



UNIVERSIDAD DE CHILE

“Santiago hacia una Ciudad Inteligente: distribución espacial de la
producción de datos en el Gran Santiago”

Universidad de Chile
Facultad Arquitectura y Urbanismo
Escuela de Geografía
2014

Profesor Guía: Dr. Enrique Aliste Almuna.
Alumno: Lic. Cristian Hernández Milla.

“Una casa es una máquina para vivir en ella” (Le Corbusier, 1923)

a.- Agradecimientos

Quisiera agradecer enormemente a mis padres. Siempre han sido el principal apoyo en todos mis proyectos. Además, decirles que no es necesario que lean toda la Tesis, porque quizá no la comprendan o no sea de vuestro interés, pero lean al menos éste párrafo donde digo que son lo más importante para mí.

A mi esposa, por sus correcciones y apoyo. Ella me impulsó a realizar esta investigación, y poder ser finalmente un geógrafo. Tú has sido mi descanso y mi preocupación desde que nos conocimos. Debes leer al menos hasta el Marco Teórico.

A mis amigos, que aún están allí. Sin Facebook no sería posible nuestra amistad. Ustedes no leerán siquiera éste párrafo, porque no puedo subir la Tesis a Internet (aunque quizá sí lo haga). Pero de todas formas lo escribo para el registro.

Finalmente a mi Profesor, Enrique Aliste. Desde que fue mi profesor en tercer año, supe que quería hacer la Tesis con él. Siempre ha querido ir más allá, y eso me agrada. Espero gane el Premio Nacional de Geografía pronto, aunque no sé si se entrega todos los años y tampoco quién entrega dicho premio.

Cristian L. Hernández Milla

b.- Resumen

La siguiente investigación busca explorar los patrones espaciales de la producción de datos en el Gran Santiago, en tanto “los datos”, como la piedra fundacional de lo que conocemos como una “Ciudad Inteligente”. Así también, se analiza la relación de dichos patrones espaciales generados por los datos, tanto activos como pasivos, con la distribución espacial del ingreso por hogar en el área estudiada, con el fin de establecer, si existe una relación entre ellos en el contexto de Santiago como una ciudad segregada. Finalmente, se entregan conclusiones en relación a la potencialidad de ciertas comunas del Gran Santiago, para transformarse en territorios que fomenten una “Sociedad Habilitada Espacialmente”, en términos del aprovechamiento de los datos para una mejor planificación y diseño de la Ciudad.

Palabras Clave: Producción de Datos, Ciudad Inteligente, Patrón Espacial, Sociedad Habilitada Espacialmente.

Contenido

Índice	
a.- Agradecimientos	3
b.- Resumen	4
Contenido	4
Capítulo I: Presentación	7
I.a- Introducción	7
I.b.- Planteamiento del Problema.....	8
I.b.1.- El Surgimiento de la Capa Sensible en las Ciudades	8
I.c.- Hipótesis de Trabajo	11
I.d.- Objetivos.....	11
I.d.1.- Objetivo General	11
I.d.2.- Objetivos Específicos	12
I.e.- Metodología.....	12
I.e.1.- Cartografiar la producción Activa de Datos para el Gran Santiago al año 2014, mapeando datos geolocalizados producidos de forma activa desde Twitter, identificando su patrón espacial.....	12
I.e.2.- Cartografiar la producción Pasiva de Datos para el Gran Santiago al año 2014, mapeando datos geolocalizados producidos en las validaciones BIP Transantiago, identificando su patrón espacial	15
I.e.3.- Comparar los patrones espaciales de la producción de datos activos, pasivos y totales, con el patrón espacial del ingreso en el Gran Santiago, con el fin de determinar si la producción de datos tiene correlación con el nivel de ingreso en las 34 comunas que componen el Gran Santiago.	16
Capítulo II: Marco Teórico	18
II.a.- Definición de la Ciudad Inteligente y la Búsqueda por la Equidad	18
II.b.- Proyectos de Ciudad Inteligente en Santiago: capas de inteligencia en una Ciudad Segregada Socioespacialmente.....	20
Capítulo III: Desarrollo	23
III.a.- La producción de Datos Activos en el Gran Santiago: <i>Tweets</i>	23
III.b.- La producción de Datos Pasivos en el Gran Santiago: Validaciones BIP.	27
III.c.- La producción Total de Datos en el Gran Santiago	32
II.d.- Relación entre la Producción de Datos y el Ingreso en las comunas del Gran Santiago. ..	38
Capítulo IV: Conclusiones	44
IV.a.- Sobre los patrones espaciales de la producción de datos y su relación con el ingreso. ..	44
IV.b.- Potencial en el Gran Santiago de favorecer la Sociedad Habilitada Espacialmente	45
Bibliografía	48
Anexos	50
A.- Mapas utilizados en esta investigación.....	50
B.- Código Javascript.....	50
C.- Tabla Total de Coordenadas Twitter y Validaciones BIP.....	50
Mapas	
Mapa 1: “Dispositivos Conectados a Internet en el Mundo”	9
Mapa 2: “Área de recolección de tweets”.....	15
Mapa 3: “Producción total de <i>Tweets</i> para un día laboral en el Gran Santiago”.....	23

Mapa 4: “Comparativa de <i>Tweets</i> producidos en Plantas Urbanas fuera del Gran Santiago”...	24
Mapa 5: “Concentración Kernel de <i>Tweets</i> para un día Laboral en el Gran Santiago”.....	25
Mapa 6: “Concentración Zonal de <i>Tweets</i> para un día Laboral a nivel de Manzana en el Gran Santiago”.....	26
Mapa 7: “Localización de Paradas de Bus y Estaciones de Metro, Transantiago”.....	28
Mapa 8: “Concentración Kernel Validaciones BIP en un día laboral en el Gran Santiago”....	30
Mapa 9: “Concentración Zonal Validaciones BIP en un día laboral a nivel de Manzana, Gran Santiago”.....	31
Mapa 10: “Concentración Kernel Producción Total de Datos para un día laboral en el Gran Santiago”.....	34
Mapa 11: “Comparativa Concentraciones Kernel de Producción de Datos en un día laboral en el Gran Santiago”.....	35
Mapa 12: ““Zonas sin Concentración Relevante en la Producción de Datos”.....	37
Mapa 13: “Concentración Zonal Total de Datos para un día laboral por Manzana, en el Gran Santiago”.....	37
Mapa 14: “Distribución del Ingreso en el Gran Santiago”.....	38
Mapa 15: “Correlación entre Ingreso Por Hogar y Producción de Datos en un día laboral, a nivel de Manzanas en el Gran Santiago”.....	39
Mapa 16: “Análisis LocalR2: <i>Tweets</i> e Ingreso Monetario Promedio”.....	40
Mapa 17: “Análisis LocalR2: Validaciones BIP e Ingreso Monetario Promedio”.....	41
Mapa 18: “Análisis StdResid: <i>Tweets</i> e Ingreso Monetario Promedio”.....	42
Mapa 19: “Análisis StdResid: Validaciones BIP e Ingreso Monetario Promedio”.....	43

Figuras

Figura 1: “Nacimiento del IoT”.....	9
Figura 2: “Modelo Comunicación con la API de Twitter”.....	13
Figura 3: “Proceso de recolección de datos en Node.js”.....	14
Figura 4: “Topología de Funciones de una Ciudad Inteligente”.....	19

Gráficos

Gráfico 1: ““Producción de <i>Tweets</i> por Comuna en un día laboral en el Gran Santiago”.....	24
Gráfico 2: “Validaciones totales BIP por Comuna en un día laboral, Gran Santiago”.....	29
Gráfico 3: “Producción Total de Datos por Comuna en un día laboral, Gran Santiago”.....	33
Gráfico 4: “Comparativa Datos Totales Producidos por Comuna en un día Laboral, Gran Santiago”.....	33

Tablas

Tabla 1: “Producción Acumulada de <i>Tweets</i> en un día laboral por Comuna en el Gran Santiago”.....	27
Tabla 2: “Validaciones BIP acumuladas por Comuna en un día laboral, Gran Santiago”.....	32
Tabla 3: “Producción Total de Datos acumulados para un día laboral en el Gran Santiago”.....	38

Capítulo I: Presentación

I.a- Introducción

La ciudad de Santiago según rankings internacionales como “*Cities in Motion*” (IESE, 2014) y nacionales como el “Ranking de Ciudades Inteligentes en Chile” (Cohen & Obediente, 2014), se sitúa como la ciudad “más inteligente” entre las ciudades latinoamericanas, y la más inteligente entre las ciudades chilenas, respectivamente. En efecto, parece existir en Chile al año del presente informe, cierto interés de dotar a la ciudad de “inteligencia”, demostrado en proyectos como “Santiago Smart City”¹, “Ciudad Inteligente para el Transporte Chile 2020”² o “Smart City Concepción”³.

Sin embargo, la denominación de “Inteligente” de una Ciudad ha generado un debate entre quienes ponen el acento en la agregación de “tecnologías inteligentes” a la Ciudad, con el fin de “sensorizar” el espacio urbano y recopilar datos (principalmente empresas productoras de tecnología como IBM, Cisco, Siemens, Ericsson, Microsoft), y entre quienes apoyan que la Ciudad es inteligente en tanto la “coordinación de los habitantes urbanos” (coordinación de la inteligencia humana), permitida por dichas TICs (Greenfield, 2013; Townsend, 2013).

La sensorización del espacio urbano a través de diversas tecnologías, y el fenómeno de coordinación de los habitantes urbanos a través de TICs, ha permitido la producción y registro de una gran cantidad de datos, los cuales son utilizados cada vez de forma más intensiva para el diseño y la planificación urbana, como es el caso del uso de datos provenientes desde *Smartphones* para el diseño urbano en US y UK (ARUP, 2013).

Los datos recopilados en entornos urbanos, pueden ser diferenciados en tanto la forma que son recolectados, ya sea de forma pasiva como activa. Los datos pasivos, o de “ambiente”, corresponden a aquellos datos que son recopilados por sensores que no necesitan de la mediación humana, como cámaras CCTV o tarjetas para el pago en el transporte público, como la tarjeta “BIP”, utilizada en el transporte público de Santiago. Por su parte, los datos activos son aquellos producidos exclusivamente en la intervención humana, como compartir contenido a través de Internet (mensajes, fotos, videos, etc.), principalmente a través de *Smartphones* (IDC – EMC, 2008).

La sensorización activa y pasiva, y los datos recolectados a través de ambos procesos, se ha transformado en un insumo vital para “hacer Ciudad” actualmente, en procesos que favorecen o inhiben la participación de los habitantes urbanos, en tanto permiten a las diversas comunidades que forman la ciudad, acceder de un modo abierto a los datos que en la Ciudad misma se producen. Así mismo, un nuevo poder emerge para quienes comandan la Ciudad, que es el de abrir los datos que diversos sensores producen registrando la actividad de sus habitantes,

1 Ver www.smartcitysantiago.cl

2 Ver <http://www.ciudadesinteligentes.gob.cl/>

3 Ver <http://innovatingcities.org/innovatingcities/chile/>

permitiendo el desarrollo de lo que se conoce como “Sociedad Habilitada Espacialmente” (Williamsom, 2011: Steudler & Rajabifard, 2012) o SES (*Spatially Enabled Society*). En este escenario, los datos se transforman en una herramienta de apropiación del espacio, de creación de territorio, y el acceso a los datos geolocalizados un derecho del habitante urbano, en tanto “creación” (Florida, 2003), o “hackeo” (Townsend, 2013), de la Ciudad.

La pregunta que cabe hacer ahora es ¿Todas las áreas de la Ciudad producen la misma cantidad de datos?, ¿Existirían áreas con una potencialidad menor de convertirse en “Ciudad inteligente”, en tanto es la producción de Datos su insumo vital? (Batty, 2013). Dichas preguntas sobre el cómo se distribuye la producción de datos en la Ciudad, cobra mayor pertinencia aún en Santiago, comprendida como una ciudad segregada socioeconómicamente, especialmente en términos de ingreso y tecnología (Hernández, 2006) ¿Será la producción de datos y su distribución en el espacio, una nueva manifestación de desigualdad territorial, en cuanto permite o inhibe el desarrollo de la Ciudad Inteligente?

La siguiente investigación busca realizar el ejercicio exploratorio de determinar los patrones espaciales resultantes de la producción de datos en el Gran Santiago, comparando dichos patrones con la configuración del Gran Santiago en tanto al ingreso por hogar, con el fin de entregar luces y abrir un debate sobre una posible nueva manifestación de segregación socioespacial, ahora en un nuevo tipo de Ciudad, la Ciudad Inteligente.

I.b.- Planteamiento del Problema

I.b.1.- El Surgimiento de la Capa Sensible en las Ciudades

Internet durante la última década ha transitado desde una red de comunicación entre humanos, a una red de comunicación entre “artefactos” denominada *Internet of Things* (IoT), o “Internet de las Cosas” (Castells, 2014), el cual se basa en la idea de una red de datos provenientes desde sensores dispuestos en el mundo real⁴, creada a partir de entregarle capacidad de comunicación a todo artefacto en el planeta (Cisco, 2011). En la IoT a diferencia de Internet y la *World Wide Web* (www), la información producida es generada por objetos, artefactos o máquinas (incluso animales) que entregan información del entorno en el que están dispuestas, como por ejemplo *Smartphones*, autos, bicicletas, un lago⁵, una casa⁶, lámparas⁷, etc⁸.

Por su parte, el aumento de los dispositivos conectados o la IoT, ha tenido un ritmo aún más vertiginoso que el de personas conectadas a la Internet mediante la WWW. Se calcula que hoy en día 3 Billones de personas acceden a Internet (ITU, 2014), es decir un 40% de la Población mundial, lo que se ha considerado un gran logro en tanto Internet como un derecho humano (ONU, 2011). Sin embargo, ya al año 2010 existían 12.5 billones de artefactos conectados al IoT

⁴ Ver <http://www.forbes.com/sites/oreillymedia/2014/03/12/tim-oreilly-talks-about-the-internet-of-things-and-beyond/>

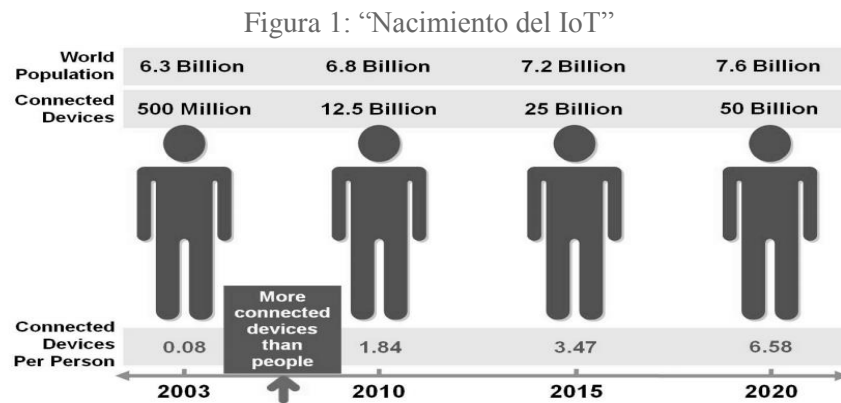
⁵ Lago George, Estado de Nueva York, Estados Unidos.

