARACHUTES; ARRANGEMENTS OR MOUNTING OF POWER PLANTS OR PROPULSION TRANSMISSIONS IN AIRCRAFT
B65D	CONTAINERS FOR STORAGE OR TRANSPORT OF ARTICLES OR MATERIALS, e.g. BAGS, BARRELS, BOTTLES, BOXES, CANS, CARTONS, CRATES, DRUMS, JARS, TANKS, HOPPERS, FORWARDING CONTAINERS; ACCESSORIES, CLOSURES, OR FITTINGS THEREFOR; PACKAGING ELEMENTS; PACKAGES

Cuadro 6: Categorías eliminadas de la base (*continued*)

IPC no relacionadas	Column2
B67D	DISPENSING, DELIVERING, OR TRANSFERRING LIQUIDS, NOT OTHERWISE PROVIDED FOR (cleaning pipes or tubes or systems of pipes or tubes B08B0009020000;emptying or filling of bottles, jars, cans, casks, barrels, or similar containers, not otherwise provided for B67C;water supply E03;pipe systems F17D;domestic hot-water supply systems F24D;measuring volume, volume flow, mass flow or liquid level, metering by volume G01F;coin-free or like apparatus G07F)
C05F	ORGANIC FERTILISERS NOT COVERED BY SUBCLASSES C05B, C05C, e.g. FERTILISERS FROM WASTE OR REFUSE
C06Q20	NA
C12N	MICROORGANISMS OR ENZYMES; COMPOSITIONS THEREOF; PROPAGATING, PRESERVING, OR MAINTAINING MICROORGANISMS; MUTATION OR GENETIC ENGINEERING; CULTURE MEDIA (microbiological testing media C12Q0001000000)
C12P	FERMENTATION OR ENZYME-USING PROCESSES TO SYNTHESISE A DESIRED CHEMICAL COMPOUND OR COMPOSITION OR TO SEPARATE OPTICAL ISOMERS FROM A RACEMIC MIXTURE
C12Q	MEASURING OR TESTING PROCESSES INVOLVING ENZYMES, NUCLEIC ACIDS OR MICROORGANISMS (immunoassay G01N0033530000); COMPOSITIONS OR TEST PAPERS THEREFOR; PROCESSES OF PREPARING SUCH COMPOSITIONS; CONDITION-RESPONSIVE CONTROL IN MICROBIOLOGICAL OR ENZYMOLOGICAL PROCESSES
G01C	MEASURING DISTANCES, LEVELS OR BEARINGS; SURVEYING; NAVIGATION; GYROSCOPIC INSTRUMENTS; PHOTOGRAMMETRY OR VIDEOGRAMMETRY (measuring liquid level G01F;A radio navigation, determining distance or velocity by use of propagation effects, e.g. Doppler effect, propagation time, of radio waves, analogous arrangements using other waves G01S)
G01G	WEIGHING
G01N1	Sampling; Preparing specimens for investigation (handling materials for automatic analysis G01N0035000000)
G01N5	Analysing materials by weighing, e.g. weighing small particles separated from a gas or liquid (G01N0009000000 takes precedence)
G01S	RADIO DIRECTION-FINDING; RADIO NAVIGATION; DETERMINING DISTANCE OR VELOCITY BY USE OF RADIO WAVES; LOCATING OR PRESENCE-DETECTING BY USE OF THE REFLECTION OR RERADIATION OF RADIO WAVES; ANALOGOUS ARRANGEMENTS USING OTHER WAVES

Cuadro 6: Categorías eliminadas de la base (*continued*)

IPC no relacionadas	Column2
G01V	GEOPHYSICS; GRAVITATIONAL MEASUREMENTS; DETECTING MASSES OR OBJECTS; TAGS (means for indicating the location of accidentally buried, e.g. snow-buried, persons A63B0029020000)
G05B19	Programme-control systems (specific applications, see the relevant places, e.g. A47L0015460000;clocks with attached or built-in means operating any device at a preselected time interval G04C0023000000;marking or sensing record carriers with digital information G06K;information storage G11;time or time-programme switches which automatically terminate their operation after the programme is completed H01H0043000000)
G06N	COMPUTER SYSTEMS BASED ON SPECIFIC COMPUTATIONAL MODELS
G06Q30	Commerce, e.g. shopping or e-commerce
G06Q30	Commerce, e.g. shopping or e-commerce
H04N	PICTORIAL COMMUNICATION, e.g. TELEVISION

Los datos del nuevo dataset son:

Cuadro 7: Descripción de la base utilizada

Concepto	Monitoreo de Suelo
Línea de negocio	InstaSoil
Query	IC:(G) FP:(soil AND crop)
N de patentes pre-procesamiento	1204
N de patentes post-procesamiento	811

En esta segunda iteración se utilizan las configuraciones que se muestran en la siguiente tabla.