⁶ Ver <http://www.hondasmarthome.com/>

⁷ Ver <http://lifix.co/>

⁸ Ver <http://www.oei.es/divulgacioncientifica/?La-invasion-de-los-objetos>

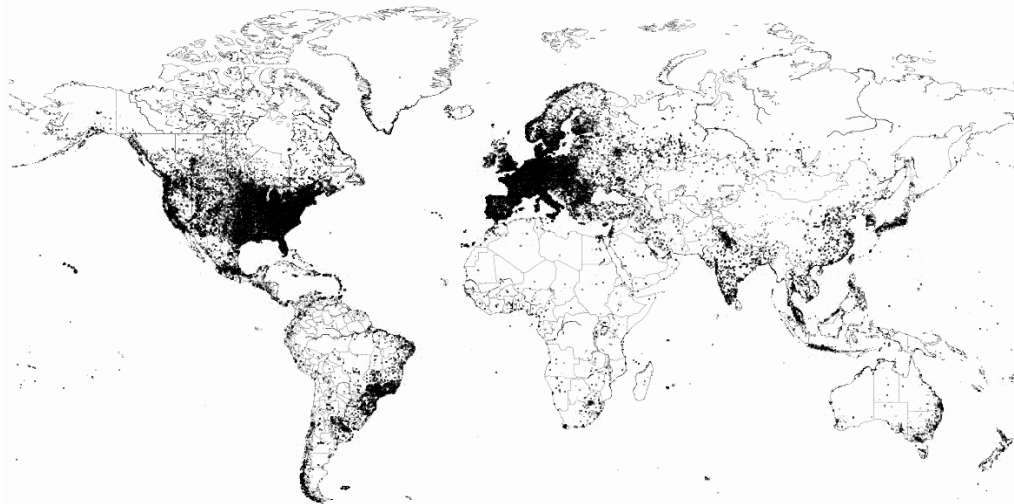
(Cisco, 2011), y más aún, ya entre los años 2008 y 2009⁹ la cantidad de artefactos conectados superaba a la población mundial, momento que se ha denominado como el “Nacimiento del IoT” (Cisco, 2011):



Fuente: Cisco, 2011.

En términos territoriales, el IoT posee una clara vocación urbana, en tanto la localización de los artefactos conectados en red se concentran en las mayores zonas urbanas concentradas y semiconcentradas del planeta, tal como se aprecia en el mapa elaborado por la compañía *Shodan*¹⁰, donde se muestra en color negro las concentraciones de todos los dispositivos conectados a Internet en el mundo:

Mapa 1: “Dispositivos Conectados a Internet en el Mundo”



Fuente: *Shodan*, 2014

Si el objetivo del Internet fue conectar personas, facilitando y creando nuevas formas de comunicación, el objetivo de la IoT es la creación de datos, transformando cada objeto en el planeta en un sensor. Como por ejemplo lo sería un automóvil, que a la vez que medio de

⁹ Se estima ese momento entre 2008 y 2009, el nacimiento de la IoT. Cisco (2011).

¹⁰ Ver <http://www.shodanhq.com/>

transporte, es un objeto sensorizado que transmite datos sobre la ruta, y que ha permitido el nacimiento del “*Driverless Car*” (McEvoy, 2014). Se estima que cada año los datos recopilados a través de la “Capa Sensible” de artefactos conectados al IoT, aumentan a un ritmo de cuatro mil por ciento (%4.300), calculando que para el año 2015 se producirá una cantidad de 7.9 ZB¹¹ de datos en el planeta (CSC, 2011).

La evolución de la IoT a través de su capacidad de sensorización del espacio y creación de datos, ha facilitado el camino para el nacimiento de proyectos que buscan crear a través de la sensorización intensiva, un “sistema nervioso” en el territorio, como es el caso de CeNSE¹² (*Central Nervous System for the Earth* o Sistema Nervioso Central del Planeta), que busca “revolucionar” la forma en la que se recolecta, comunica y analiza información. Tal como menciona *Hewlett-Packard* líder del proyecto, “CeNSE consiste en una red de inteligencia superior de billones de sensores a nano escala, diseñados para sentir, oler, ver y escuchar qué pasa en el mundo” (Hewlett-Packard, 2013).

Otro ejemplo de la utilización de sensores, ahora para la gestión de la Ciudad, es el proyecto CORIO¹³ (Centro de Comando de la Prefectura de Río de Janeiro), que nace de la asociación del gobierno de la Ciudad e IBM, inaugurado en 2012. Según nota de prensa¹⁴, CORIO tiene por objetivo “recopilar datos de diversas agencias gubernamentales, a través de sensores, cámaras de video y/o dispositivos geolocalizados, con el objetivo de analizar dichos datos con algoritmos computacionales, identificando patrones y tendencias, creando lo que se denomina un <software sensible>”.

Otro caso de utilización de los datos para el planeamiento de la Ciudad es Londres, donde los datos de transporte público (TFL¹⁵), permanecen abiertos para el uso de quien desee aprovecharlos, ya sea con fines académicos o incluso comerciales. Los datos recopilados por TFL, han sido útiles para laboratorios de investigación como el *Centre for Advanced Spatial Analysis* (CASA)¹⁶; en lo que ellos denominan la “Nueva Ciencia de la Ciudad” (Batty, 2011), así como también para compañías como “*Citymapper*”, creadores de una de las aplicaciones móviles de más uso en Londres, que entrega respuestas en tiempo real sobre rutas de viaje¹⁷.

En Chile, un ejemplo de datos para la planificación de la Ciudad, son los datos provenientes desde las validaciones o pagos con la tarjeta BIP, la cual utiliza un RFID (*Radio Frequency Identification*)¹⁸, para los registros de carga y pagos en el Transantiago. Dichos datos sirven para la mejora continua del sistema, como el cálculo de las frecuencias y cálculo del pago. Sin embargo dichos datos permanecen cerrados, sin acceso público a ellos.

11 ZB = 1.000.000.000 TeraBytes (TB) - 1 TB = 1.000 GigaBytes (GB).

12 Ver <http://www8.hp.com/us/en/hp-information/environment/cense.html#.VD70bmddWHQ>

13 Ver <http://www.rio.rj.gov.br/web/corio>

14 *New York Times*: “Mission Control, Built for Cities”, 2012.

15 Ver <https://www.tfl.gov.uk/info-for/open-data-users/>

16 Ver <http://www.bartlett.ucl.ac.uk/casa>

17 Ver <https://citymapper.com>

18 Tecnología inalámbrica que permite la comunicación entre un lector y una etiqueta.

Los datos generados a través de los artefactos conectados en red) en nuestras ciudades, han abierto la discusión sobre el acceso a dichos datos por parte de los habitantes urbanos, en tanto la gran cantidad de datos producidos en entornos urbanos, permitiría el acceso a la ciudad (Townsend, 2013), al dotar a las distintas comunidades que habitan el territorio urbano de la capacidad de intervenir el espacio, creando ciudad, en un proceso de apropiación o creación de territorio urbano. A dicho proceso de apropiación del espacio, a través de los datos, lo denominamos "Sociedad Habilitada Espacialmente" (Rajabifard, 2009; Roche 2013).

Así, la Ciudad se transforma en un lugar en el cual se deben generar procesos de coordinación de los habitantes urbanos para el acceso a los datos que la capa sensible de la Ciudad genera, en tanto dicho acceso a los datos, provee acceso a la Ciudad (Williams et al, 2014;Kung, 2014).

En este escenario, la siguiente investigación busca realizar un ejercicio exploratorio usando como fuente dos tipos de datos, a saber, los datos provenientes desde la creación de datos geolocalizados en la red social *Twitter*, y los datos provistos por la Subsecretaría de Transporte para las Validaciones BIP, con el fin de explorar la distribución espacial de la producción de datos en Santiago, y cómo se relaciona dicho patrón espacial con el ingreso a nivel comunal para el Gran Santiago.

Si la respuesta es positiva, podríamos estar frente a un indicio de una nueva expresión de segregación espacial, la cual debería ser corregida en actuales y futuros planes de Ciudad Inteligente para el Gran Santiago. De ser la respuesta negativa, podríamos encontrarnos frente a una posibilidad, en la cual comunas con menores ingresos, puedan generar planes de Ciudad Inteligente, habilitando espacialmente a sus vecinos, entregándoles una nueva herramienta de uso y acceso a la Ciudad.

I.c.- Hipótesis de Trabajo

La hipótesis de trabajo plantea que la producción de datos en el Gran Santiago, medida en términos de datos activos y pasivos, presenta un patrón con tendencia a tener directa relación con el patrón espacial del ingreso, en tanto los ingresos y su distribución en el espacio urbano del Gran Santiago, se ha transformado en uno de las principales causas de explicación de la configuración espacial de diversos fenómenos en la Ciudad.

I.d.- Objetivos

I.d.1.- Objetivo General

El objetivo general de la presente investigación, es determinar los patrones espaciales generados por la producción de datos activos y pasivos en el Gran Santiago, al año 2014. Además, comparar dichos patrones espaciales con la distribución del ingreso a nivel de hogar para las 34 comunas que componen el Gran Santiago, discutiendo sobre la relación de la producción de datos y los efectos de la distribución de los ingresos en dicho patrón espacial.

I.d.2.- Objetivos Específicos

A.- Cartografiar la producción Activa de Datos para el Gran Santiago al año 2014, mapeando datos geolocalizados producidos de forma activa desde Twitter, identificando su patrón espacial.

B.- Cartografiar la producción Pasiva de Datos para el Gran Santiago al año 2014, mapeando datos geolocalizados producidos en las validaciones BIP Transantiago, identificando su patrón espacial.

C.- Comparar los patrones espaciales de la producción de datos activos, pasivos y totales, con el patrón espacial del ingreso en el Gran Santiago, con el fin de determinar si la producción de datos tiene correlación con el nivel de ingreso en el Gran Santiago.

I.e.- Metodología

I.e.1.- Cartografiar la producción Activa de Datos para el Gran Santiago al año 2014, mapeando datos geolocalizados producidos de forma activa desde Twitter, identificando su patrón espacial.

Para cartografiar los datos activos, se utiliza *Twitter*, y la producción de mensajes geolocalizados a través de este servicio, utilizando la API provista por dicha compañía para desarrolladores. Para la recolección de datos, se utiliza el ambiente de programación *Node.js*, y el código *Javascript* provisto por Steven Gray¹⁹ (*University College of London*). La definición y explicación del proceso se describe a continuación:

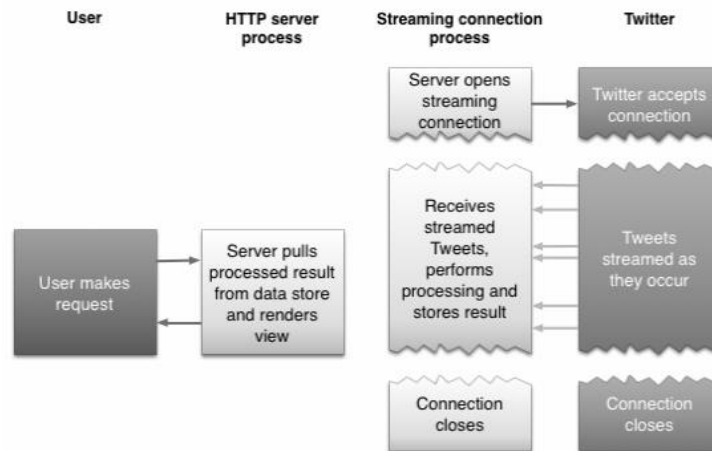
API: siglas para definir la *Application Programming Interface*, o en español conocido como “Interfaz de Programación”. A través de una API accedemos a las bases de datos, provistos por un servicio en Internet (web o móvil). En este caso se utilizará la API de *Twitter*, en específico la “*Streaming API – Public Stream*”, que entrega una muestra del total de *Tweets* generados, sin restricción en la cantidad de contenido que entrega, es decir, puede leer desde un *tweet* a millones de *tweets*, entregando la muestra correspondiente para cada número respectivamente²⁰.

A continuación se presenta un modelo que explica el funcionamiento de la *Streaming API*:

¹⁹ Ver <https://stevenjamesgray.com>

²⁰ Ver <https://dev.twitter.com/streaming/overview>

Figura 2: “Modelo Comunicación con la API de Twitter”

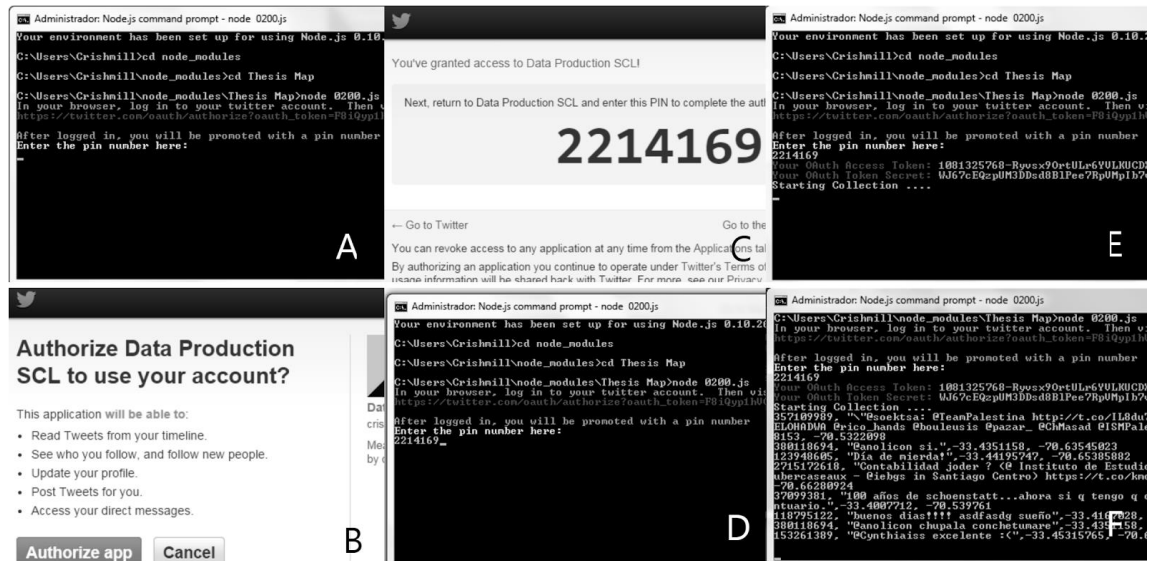


Fuente: Twitter, 2014

El funcionamiento básico radica en realizar “requerimientos” o consultas a través de un diálogo HTTP. Luego del requerimiento al servidor de Twitter, se entrega un acceso al *streaming* requerido, en este caso todos los *tweets* producidos en un polígono que cubre el Gran Santiago. Aceptado el *streaming* de datos por parte de Twitter, se almacenan cuatro variables, a saber: “id de usuario”; “contenido del *tweet*”; y “coordenadas de longitud y latitud” en donde se produjo dicho mensaje o *tweet*. Éstas variables se almacenan automáticamente en formato de texto CSV, en orden cronológico.

Node.js: node.js es un “ambiente de programación” que utiliza el lenguaje *Javascript* para crear aplicaciones que se comunican a nivel de servidor con diversos servicios. Las siguientes imágenes muestran la forma en la que el proceso de requerimientos de tweets geolocalizados ocurrió para esta investigación:

Figura 3: “Proceso de recolección de datos en Node.js”



Fuente: Elaboración propia, 2014.

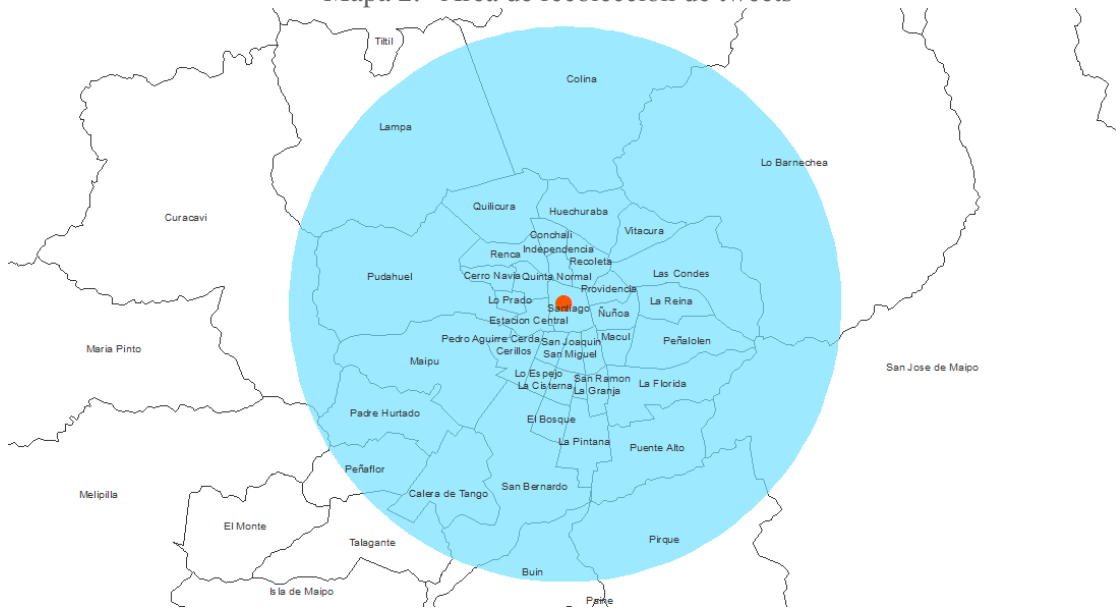
Son seis momentos desde que se ejecuta el código *Javascript* que contiene los requerimientos HTTP a la *Streaming API* de Twitter. Al momento de ejecutar el comando, como se aprecia en el momento “A”, la API Twitter, requiere de la autorización a través de un código PIN, que cual se obtiene en la Web de Twitter, como se aprecia en el momento “B” y “C”. Posteriormente, luego de ingresar el código PIN, la aplicación aprueba los privilegios de desarrollador, con el fin de comenzar la recolección de datos, como se aprecia en el momento “E”. Finalmente, los tweets comienzan a ser recolectados con la información que hemos requerido, en formato CSV.

Javascript: el comando escrito en el lenguaje de programación *Javascript*, que fue utilizado en ésta investigación, es el creado por Steven Gray, el cual se encuentra dispuesto al uso público en el repositorio de código abierto *Github*²¹. El código puede descargarse en los anexos de éste informe.

Área de Recolección de Datos: el área de recolección de tweets, corresponde a un punto específico del mapa, al cual se le asigna un área de cobertura. Twitter en este caso, genera un polígono con centro en las coordenadas entregadas, y con la distancia asignada. Para esta investigación el centro está representado por la coordenada -33.45 , -70.66 . El área tiene una extensión de 30 kilómetros, tal como se aprecia en el siguiente mapa:

21 Ver <https://github.com>

Mapa 2: “Área de recolección de tweets”



Fuente: Elaboración propia, 2014.

Tiempo de recolección de datos: la recolección de tweets se realizó para un día completo, siendo escogido un día laboral al azar, posteriormente a la realización de tres pruebas técnicas, con el fin de afinar el instrumento de recolección. La razón de escoger un día laboral, es que los datos pasivos recolectados, en este caso las Validaciones BIP, poseen dicho recorte temporal (Día Laboral Promedio). Se recolectaron 75.394 tweets, al día 1 de Septiembre del año 2014.

Análisis de los Patrones Espaciales: se procede a convertir el archivo CSV generado en la recolección de *Tweets*, a formato SHAPE, utilizando el software abierto *QGIS*²². Una vez obtenido el archivo SHAPE, y ser transformado a coordenadas WGS84 (19S), se procede al análisis de la información en ArcMap 10.1; generando un mapa de “Densidad de *Kernel*”. Posteriormente se realiza un análisis “Estadístico Zonal” en ArcMap 10.1, con el fin de cruzar la información obtenida de puntos de producción de información en Twitter, con la división a nivel de Manzanas del Gran Santiago, comparando posteriormente el patrón espacial resultante con la distribución del ingreso a nivel de Manzanas para el área estudiada.

I.e.2.- Cartografiar la producción Pasiva de Datos para el Gran Santiago al año 2014, mapeando datos geolocalizados producidos en las validaciones BIP Transantiago, identificando su patrón espacial.

Se utilizan los datos abiertos provistos por la Subsecretaría de Transportes, a través del portal de datos públicos “Datos.gob.cl”²³. En específico, se utilizan las bases de datos correspondientes al

22 Ver <http://www.qgis.org/>

23 Ver <http://datos.gob.cl/>

“Feed GTFS del Transantiago”²⁴, y “Matrices Viajes, Subidas y Bajadas del Transantiago”²⁵. Para ambos casos, se utiliza la información de Buses y Metro.

Desde la combinación de ambas bases de datos, se obtienen los datos de “Validaciones BIP” para Buses y Metros, en todas las paradas de Buses del Gran Santiago, y las estaciones de Metro. Por “Validaciones BIP”, se entiende al momento en que el usuario de transporte acerca su tarjeta al validador, con el fin de pagar por su viaje, tanto al inicio, como al momento de realizar trasbordos. No se consideró la información sobre las Bajadas, ya que no se realiza validación con Tarjeta BIP en dicho momento. El dato (bajadas) corresponde a una inferencia obtenida de los datos de subidas y las coordenadas GPS década Bus, más que a una medición real de datos producidos (Fisher, 2010; Miranda, 2011).

La información dispuesta en formato GTFS corresponde a una colección de archivos en formato TXT (texto), que contiene información geolocalizada de paradas (buses y metro), validaciones, rutas y zonas. Para el caso de ésta investigación, se utiliza el archivo “Stops”, que contiene la información tanto de buses (10.551 paraderos), como también estaciones de metro (100 estaciones²⁶). Dicha información se combina con la información de las validaciones con tarjeta BIP que se encuentra en la matriz de viajes, subidas y bajadas. Dicha información está dispuesta para “un día laboral promedio” del año 2012. Lamentablemente no se encuentra disponible al público información más actualizada²⁷.

La tabla resultante corresponde a 10.651 puntos georeferenciados que corresponde a paradas de buses y estaciones de metro, cada cual con la cantidad total de Validaciones para un día promedio del año 2012. Dichos datos son procesados a través de QGIS, con el fin de ser transformados a un archivo SHAPE de puntos, tal como se realizó para el caso de los datos de Twitter.

Se realiza luego un análisis de concentración Kernel utilizando ArcMap 10.1. Finalmente, se realiza un análisis “Estadístico Zonal”, en ArcMap 10.1, con el fin de cruzar la información obtenida de puntos de producción de información en Twitter, con la división a nivel de Manzanas del Gran Santiago.

I.e.3.- Comparar los patrones espaciales de la producción de datos activos, pasivos y totales, con el patrón espacial del ingreso en el Gran Santiago, con el fin de determinar si la producción de datos tiene correlación con el nivel de ingreso en las 34 comunas que componen el Gran Santiago.

24 Ver <http://datos.gob.cl/datasets/ver/1587>

25 Ver <http://datos.gob.cl/datasets/ver/1655>

26 Según Metro existen 108 estaciones de metro, sin embargo, se cuentan ocho estaciones en distintas líneas, lo que hace aumentar el número, desde 100 estaciones reales, a 108 estaciones en distintas líneas.

27 Si bien se pudo hacer gestiones para obtener información más actualizada, resulta relevante indicar que éste ejercicio académico, implica reconocer los niveles de “Inteligencia Espacial” que presenta Santiago, es decir, qué datos y cómo están dispuestos al público en general para su uso, sin buscar consideraciones especiales para el acceso a los datos, al ser un derecho de la Sociedad Habilitada Espacialmente.

Para la comparación de los patrones espaciales producidos desde la producción de datos, tanto activos, pasivos y totales, se realiza un análisis de correlación estadística tomando los conjuntos de datos de ingreso a nivel de Manzanas, y la concentración de producción de datos para cada caso (activa, pasiva, total). En específico, se compara el nivel de ingresos predominante para cada Manzana, en base a la clasificación ABC1, C2, C3, D, E, entregada por Adimark, en base al Censo de Población y Vivienda 2002. Para el análisis de los datos, se utilizó el software Excel.

La fórmula utilizada es la siguiente:

$$Correl(X, Y) = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Donde x e y son las medias de muestra PROMEDIO (matriz1) y PROMEDIO (matriz2).

Además, se realiza un análisis GWR (Charlton, Fotheringham, Brunson, 2006), considerando las variables de Ingreso Monetario por Hogar, a nivel comunal para el año 2009, entregado por CASEN, y la producción de datos por comuna, tanto activa como pasiva, con el fin de obtener la correlación final entre las variables.

El análisis GWR (*Geographically Weighted Regression* - Regresión Geográficamente Ponderada), permite generar un análisis de regresión para cada unidad espacial (en este caso comunas. En específico se realizan dos tipos de análisis:

R2 (LocalR2): R cuadrado es una medida de la bondad de ajuste. Su valor varía de 0,0 a 1,0; se prefieren los valores más altos. Se puede interpretar como la proporción de varianza de la variable dependiente que da cuenta el modelo de regresión. El denominador para el cálculo de R2 es la suma de los valores de la variable dependiente cuadrada. Agregar una variable explicativa extra al modelo no altera el denominador pero sí altera el numerador; esto brinda la impresión de mejora en el ajuste del modelo que puede no ser real. Consulte el R2 ajustado a continuación.

Valores Residuales (StdResid): The residuals represent the amount of variability that remains unexplained by the model (i.e. variability on the dependent variable not explained by the independent variables). The standard deviations of the residuals also show the locations where the model over and unpredicted the value of the dependent variable

Capítulo II: Marco Teórico

II.a.- Definición de la Ciudad Inteligente y la Búsqueda por la Equidad

La *Smart City* o Ciudad Inteligente, se presenta como un término acuñado por grandes compañías productoras de tecnología (Greenfield, 2013; Townsend, 2013), tanto para la producción de proyectos de agregación de capas de inteligencia en ciudades ya existentes, como también para planes de creación de ciudades inteligentes desde cero, tal como es el caso de “Smart City India”²⁸. Así lo menciona Pallot (et al, 2011): “Una encuesta reciente de Forrester, establece que las soluciones de Ciudad Inteligente, son iniciativas empujadas más por vendedores (de tecnología), que propuestas de gobiernos”. Así, la Ciudad Inteligente se presenta como una obra tecnológico-arquitectónica, como un “emprendimiento”, que crea nuevas ciudades²⁹, o genera nuevos espacios en la ciudad, potenciando lo que nuevos artefactos tecnológicos pueden hacer por el habitante urbano, en algo al parecer “sin peso”, que llamamos territorio. Una vuelta, a lo propuesto por *Le Corbusier* (1923), donde las “máquinas”, como el auto o la casa, están en el centro de la planificación del territorio, y no el habitante urbano.

En este escenario, las definiciones de la Ciudad Inteligente provienen en gran medida desde las empresas y sus departamentos de marketing, poniendo énfasis en las bondades de sus tecnologías para resolver las problemáticas urbanas, como es el caso de la visión de IBM de las Ciudades Inteligentes³⁰: “Un mundo urbanizado significa que las ciudades están ganando un gran control sobre su desarrollo, económico y político. Las Ciudades están ganando además empoderamiento tecnológico, así los sistemas centrales en los cuales se basan, se están interconectando, permitiéndoles agregar nuevos niveles de inteligencia (...) Para aprovechar oportunidades y construir prosperidad sustentable, las ciudades necesitan transformarse en inteligentes.”.

La Ciudad Inteligente por otro lado, sería aquella que es capaz de resolver al mismo tiempo problemas urbanos en diversos sistemas al interior de la Ciudad (Giffinger, 2007). Cambiando la noción de “tiempo de respuesta” que una Ciudad puede hoy en día entregar para solucionar las problemáticas que en el día a día todo territorio urbano debe enfrentar, como el manejo del agua, movimiento de personas, o la distribución de las comunicaciones y energía. En este nuevo “Sistema Operativo de la Ciudad”, tal como lo plantea el objetivo de la compañía *LivingPlanIT*³¹, al igual que en un computador se pueden procesar diversas respuestas al mismo tiempo, en cerebros urbanos que poseen la forma de “comandos centrales” (Greenfield, 2013).