Cuadro 8: Métodología utilizada

Aspecto	Configuraciones
Columna utilizada	Title
Limpieza de texto	Si aplica
Utilización de TF-IDF para ponderar TDMMatrix	No
Distancia	Euclidian
Escalamiento multidimensional	Si - 2D

Al implementar AP Clustering se obtienen los siguientes resultados:

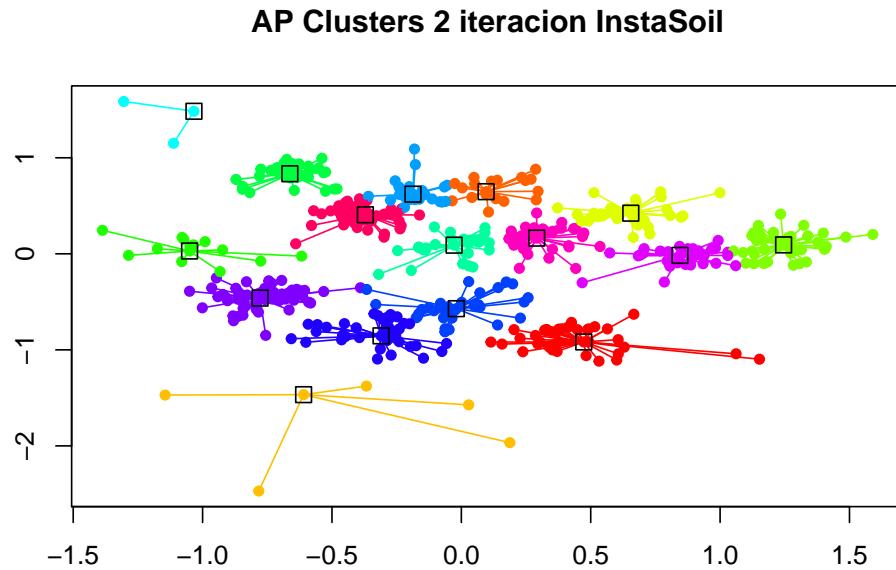


Figura 2: Resultados con AP Clustering

Cuadro 9: Métricas del cluster

Producto	# de Clusters	N de patentes	Net Similarity	Input Preference	# de Iteraciones
Instasoil	16	538	-35.8	-1.13	140

Y al implementar la metodología para la definición de cada uno de los cluster se obtiene lo siguiente

Cuadro 10: Palabras más relevantes en la columna para cada cluster la base de InstaSoil

Cluster	N	Keywords	N Keywords Relacionadas
1	38	SYSTEM, CROP, MONITORING, ENVIRONMENT, GROWTH, BASED, FIELD, INFORMATION, ROOT	1
2	20	METHOD, SOIL, AGRICULTURAL, BASED, SYSTEM, DATA, AGRICULTURE, DEVICE, MANAGEMENT	1

Cuadro 10: Palabras más relevantes en la columna para cada cluster  
la base de InstaSoil (*continued*)