Así la Ciudad Inteligente, es catalogada en cuanto a las funciones que debiera cumplir para ser considerada inteligente. Al respecto Batty (et al, 2012), nos menciona que las Ciudades “solo pueden ser inteligentes, si existen funciones inteligentes dispuestas a integrarse, y sintetizar estos

28 Ver <http://www.financialexpress.com/news/india-will-have-3040-smart-cities-in-the-next-five-years/1292527>

29 Insertar los nombres de las ciudades del Adam.

30 “A vision of smarter cities”, IBM. 2009.

31 Ver http://living-planit.com/UOS_overview.htm

datos (producidos por aquellas funciones inteligentes), a algún propósito, como mejorar la eficiencia, equidad, sustentabilidad y calidad de vida”. Al respecto el autor nos propone seis funciones básicas de una Ciudad Inteligente:

Figura 4: “Topología de Funciones de una Ciudad Inteligente”

<p>SMART ECONOMY (Competitiveness)</p> <ul style="list-style-type: none"> ▪ Innovative spirit ▪ Entrepreneurship ▪ Economic image & trademarks ▪ Productivity ▪ Flexibility of labour market ▪ International embeddedness ▪ <i>Ability to transform</i> 	<p>SMART PEOPLE (Social and Human Capital)</p> <ul style="list-style-type: none"> ▪ Level of qualification ▪ Affinity to life long learning ▪ Social and ethnic plurality ▪ Flexibility ▪ Creativity ▪ Cosmopolitanism/Open-mindedness ▪ Participation in public life 	<p>SMART GOVERNANCE (Participation)</p> <ul style="list-style-type: none"> ▪ Participation in decision-making ▪ Public and social services ▪ Transparent governance ▪ <i>Political strategies & perspectives</i>
<p>SMART MOBILITY (Transport and ICT)</p> <ul style="list-style-type: none"> ▪ Local accessibility ▪ (Inter-)national accessibility ▪ Availability of ICT-infrastructure ▪ Sustainable, innovative and safe transport systems 	<p>SMART ENVIRONMENT (Natural resources)</p> <ul style="list-style-type: none"> ▪ Attractivity of natural conditions ▪ Pollution ▪ Environmental protection ▪ Sustainable resource management 	<p>SMART LIVING (Quality of life)</p> <ul style="list-style-type: none"> ▪ Cultural facilities ▪ Health conditions ▪ Individual safety ▪ Housing quality ▪ Education facilities ▪ Touristic attractivity ▪ Social cohesion

Fuente: Batty 2012.

Las preocupaciones en una Ciudad Inteligente trascienden a la mera incorporación de tecnologías para resolver problemas en general, la Ciudad Inteligente, sería una “síntesis de infraestructura (o capital físico), con la disponibilidad y calidad de comunicar conocimiento, e infraestructura social” (Batty *et al*, 2012).

Las seis dimensiones de la Ciudad Inteligente, parecen ser corroboradas por Roche (*et al*, 2013), además de Greco & Bancardino (2012), y tienen su origen en “*Horizon 2020 Urban Forum, Digital Agenda*” (2012), hacia donde apunta las dimensiones de Ciudad Inteligente, es hacia la consideración de lo inteligente de las ciudades, en tanto sus habitantes, más que sus infraestructuras. Al respecto Roche (*et al*, 2013), menciona: “una ciudad no puede satisfacer los requerimientos de ser inteligente en el más comprensivo de los sentidos, a menos que la mejora tecnológica, y las soluciones TIC llevadas adelante, que consideran hacer inteligente la ciudad, además contribuyan al empoderamiento de la población urbana, permitiendo una <Sociedad Habilitada Espacialmente>”.

Como se definió anteriormente, la Sociedad Habilitada Espacialmente o SES (en sus siglas en inglés), se presenta como la condición principal y básica de una Ciudad Inteligente, surgiendo además un concepto más general, ahora asociado al territorio, denominado “*Spatial Intelligence*” o “Inteligencia Espacial” (Gruen, 2013; Rajabifard, 2012; Stefanidis, 2011; Williamson, 2011), referido a la habilidad de una comunidad de usar su Capital Intelectual, sus instituciones de innovación, espacio físico e infraestructura inteligente, para crear ambientes inteligentes que de

forma más inteligente, se enfrenten a desafíos tales como competitividad, sustentabilidad e inclusión (Kominos, 2011).

Por lo tanto, en la búsqueda por una definición de Ciudad Inteligente, transitamos desde la mera incorporación de sensores inteligentes, y la producción de datos a través de ellos, a cómo las comunidades se empoderan con dicha información, para en un proceso de participación efectiva, e incluso autodeterminada, puedan utilizar su capital cultural y redes institucionales, espacio físico y la infraestructura misma, para el desarrollo de una inteligencia cuyo fin último sea la “Equidad”, en el entendido que la Ciudad Inteligente no puede ser inteligente para “algunos”, sino que para todos sus habitantes (Greco & Bancardino, 2014).

La tarea de la construcción de la Ciudad Inteligente, se transforma para todos los actores que intervienen en el espacio urbano, en la búsqueda de un equilibrio entre dos inteligencias: la inteligencia humana, y la inteligencia de las máquinas o redes digitales (Pallot *et al*, 2011), para el logro de dicha equidad, en lo que sería la nueva expresión del “Derecho a la Ciudad” (Greenfield, 2013).

En este contexto, pasamos desde la discusión propia de la Ciudad Informacional (Castells, 2002), donde el acceso se medía en términos de “penetración de artefactos de conectividad entre personas” (computadores, teléfonos inteligentes, tabletas, etc.); a la medición de una condición propia de comunidades que viven en la Sociedad que se transforma en la economía del IoT y el Big Data, que es el derecho que tenemos al uso de la información que las diversas tecnologías producen, registrando nuestra vida diaria.

Como lo expresa Sassen³², usando la expresión del software libre para la construcción de nuestras ciudades: “La Ciudad es un sistema abierto”. La comunicación no puede ser coartada entre la inteligencia humana colaborativa y la inteligencia artificial o digital, a través del impedimento en el acceso a los datos por los gobiernos o las empresas, eso sería deurbanizar nuestras ciudades. Si la Ciudad posee un sistema cerrado en cuanto a la tecnología que usamos para su “registro y procesamiento”, estaría condenada a la obsolescencia, tal como nuestros computadores o nuestros *Smartphones*.

II.b.- Proyectos de Ciudad Inteligente en Santiago: capas de inteligencia en una Ciudad Segregada Socioespacialmente

Resulta interesante repasar las visiones y definiciones que los proyectos chilenos sobre Ciudades Inteligentes, presentan. En términos generales corresponden a definiciones operativas sobre lo que debiera ser un proyecto de Ciudad Inteligente para Chile, concentrada cada iniciativa en una arista muy específica:

³² “*The Future of Smart Cities*”, 2011. (Conferencia) Ver <http://videos.liftconference.com/video/2895375/saskia-sassen-the-future-of-smart-cities>

- i. “(las Ciudades Inteligentes) son ciudades pensadas para las personas, que a través del trabajo colaborativo, herramientas de innovación y tecnologías, implementen soluciones integrales y sustentables para entregar mejores servicios para el transporte.” (“Estrategia de Ciudad Inteligente para el Transporte, Chile 2020”. Subsecretaría de Transporte, Gobierno de Chile, 2014).
- ii. “Son ciudades que, por medio de la aplicación de la tecnología en sus diferentes ámbitos, se transforman en localidades más eficientes en el uso de sus recursos, ahorrando energía, mejorando los servicios entregados y promoviendo un desarrollo sustentable, solucionando los principales problemas a los que se ven enfrentados los ciudadanos; logrando de esta forma, que las personas mejoren su calidad de vida” (“Ranking de Ciudades Inteligentes en Chile”. Fundación País Digital/Universidad del Desarrollo -UDD, 2014)
- iii. “Una ciudad se puede definir como "inteligente" o como "inteligentemente eficiente", cuando la inversión social, el capital humano, las comunicaciones, y las infraestructuras, conviven de forma armónica con el desarrollo económico sostenible, apoyándose en el uso y la modernización de nuevas tecnologías (TIC), y dando como resultado una mejor calidad de vida y una gestión prudente de los recursos naturales, a través de la acción participativa y el compromiso de todos los ciudadanos.” (“Ciudad Inteligente y con sentido Ciudadano”. Subsecretaría de Telecomunicaciones, Gobierno de Chile, 2014)
- jjj. “Smart City Santiago es un proyecto pensado para mitigar el impacto de los emplazamientos humanos en el entorno y es el primer núcleo urbano sustentable de Chile. La implementación de una *Smart Grid* o red inteligente es uno de los aspectos centrales para concebir un proyecto urbano que opere de manera integrada. A través de las redes inteligentes, la energía eléctrica puede ser gestionada a distancia y de manera más eficiente”. (“Smart City Santiago, Ciudad Empresarial, 2014).

En las tres definiciones de las iniciativas chilenas sobre ciudades inteligentes, existe un consenso en que la tecnología y sus nuevas aplicaciones, ahora en la ciudad, serían la respuesta a problemas como la energía, el transporte o el desarrollo sustentable, a través de una alianza entre gobierno, empresa, academia y ciudadanía. Sin embargo, a excepción del documento presentado por el MOP, “Estrategia de Ciudad Inteligente para el Transporte, Chile 2020”, si bien se mencionan conceptos como “Datos”, “Acceso Libre a los Datos”, no se mencionan conceptos aún más importantes en una Ciudad Inteligente como “Equidad”, “Sociedad Habilitada Espacialmente”, o “Inteligencia Espacial”. Tal parece ser que el uso de tecnologías, en el sentido de “artefactos”, y la libre disposición de datos, sin incentivar a la población a su uso, facilitaría de algún modo condiciones que produciría la Ciudad, de carácter negativo, y que el hecho de incorporar tecnología las resolvería.

Sin embargo, lo que es aún más preocupante es que no existe una concepción de lo que se entiende por Ciudad en estos proyectos. El espacio urbano parece ser una función específica como la Movilidad o la Energía, en una visión funcionalista de la ciudad industrial planificada por *Le Corbusier* (De Stefani, 2009).

A este respecto ¿Qué ocurre al tratar de implementar un proyecto de *Smart City*, sin hacerse cargo de la Ciudad en la que se intenta desarrollar dicho emprendimiento? En Santiago, por ejemplo, la Ciudad Inteligente debe hacerse cargo de una ciudad en donde existe poca o nula mezcla de población de distintos niveles o estratos socioeconómicos, representativo también de las principales ciudades latinoamericanas (Rodríguez, 2001). Dicha segregación socioespacial de Santiago, provoca una serie de efectos en diversas dimensiones sociales, demostrado también en un distinto nivel de acceso a TICs (Hernández, 2006), una de las cualidades en la que parece basarse la Ciudad Inteligente, asumiendo al parecer que la población tiene el mismo nivel de acceso a Internet, y artefactos de comunicación, que en este caso, serían de vital importancia para poder acceder a los datos producidos por los habitantes de la Ciudad Inteligente.

En específico, propuestas como el transporte o la movilidad urbana de Ciudad Inteligente en Chile, como la presentada por el MOP, parecen obviar fenómenos como las migraciones diarias a lugares de trabajo y estudio (Elgueta 2006, Ortiz, 2006), o el desarrollo incipiente de policentros en el Gran Santiago (Scolano, 2005; Ortíz, 2005).

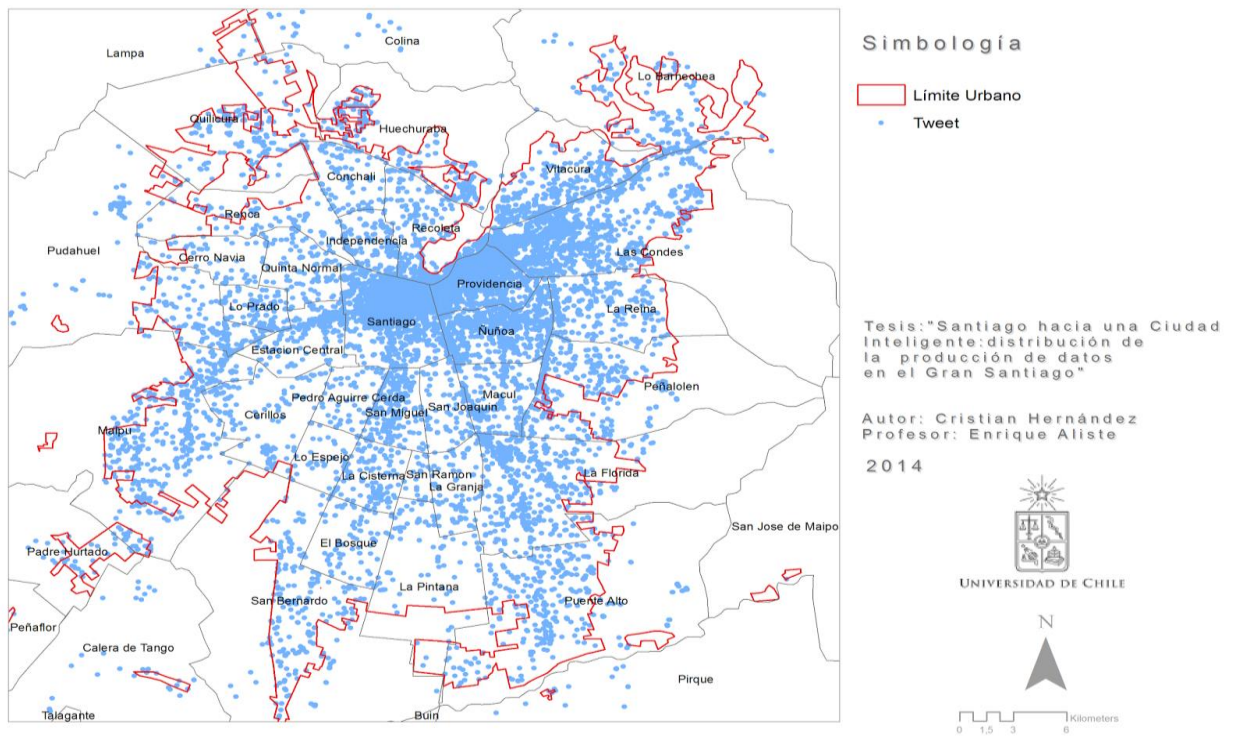
Finalmente, y como se desarrollará más adelante en ésta investigación, no es posible generar un proyecto de Ciudad inteligente, si no se considera la ciudad y el poder que reside en su diseño y planificación (Lefebvre, 1968), en tanto el territorio es resultado de la apropiación social que hacemos del espacio (Aliste, 2010), y que el concepto en sí mismo de Ciudad Inteligente, apela por ende, a una apropiación que reviste equidad, ahora desde el punto de vista del acceso a datos para “*hackear*” la ciudad.

Capítulo III: Desarrollo

III.a.- La producción de Datos Activos en el Gran Santiago: *Tweets*.

Se recolectaron *tweets* (mensajes personales de cada usuario en Twitter), para un día laboral aleatorio, en este caso el 1 de Septiembre del 2014, pudiendo recopilar alrededor de 74 mil *tweets*, representados en puntos para el área estudiada. En un primer momento, se identifica que en la planta urbana consolidada del Gran Santiago, se concentra el 88% de la producción de datos en Twitter³³, lo que habla de un fenómeno que se circunscribe en mayor medida a los límites de la planta urbana, como se aprecia en la cartografía:

Mapa 3: “Producción total de *Tweets* para un día laboral en el Gran Santiago”

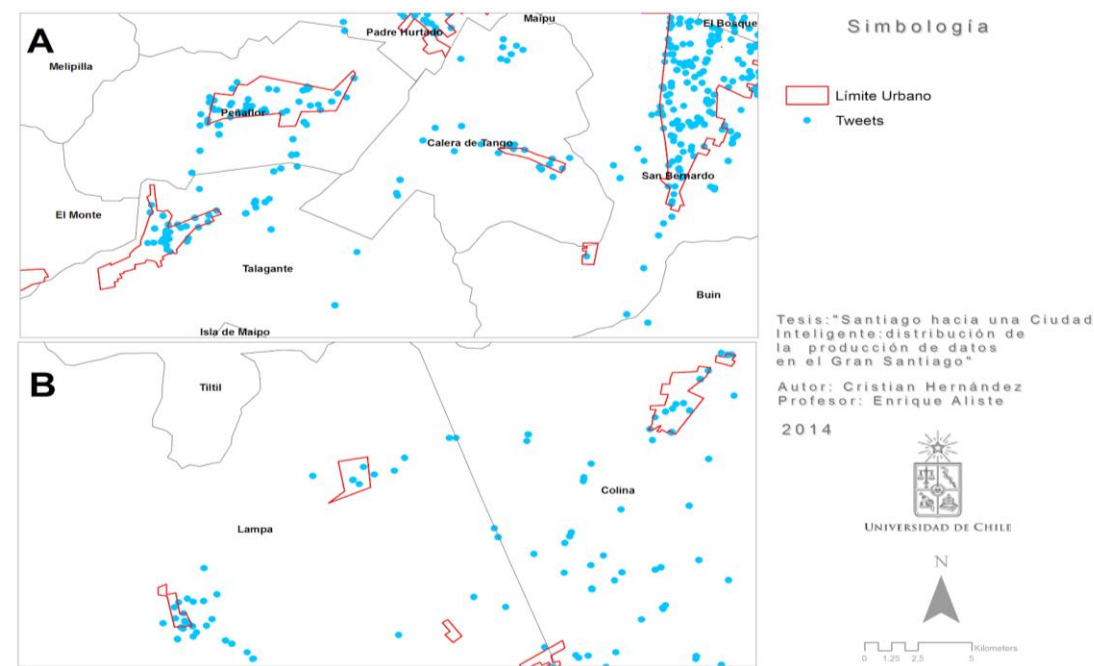


Elaboración propia con datos recolectados desde Twitter API, 2014.

Sin embargo, llama la atención, que a pesar de la alta concentración de *tweets* en la planta urbana consolidada del Gran Santiago, existe una tendencia a la concentración de *tweets* en plantas urbanas dispersas en la Región Metropolitana, como es el caso de Padre Hurtado, Peñaflo, Talagante y Calera de Tango (A), y en menor medida, Lampa y Colina (B). Tal como se aprecia en el siguiente mapa:

33 Cálculo realizado en ArcMap: Estadística Zonal.

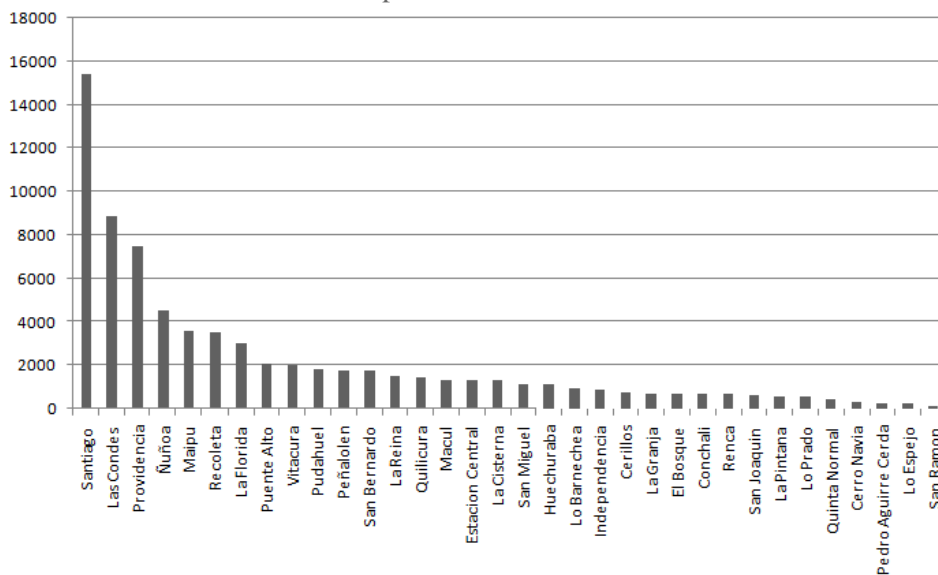
Mapa 4: “Comparativa de *Tweets* producidos en Plantas Urbanas fuera del Gran Santiago”



Elaboración propia con datos recolectados desde Twitter API, 2014..

En términos de producción de datos desde *Twitter*, las comunas de Santiago, Las Condes, Providencia, Ñuñoa y Maipú, lideran. Mientras tanto, comunas como San Ramón, Lo Espejo, Pedro Aguirre Cerda o Cerro Navia, son las comunas que menos datos producen. El orden decreciente de producción de datos desde *Twitter* por comuna se aprecia en el siguiente gráfico:

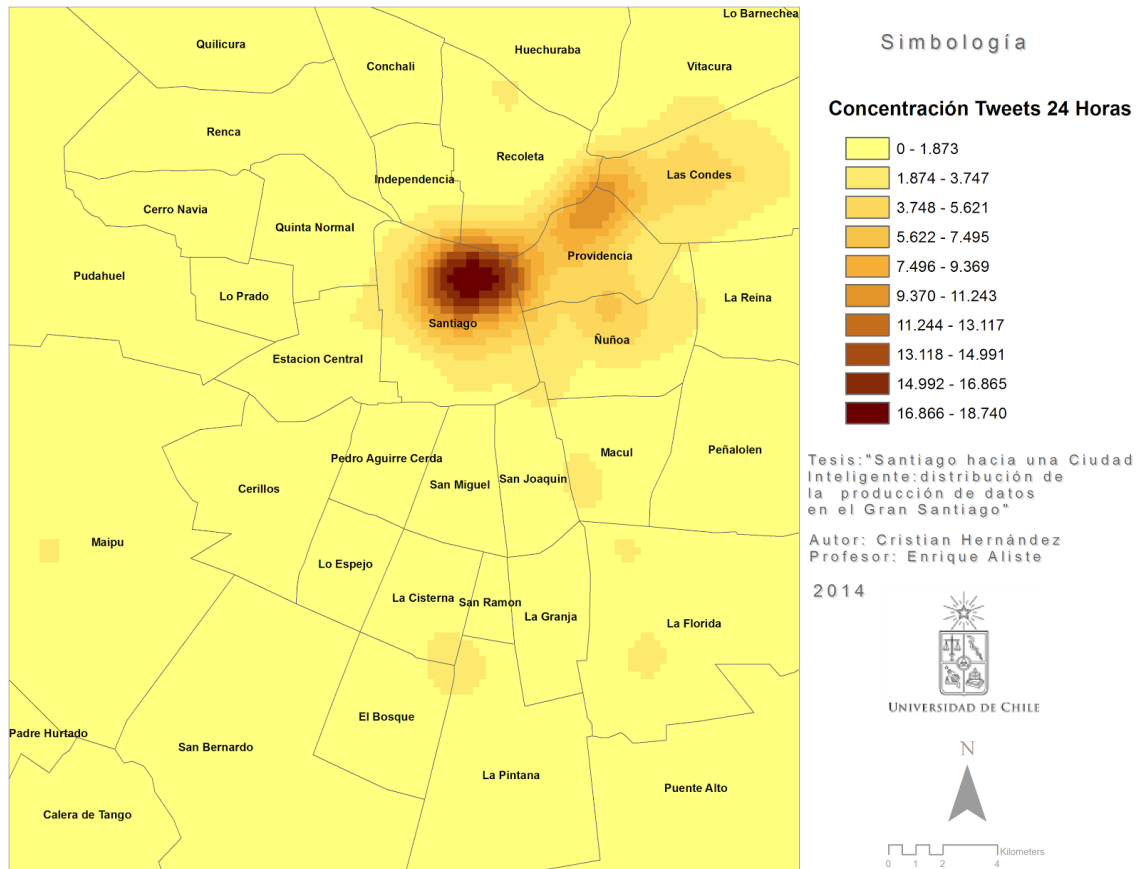
Gráfico 1: “Producción de *Tweets* por Comuna en un día laboral en el Gran Santiago”



Elaboración propia en base a datos recolectados desde Twitter API. 2014.

Como se observa en el Mapa 5 y en el Gráfico 1, existe una marcada concentración de la producción de datos activos en la comuna de Santiago, una concentración que se expande hacia Providencia y Las Condes, y en menor medida hacia Ñuñoa. Resulta interesante observar cómo la “mancha” de producción de datos activos, tiene una presencia dispersa en pequeñas concentraciones, como lo vemos en San Joaquín/Macul; dos concentraciones en La Florida; un nodo de concentración entre las comunas de San Ramón, El Bosque, La Pintana y La Cisterna; un nodo en Maipú; una pequeño en Pudahuel; y finalmente uno presente en Quilicura y otro en el límite entre Recoleta y Huechuraba.

Mapa 5: “Concentración Kernel de *Tweets* para un día Laboral en el Gran Santiago”



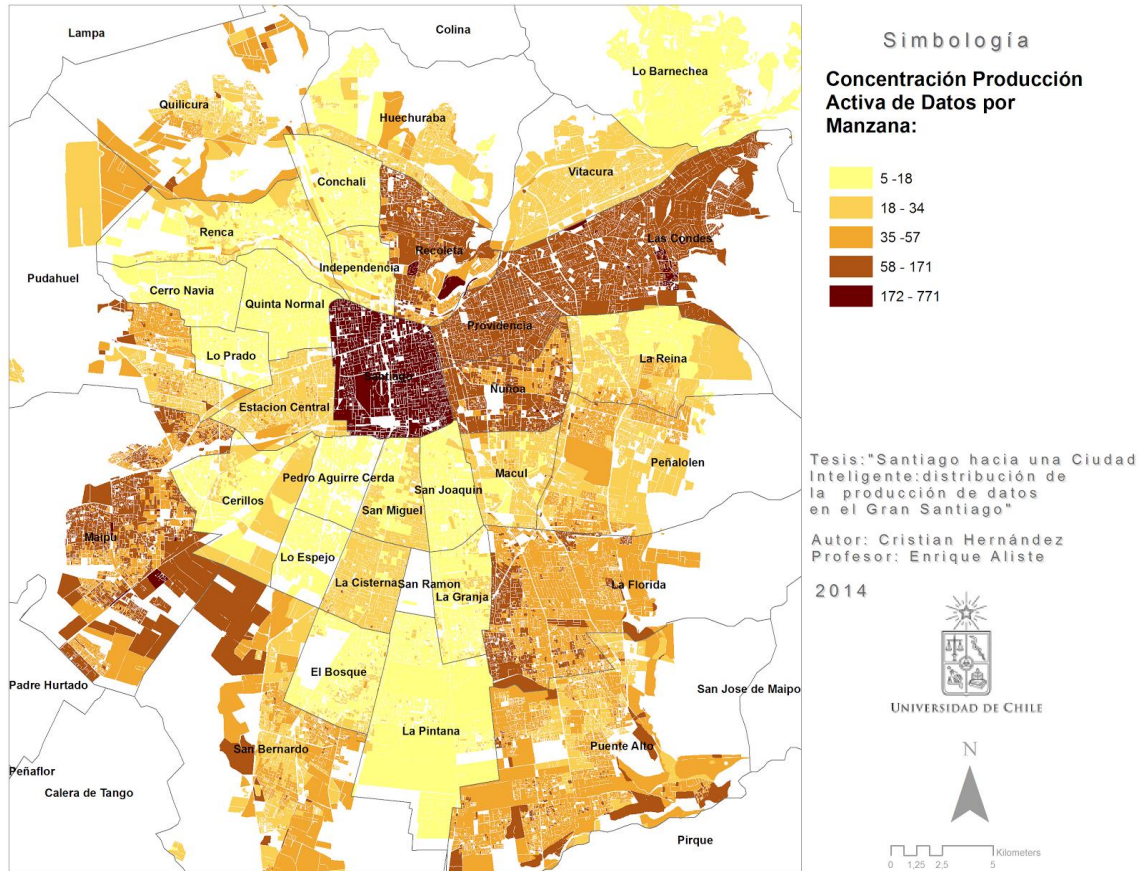
Elaboración propia con datos recolectados desde Twitter API, 2014.