Cluster	N	Keywords	N Keywords Relacionadas
3	6	CROP, SYSTEM, IMAGE, METHOD, PROCESSING, PROGRAM, ACQUISITION, ESTIMATION, INFORMATION, MAPPING, REGIONAL	0
4	38	SYSTEM, AGRICULTURAL, METHOD, SOIL, BASED, MANAGEMENT, MONITORING	2
5	41	SYSTEM, AGRICULTURAL, BASED, INTELLIGENT, MONITORING, INTERNET, DATA	1
6	14	METHOD, SOIL, CROP, MOISTURE, WATER, BASED, EVALUATING, MONITORING	4
7	40	SOIL, METHOD, DEVICE, MOISTURE, CONTENT, DETECTION, FIELD	2
8	40	DEVICE, MONITORING, WATER, APPARATUS, CONTROL, CROP, CYBER, FARM, FARMING, FARMLAND, FERTILIZER, IRRIGATION, METHOD, NITROGEN	3
9	3	METHOD, SOIL, CONDITIONER, CONTAMINATED, CONTENT, COPPER, MOISTURE, PROPERTY	2
10	28	SOIL, METHOD, WATER, DEVICE, MOISTURE, SYSTEM, CONTROL, FERTILIZER, MONITORING	4
11	34	CROP, SYSTEM, METHOD, SOIL, WATER, BASED, GROWTH, INTELLIGENT	2
12	40	CROP, DEVICE, MONITORING, GROWTH, INFORMATION, FIELD, ENVIRONMENT	1
13	60	METHOD, CROP, DEVICE, BASED, WATER, CROPS, FIELD	1
14	35	SYSTEM, CONTROL, MONITORING, GREENHOUSE, INTELLIGENT, IRRIGATION, DATA, MANAGEMENT	1
15	37	SYSTEM, METHOD, IRRIGATION, AGRICULTURAL, DEVICE, INFORMATION, MANAGEMENT, MONITORING	1
16	64	METHOD, BASED, DEVICE, SOIL, IRRIGATION, WATER, APPLICATION, MONITORING	3

Cuadro 11: Títulos de los clusters relacionados en la base de InstaSoil

Cluster	IPC	Definición	N
1	G01D21	Measuring or testing not otherwise provided for	13
1	G01N33	Investigating or analysing materials by specific methods not covered by groups G01N0001000000-G01N0031000000	11
1	G01N21	Investigating or analysing materials by the use of optical means, i.e. using sub-millimetre waves, infrared, visible or ultraviolet light (G01N0003000000-G01N0019000000 take precedence)	7
1	G05D27	Simultaneous control of variables covered by two or more of main groups G05D0001000000-G05D0025000000	3
1	G06Q50	Systems or methods specially adapted for specific business sectors, e.g. utilities or tourism (healthcare informatics G16H)	3
1	A01G1	NA	2
1	A01G25	Watering gardens, fields, sports grounds or the like (special apparatus or adaptations for fertilising-liquids A01C0023000000; nozzles or outlets, spraying apparatus B05B)	2
1	G01B11	Measuring arrangements characterised by the use of optical means (instruments of the types covered by group G01B0009000000 per se G01B0009000000)	2
1	G01N30	Investigating or analysing materials by separation into components using adsorption, absorption or similar phenomena or using ion-exchange, e.g. chromatography (G01N0003000000-G01N0029000000A take precedence)	2
1	G05B13	Adaptive control systems, i.e. systems automatically adjusting themselves to have a performance which is optimum according to some preassigned criterion (G05B0019000000 takes precedence; machine learning G06N0020000000)	2
1	G08C17	Arrangements for transmitting signals characterised by the use of a wireless electrical link	2
1	H04W4	Services specially adapted for wireless communication networks; Facilities therefor	2
2	G06Q10	Administration; Management	4
2	G01N27	Investigating or analysing materials by the use of electric, electrochemical, or magnetic means (G01N0003000000-G01N0025000000 take precedence; measurement or testing of electric or magnetic variables or of electric or magnetic properties of materials G01R)	2
2	G06F16	Information retrieval; Database structures therefor; File system structures therefor	2
2	G06F17	Digital computing or data processing equipment or methods, specially adapted for specific functions (information retrieval, database structures or file system structures therefor G06F0016000000)	2
3	G06T1	General purpose image data processing	1
3	G06T11	2D [Two Dimensional] image generation	1
3	G06T7	Image analysis	1
3	H04L29	Arrangements, apparatus, circuits or systems, not covered by a single one of groups H04L0001000000-H04L0027000000	1
3	H04W84	Network topologies	1
8	G06F19	NA	11 / 11 5
8	G05D9	Level control, e.g. controlling quantity of material stored in vessel	3
8	G01D11	Component parts of measuring arrangements not specially adapted for a specific variable (G01D0013000000, G01D0015000000 take precedence)	2
8	G01W1	Meteorology	2
8	G06K9	Methods or arrangements for reading or recognising printed or written characters or for recognising patterns, e.g. fingerprints (methods or arrangements for graph-reading or for converting the pattern of	2