En término de la producción de datos desde *Twitter*, y su expresión a nivel de Manzanas en el Gran Santiago, podemos ahora observar cómo se distribuye los datos, ahora al interior de cada comuna. Llama la atención que en Santiago, la totalidad de los que se producen a nivel de manzanas, presentan las más altas concentraciones 172-771 T/M³⁴, lo que invitaría en análisis posteriores a realizar una categorización específica para dicha comuna. Lo mismo ocurre con la comuna de Providencia, aunque en un nivel de concentración menor, donde todas las Manzanas de Providencia, presentan una concentración entre 58-171 T/M. Las Condes, por su parte,

³⁴ T/M = Tweets por Manzana.

presenta algunas Manzanas con niveles de concentración entre 172-771 T/M, con todas sus otras manzanas en un nivel 58-171 T/M. Casos más dispersos, con densidades altas, resultan Recoleta, Ñuñoa y Maipú; con manzanas entre los niveles de concentración 18-34 T/M y 58-171 T/M. En el otro extremo, comunas como Lo Espejo o Cerro Navia, presentan la totalidad de sus Manzanas en niveles de concentración de la producción de datos en un nivel de concentración 5-18 T/M. El panorama general se puede apreciar en el siguiente mapa:

Mapa 6: “Concentración Zonal de *Tweets* para un día Laboral a nivel de Manzana en el Gran Santiago”



Elaboración propia con datos recolectados desde Twitter API, 2014.

En términos generales, podemos apreciar que la producción de datos desde *Twitter* se concentra en un 80% en 14 comunas de las 34 que componen el Gran Santiago. Por otro lado, la comuna que produce más datos (Santiago), produce 107 veces más datos activos que San Ramón, la comuna que produce menos datos. Si realizamos el corte en el 50% de los datos producidos, nos encontramos que tan solo cuatro comunas del Gran Santiago, producen la mitad de los datos activos en un día laboral, las cuales son: Santiago, Las Condes, Providencia y Ñuñoa. Esto explica la alta concentración observada, tanto en el análisis de concentraciones *Kernel*, como en el análisis por Manzanas. La cantidad de datos producidos, el porcentaje que representa en el Gran Santiago, y el porcentaje acumulado por comunas, lo vemos en la siguiente tabla:

Tabla 1: “Producción Acumulada de *Tweets* en un día laboral por Comuna en el Gran Santiago”

Porcentaje de Producción de Datos Activos a nivel Comunal en el Gran Santiago									
N°	COMUNA	QDA	% del Total	% Acumulado	N°	COMUNA	QDA	% del Total	% Acumulado
1	Santiago	15424	21%	21%	18	San Miguel	1142	2%	87%
2	Las Condes	8857	12%	33%	19	Huechuraba	1119	2%	89%
3	Providencia	7511	10%	43%	20	Lo Barnechea	913	1%	90%
4	Ñuñoa	4523	6%	49%	21	Independencia	910	1%	91%
5	Maipú	3572	5%	54%	22	Cerrillos	755	1%	92%
6	Recoleta	3524	5%	59%	23	La Granja	701	1%	93%
7	La Florida	3046	4%	63%	24	El Bosque	695	1%	94%
8	Puente Alto	2100	3%	66%	25	Conchalí	684	1%	95%
9	Vitacura	2024	3%	69%	26	Renca	683	1%	96%
10	Pudahuel	1828	2%	71%	27	San Joaquín	619	1%	97%
11	Peñalolén	1780	2%	74%	28	La Pintana	560	1%	97%
12	San Bernardo	1746	2%	76%	29	Lo Prado	542	1%	98%
13	La Reina	1505	2%	78%	30	Quinta Normal	410	1%	99%
14	Quilicura	1475	2%	80%	31	Cerro Navia	322	0%	99%
15	Macul	1353	2%	82%	32	Pedro Aguirre Cerda	255	0%	99%
16	Estación Central	1322	2%	84%	33	Lo Espejo	242	0%	100%
17	La Cisterna	1308	2%	85%	34	San Ramón	143	0%	100%

Elaboración propia con datos recolectados desde Twitter API, 2014.

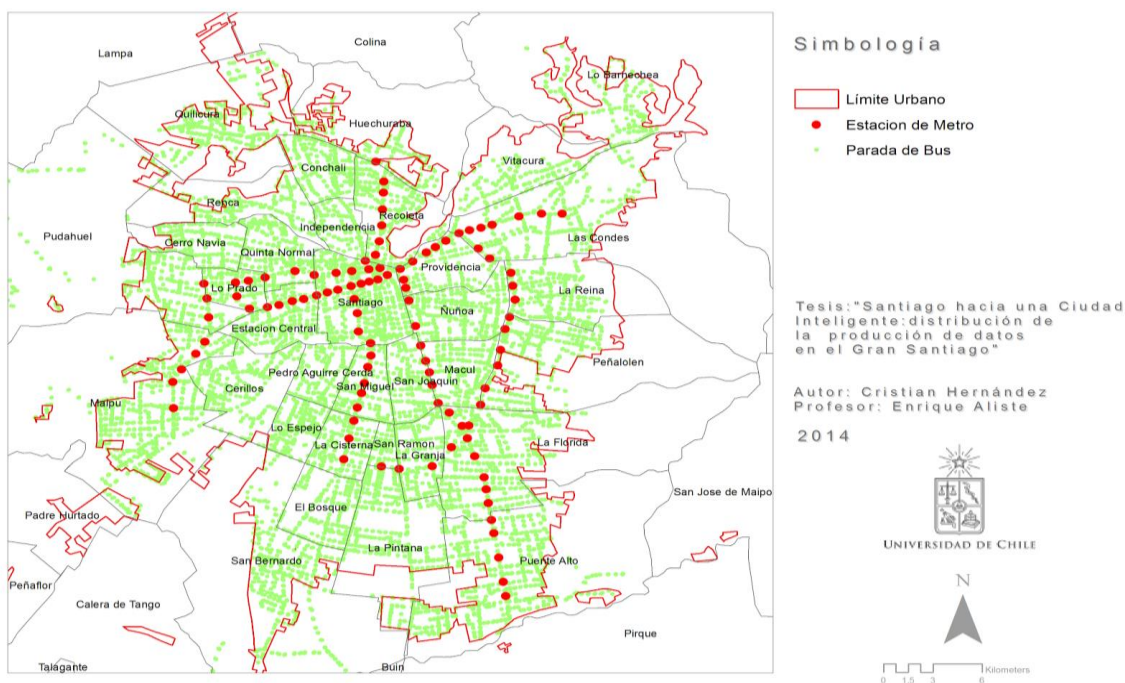
III.b.- La producción de Datos Pasivos en el Gran Santiago: Validaciones BIP.

El entramado de paraderos del Transantiago y las estaciones de Metro, corresponden a la red de puntos en los cuales se disponen los sensores que captan datos provenientes desde las Validaciones BIP de los usuarios del sistema de transporte de Santiago. En total, sumando Buses y Metro, los puntos (paradas y estaciones) suman 10.651³⁵, en donde se realizan un total de 5.471.380 validaciones para el día estudiado³⁶.

35 100 Estaciones de Metro, y 10.551 Paraderos de Bus.

36 Día Laboral Promedio entregados como datos abiertos de libre acceso por la Subsecretaría de Transportes a través del portal “datos.gob.cl”

Mapa 7: “Localización de Paradas de Bus y Estaciones de Metro, Transantiago”



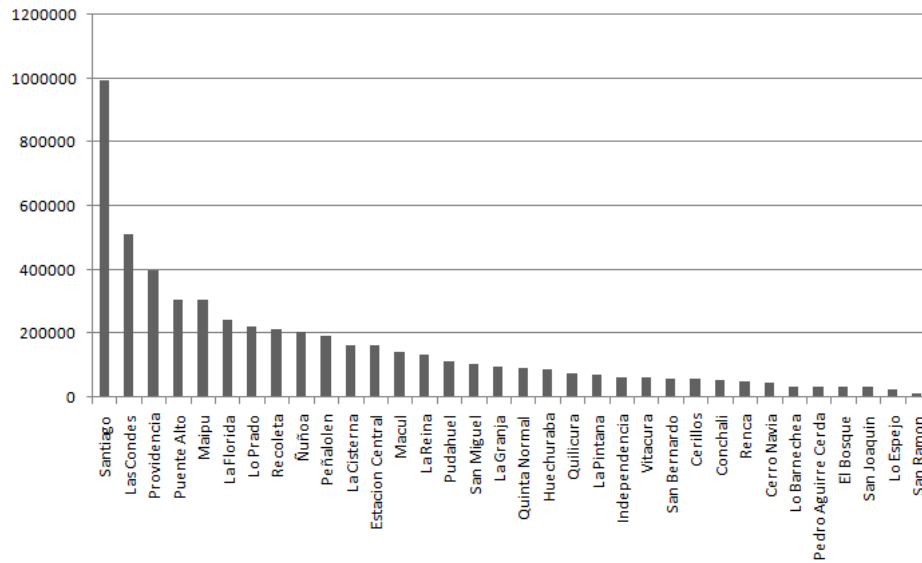
Mapa 8: Elaboración propia con datos recolectados desde Datos.gob.cl

A excepción de ciertos “corredores”, como en Pudahuel y San Bernardo, que sobrepasan el límite urbano, se aprecia el hecho que Transantiago es un sistema de transporte público pensado para la planta urbana consolidada, que representa el Gran Santiago.

En este caso, y a diferencia de la información proveniente desde Twitter, donde una coordenada en el espacio representa un dato (uno a uno), cada punto en el área estudiada correspondiente a una estación de Metro o a una Parada de Bus, tiene un peso distinto (uno es a muchos), el cual está determinado por la cantidad de validaciones realizadas en cada uno de ellos. Es decir, el número de validaciones realizadas en cada parada de Bus y estación de Metro, entregan un peso distinto, que provoca la forma de la concentración de los datos. Por ejemplo, tenemos paradas de Bus con 3 validaciones BIP en el día estudiado, mientras que una estación de Metro puede superar las 74.133 validaciones al día, como la estación “La Cisterna”.

Las frecuencias de las “Validaciones BIP” para el Gran Santiago por Comuna, son las siguientes:

Gráfico 2: “Validaciones totales BIP por Comuna en un día laboral, Gran Santiago”

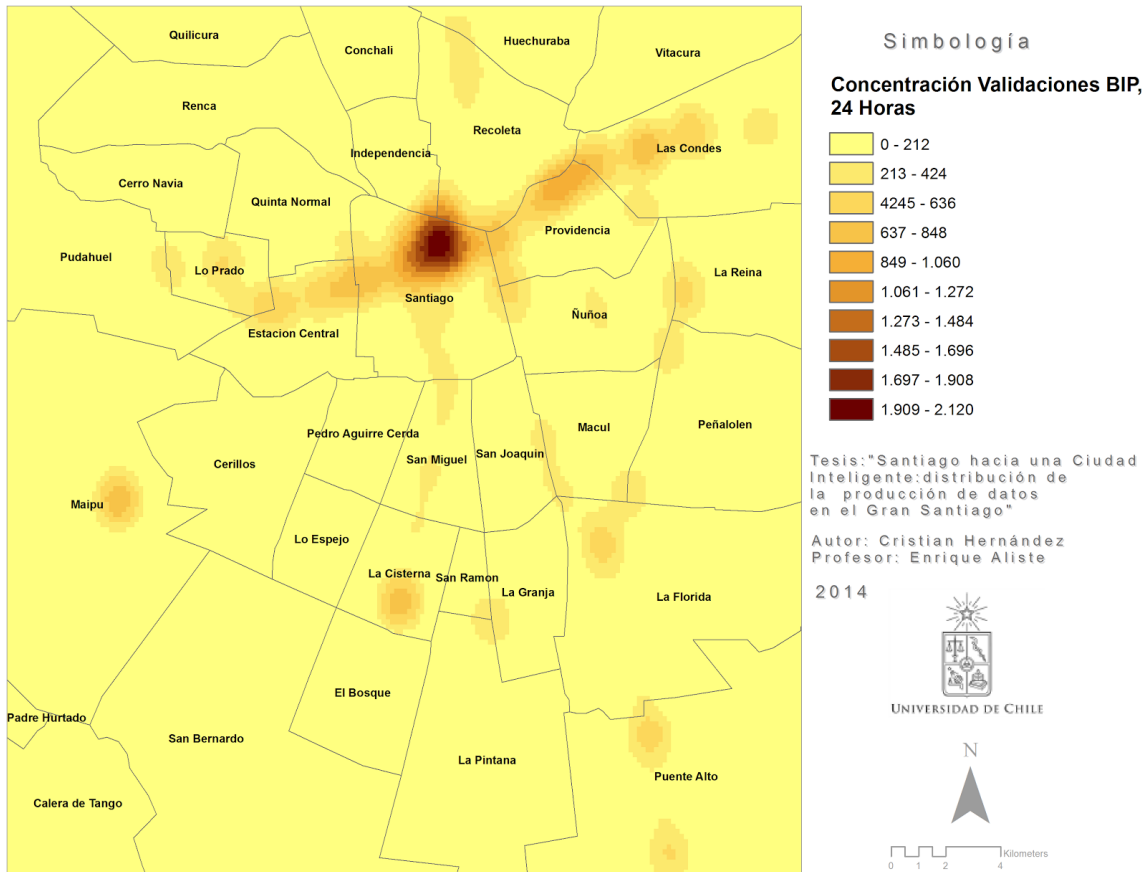


Elaboración propia con datos recolectados desde Datos.gob.cl

Nuevamente Santiago es la comuna que en término de frecuencias, concentra la mayor cantidad de validaciones BIP, provenientes desde Buses y Metro, seguido también en un caso que se reitera por Las Condes y Providencia. Ahora bien, en el caso de la producción pasiva de datos, aparecen comunas como Puente Alto y Maipú con altas frecuencias en las validaciones BIP, probablemente por la gran cantidad de población que habita dichas comunas, comparadas con las comunas del resto del Gran Santiago. En el otro extremo, tenemos comunas como San Ramón (nuevamente en última posición), Lo Espejo, San Joaquín y el Bosque, con las menores frecuencias en las validaciones BIP.

En términos de la concentración de la producción de datos provenientes desde Validaciones BIP, tenemos que se presenta un patrón que sigue en términos generales, el corredor “Línea 1” del Metro, o el eje vial de Buses, “Pajaritos, Alameda, Providencia, Apoquindo”; con una alta concentración de validaciones BIP en la comuna de Santiago, como ya se ha mencionado. Así mismo, se abren “troncales” de datos pasivos al sur del eje ya mencionado, en específico el comprendido por “San Diego-Gran Avenida”, “Vicuña Mackenna”, y “Américo Vespucio”. Otros puntos relevantes desde el punto de vista de la concentración, los encontramos en Maipú (Metro Plaza de Maipú), Pudahuel (estación Pudahuel), o Recoleta (entre estación Dorsal y Américo Vespucio Norte). La concentración de datos producidos en las Validaciones BIP, la podemos observar en el siguiente mapa:

Mapa 8: “Concentración Kernel Validaciones BIP en un día laboral en el Gran Santiago”



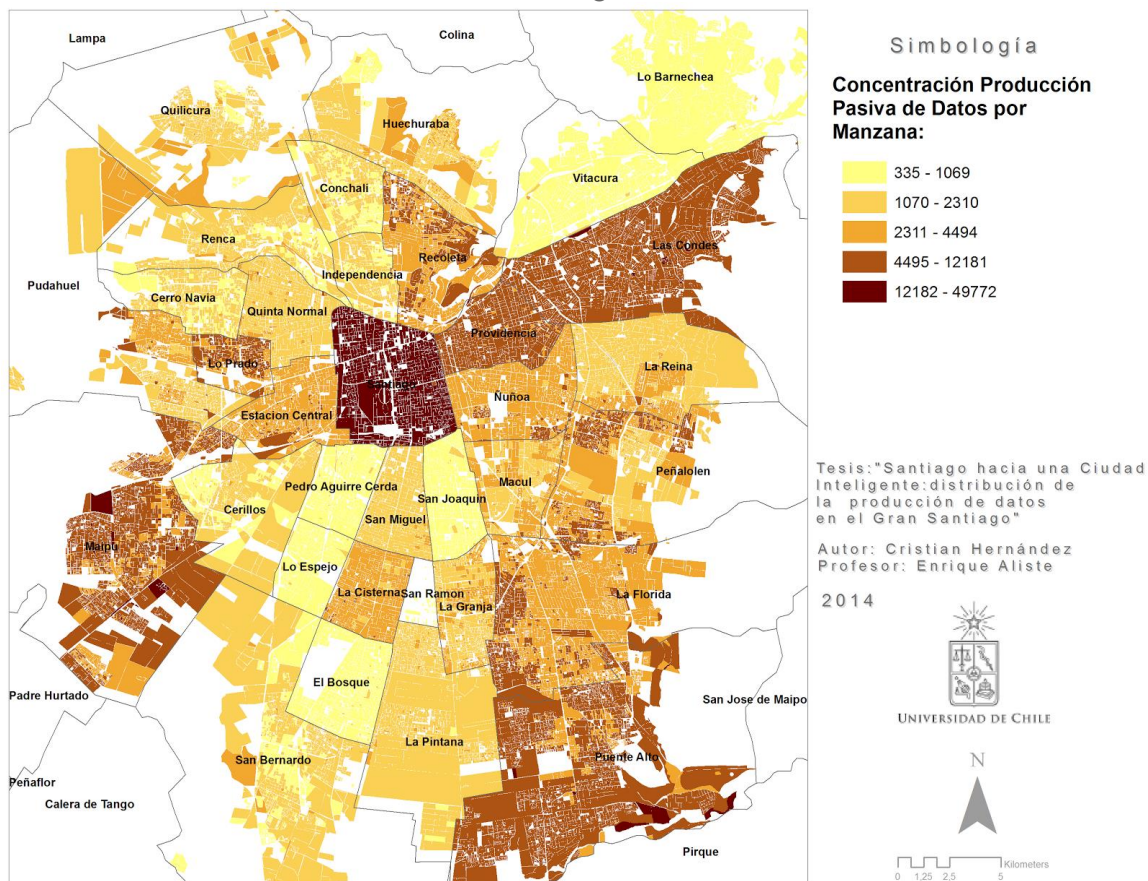
Elaboración propia con datos recolectados desde Datos.gob.cl

En términos de la concentración de la producción de datos provenientes desde Validaciones BIP, tenemos nuevamente una homogeneidad en las manzanas de la comuna de Santiago, en tanto todas ellas poseen los más altos niveles de concentración de la producción de datos. A grandes rasgos, podemos observar que existen cuatro grandes zonas de validaciones BIP³⁷, fuera de la concentración mayor que corresponde a Santiago, a saber: (i) Eje Providencia - Las Condes; (ii) Eje La Florida - Puente Alto; (iii) Eje Estación Central - Maipú; y (iv) Eje Recoleta - Huechuraba.

Podemos apreciar dicho patrón territorial, en la siguiente cartografía:

37 Cabe recordar que la presente investigación no tiene como fin describir un patrón de movilidad urbana, en términos de “Origen-Destino”, solamente considerar la cantidad de validaciones BIP realizadas en cada comuna, cuyos viajes pueden ser de término o inicio.

Mapa 9: “Concentración Zonal Validaciones BIP en un día laboral a nivel de Manzana, Gran Santiago”



Elaboración propia con datos recolectados desde Datos.gov.cl

Del total de 34 comunas que componen el Gran Santiago, el 80% de la producción de datos pasivos, se presenta en 15 comunas. Si hablamos del 50% del total de las validaciones BIP, consideradas como la información pasiva en este estudio, podemos decir que se concentra en tan solo 6 comunas del Gran Santiago. La comuna de Santiago, que es la comuna que concentra la mayor parte de la producción de datos pasivos, produce en total 75 veces más datos que San Ramón, la comuna que produce menor cantidad de datos pasivos.

Tabla 2: “Validaciones BIP acumuladas por Comuna en un día laboral, Gran Santiago”

Porcentaje de Producción de Datos Pasivos a nivel Comunal en el Gran Santiago									
N°	COMUNA	QDP	% del Total	% Acumulado	N°	COMUNA	QDP	% del Total	% Acumulado
1	Santiago	995441	18%	18%	18	Quinta Normal	93581	2%	85%
2	Las Condes	511683	9%	28%	19	Huechuraba	87042	2%	87%
3	Providencia	397047	7%	35%	20	Quilicura	76594	1%	88%
4	Puente Alto	306005	6%	41%	21	La Pintana	70220	1%	89%
5	Maipú	304479	6%	47%	22	Independencia	63738	1%	91%
6	La Florida	241589	4%	51%	23	Vitacura	63673	1%	92%
7	Lo Prado	223532	4%	55%	24	San Bernardo	60826	1%	93%
8	Recoleta	214213	4%	59%	25	Cerrillos	56970	1%	94%
9	Ñuñoa	204384	4%	63%	26	Conchalí	53426	1%	95%
10	Peñalolén	192330	4%	66%	27	Renca	51909	1%	96%
11	La Cisterna	165603	3%	70%	28	Cerro Navia	47061	1%	97%
12	Estación Central	162127	3%	73%	29	Lo Barnechea	34965	1%	97%
13	Macul	141419	3%	75%	30	Pedro Aguirre Cerda	34409	1%	98%
14	La Reina	134964	2%	78%	31	El Bosque	32736	1%	99%
15	Pudahuel	112944	2%	80%	32	San Joaquín	31766	1%	99%
16	San Miguel	105240	2%	82%	33	Lo Espejo	24679	0%	100%
17	La Granja	94653	2%	83%	34	San Ramón	13300	0%	100%

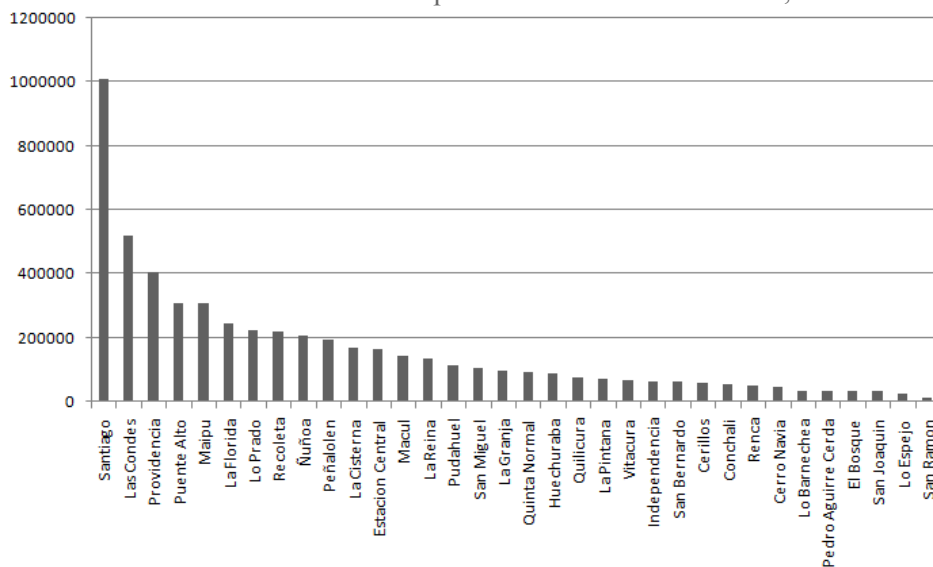
Elaboración propia con datos recolectados desde Datos.gob.cl, 2014.

III.c.- La producción Total de Datos en el Gran Santiago

En términos agregados, sumando la cantidad de datos producidos de forma activa (tweets) y de forma pasiva (Validaciones BIP) para las comunas del Gran Santiago, la primera observación es la comparación en la cantidad de datos producidos activamente y pasivamente. Concretamente, para un día laboral, en Santiago se producen 73.593 datos activos, provenientes en este caso, de Twitter; mientras que datos producidos por Validaciones BIP son 5.404.548. Es decir, que la producción activa de datos, en términos absolutos, representa tan solo un 1% de la información pasiva. Esto reviste gran relevancia, desde el punto de vista de la consideración de la sensorización en la ciudad, ya que la mayor parte de los datos que se producen en la Ciudad son datos recopilados de forma pasiva, es decir, a través de sensores dispuestos en la planta urbana, y no a través de la interacción entre habitantes de la Ciudad utilizando *Smartphones*, por ejemplo.

Se observa además que la frecuencia de producción de datos totales, tienen una clara influencia en cuanto la producción de datos pasiva para cada comuna, en tanto los datos provenientes desde Validaciones BIP son considerablemente más en términos absolutos, en relación a los datos provenientes desde *Twitter*. Al respecto, tenemos la siguiente distribución:

Gráfico 3: “Producción Total de Datos por Comuna en un día laboral, Gran Santiago”

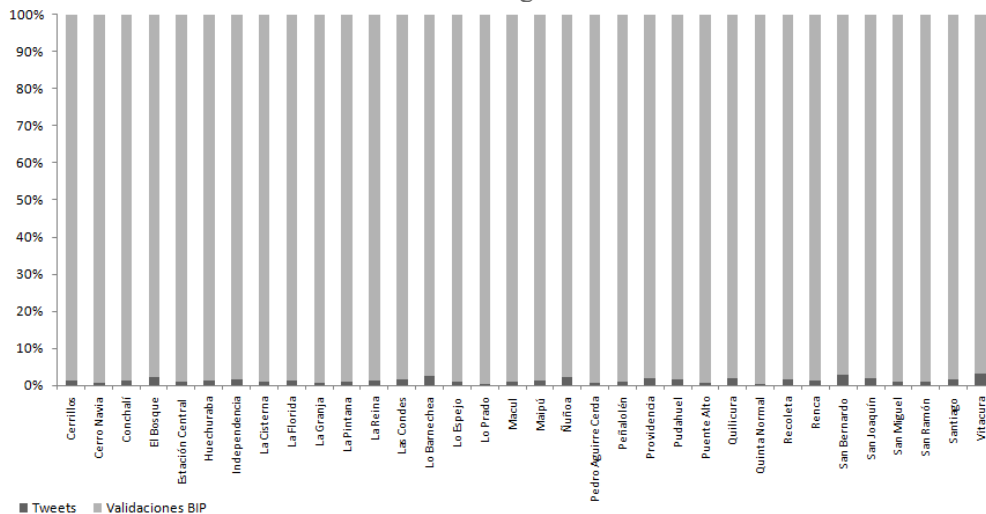


Elaboración propia, 2014.

La comuna de Santiago, tanto en términos activos como pasivos, es la comuna que produce más datos, seguido por Las Condes, Providencia, Puente Alto y Maipú. Las cinco comunas que menos datos producen en el Gran Santiago, son las comunas de Pedro Aguirre Cerda, El Bosque, San Joaquín, Lo Espejo y San Ramón.

A continuación se aprecia lo antes mencionado, en términos del escaso porcentaje que los datos producidos de forma activa, representan en relación a los datos pasivos:

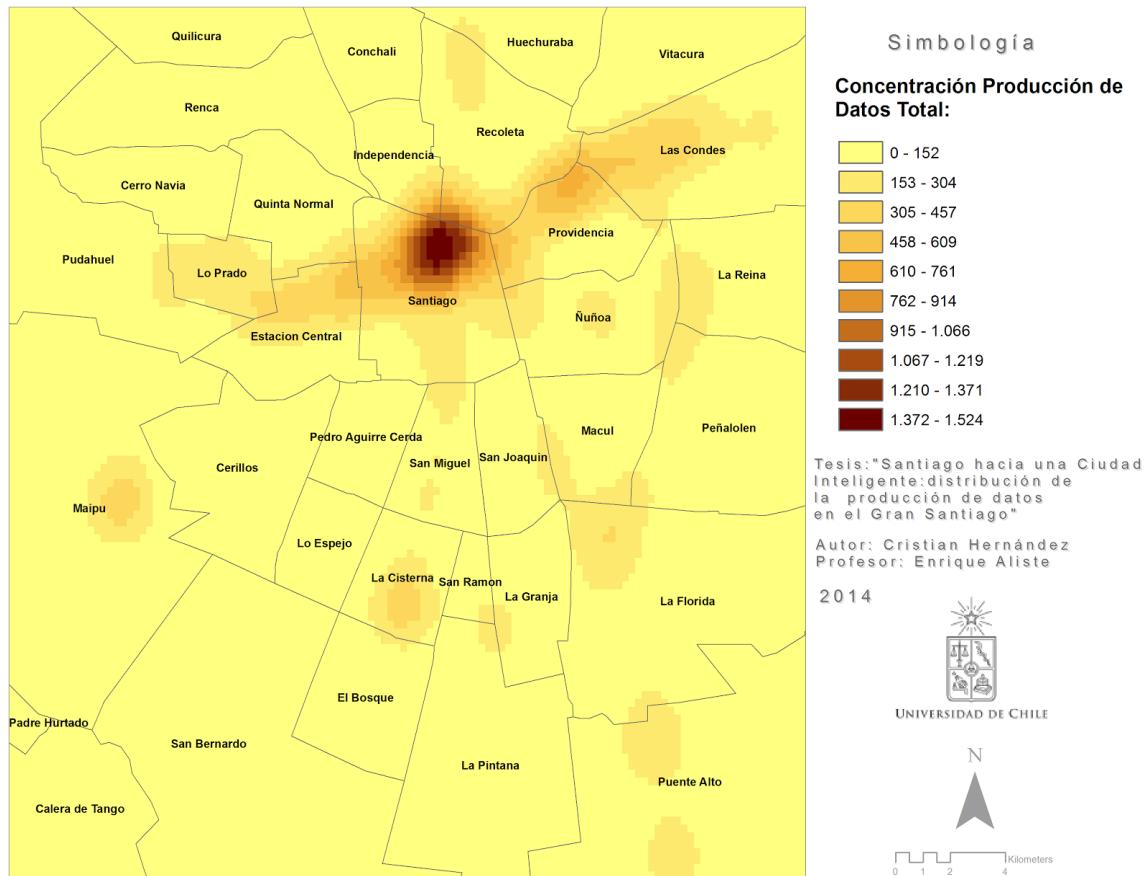
Gráfico 4: “Comparativa Datos Totales Producidos por Comuna en un día Laboral, Gran Santiago”



Fuente: Elaboración Propia, 2014

Si miramos la concentración espacial producida por dicha frecuencia en la producción de datos totales, tenemos lo siguiente:

Mapa 10: “Concentración Kernel Producción Total de Datos para un día laboral en el Gran Santiago”

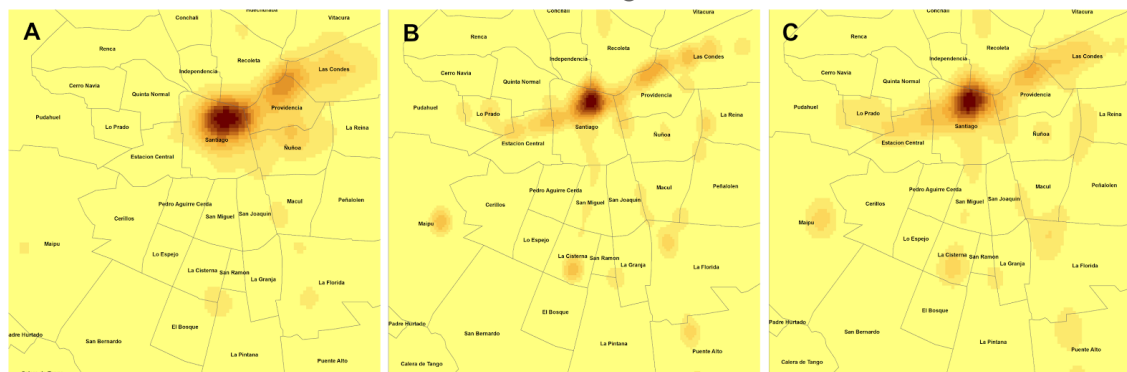


Fuente: Elaboración propia, 2014.

Tenemos un eje Oeste - Este, de producción de datos, que va desde la comuna de Lo Prado, hasta la comuna de Las Condes, con un marcado centro en la comuna de Santiago. Luego, apreciamos los ejes que son producidos por la producción de datos pasivos, como los tres ejes que se dirigen al sur, desde el eje principal “Lo Prado - Las Condes”; a saber: “Santiago - La Cisterna”, “Providencia - Puente Alto”, “Las Condes - La Florida”; y concentraciones aisladas como Maipú o el límite Recoleta - Huechuraba.

En términos comparativos, tenemos que los patrones de concentración de la producción de datos activos, pasivos y totales, es la que se aprecia en el Mapa 12. Podemos ver cómo la producción de datos activa (CPA), posee una concentración mayor a la producción de datos pasivos (CPP), en tanto ambas poseen un centro en Santiago, con la más alta concentración. En tanto la producción total, se muestra siguiendo los patrones de concentración de la producción de datos pasivos, pero abarcando mayor superficie, al agregarle los datos producidos activamente.

Mapa 11: “Comparativa Concentraciones Kernel de Producción de Datos en un día laboral en el Gran Santiago”



A: Tweets; B: Validaciones BIP; C: Producción Total de Datos. Elaboración propia, 2014

Por otro lado, la producción total de datos, generada a partir de datos activos y pasivos, refuerza ciertos “centros periféricos”, como lo es Maipú (Plaza de Maipú); La Cisterna (Metro La Cisterna), Recoleta - Huechuraba (entre Metro Dorsal y Metro Américo Vespucio Norte); la doble concentración producida en Puente Alto (Metro Las Mercedes y Metro Hospital Sotero del Río); la concentración en La Florida (Metro Vicuña Mackenna y Bellavista de La Florida); La Reina - Peñalolén (entre Metro Príncipe de Gales y Grecia); Ñuñoa, fuera de ser parte de la concentración central, posee otro centro en Plaza Ñuñoa; y finalmente el centro San Ramón - La Pintana - La Granja (Estación Santa Rosa).

Resulta interesante el caso de la concentración Plaza Ñuñoa, en tanto es la única que no representa una concentración sin ser una estación de Metro, aunque ya se encuentra proyectada su construcción³⁸, por lo que se esperaría que se transforme en una extensión de la concentración central.

Otro caso interesante es la existencia de dos “corredores sin producción de datos” que se encuentran entre las estaciones de metro Santiago Bueras y Laguna Sur (Maipú hacia Lo Prado), y entre las estaciones “Rojas Magallanes” y “San José de La Estrella” (La Florida). Dichos corredores se presentan como los únicos que no presenta concentración en la producción de información total. Si bien existen estaciones aisladas de Metro que no son parte de las concentraciones, como “Cementerios”, “Rodrigo de Araya” o “La Granja”, dichos ejes representan grandes extensiones de ciudad sin producción de datos. Dichas zonas se pueden apreciar en la siguiente cartografía:

38 Ver <http://diario.latercera.com/2011/03/30/01/contenido/santiago/32-64107-9-nunoo-tendra-seis-nuevas-estaciones-de-metro-en-2017.shtml>

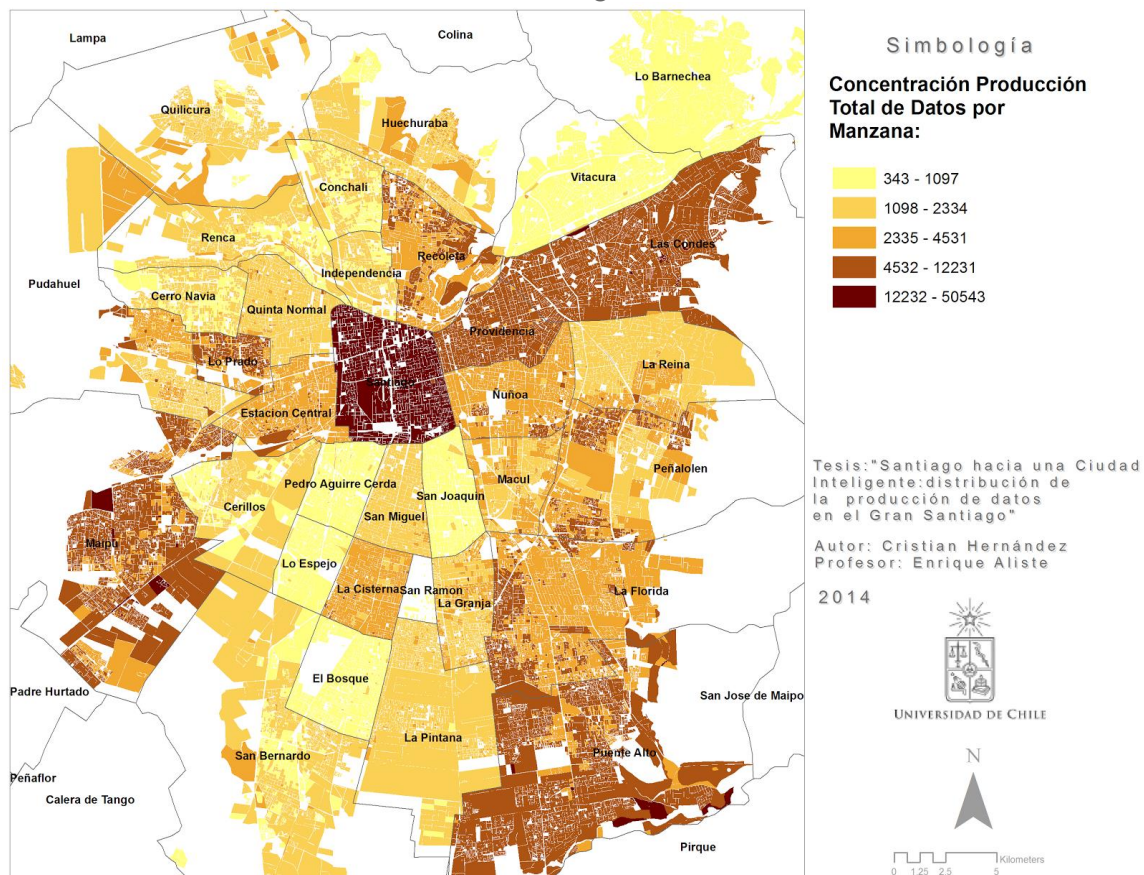
Mapa 12: “Zonas sin Concentración Relevante en la Producción de Datos”



Fuente: elaboración propia, 2014

En términos de la concentración de producción de datos totales, a nivel de manzanas, tenemos que aparte de la concentración mayor de Santiago, se presentan importantes concentraciones, en las comunas de Puente Alto, Maipú y Las Condes. Otras comunas del pericentro también presentan importantes concentraciones, como Recoleta, Estación Central y Lo Prado. Pedro Aguirre Cerda y Lo Espejo, se presentan como un corredor de escasa concentración de producción, en efecto, en su totalidad sus manzanas poseen las más altas concentraciones (Mapa 13).

Mapa 13: “Concentración Zonal Total de Datos para un día laboral por Manzana, en el Gran Santiago”



Elaboración propia, 2014.

A nivel de concentraciones acumuladas por comuna, tenemos que Santiago, la comuna que produce más datos, tanto activos como pasivos y totales, produce 75 veces más datos que San Ramón. El 80% de la producción total de los datos es producida por 15 comunas del Gran Santiago, y el 50% de los datos es producido por 6 comunas, que son Santiago, Las Condes, Providencia, Puente Alto, Maipú y la Florida, tal como se aprecia en la Tabla 3:

Tabla 3: “Producción Total de Datos acumulados para un día laboral en el Gran Santiago”

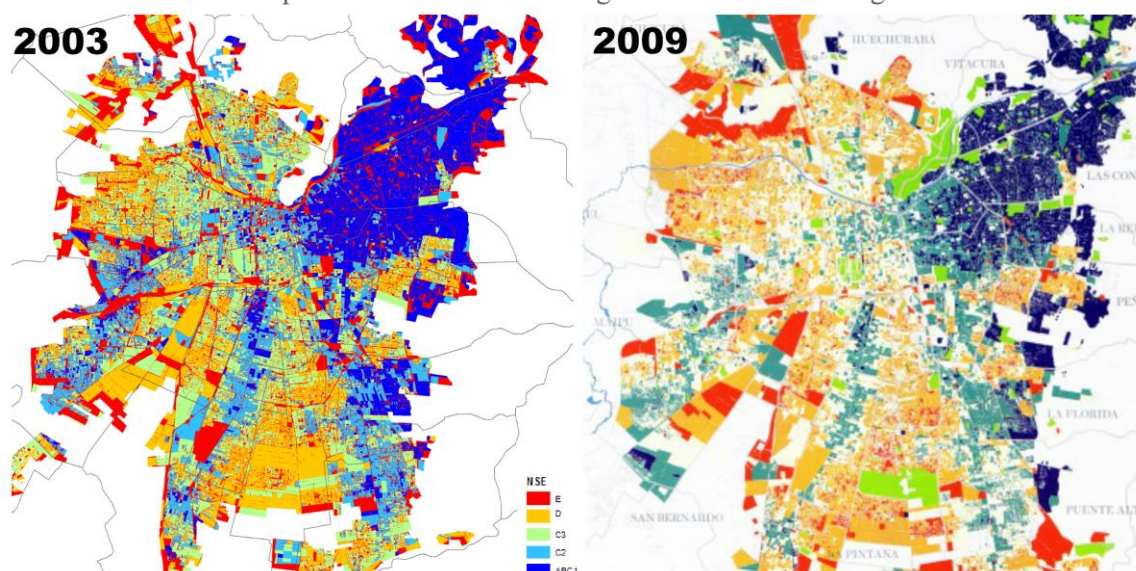
Porcentaje de Producción de Datos Total a nivel Comunal en el Gran Santiago									
N°	COMUNA	QDT	% del Total	% Acumulado	N°	COMUNA	QDT	% del Total	% Acumulado
1	Santiago	1010865	18%	18%	18	Quinta Normal	93991	2%	85%
2	Las Condes	520540	10%	28%	19	Huechuraba	88161	2%	87%
3	Providencia	404558	7%	35%	20	Quilicura	78069	1%	88%
4	Puente Alto	308105	6%	41%	21	La Pintana	70780	1%	89%
5	Maipú	308051	6%	47%	22	Vitacura	65697	1%	91%
6	La Florida	244635	4%	51%	23	Independencia	64648	1%	92%
7	Lo Prado	224074	4%	55%	24	San Bernardo	62572	1%	93%
8	Recoleta	217737	4%	59%	25	Cerrillos	57725	1%	94%
9	Ñuñoa	208907	4%	63%	26	Conchalí	54110	1%	95%
10	Peñalolén	194110	4%	66%	27	Renca	52592	1%	96%
11	La Cisterna	166911	3%	70%	28	Cerro Navia	47383	1%	97%
12	Estación Central	163449	3%	73%	29	Lo Barnechea	35878	1%	97%
13	Macul	142772	3%	75%	30	Pedro Aguirre Cerda	34664	1%	98%
14	La Reina	136469	2%	78%	31	El Bosque	33431	1%	99%
15	Pudahuel	114772	2%	80%	32	San Joaquín	32385	1%	99%
16	San Miguel	106382	2%	82%	33	Lo Espejo	24921	0%	100%
17	La Granja	95354	2%	83%	34	San Ramón	13443	0%	100%

Elaboración propia, 2014.

II.d.- Relación entre la Producción de Datos y el Ingreso en las comunas del Gran Santiago.

Para el análisis de la comparación entre el Ingreso por hogar y la producción de datos en el Gran Santiago, se utiliza la distribución de ingresos a nivel de manzanas provista por Adimark, en base a los datos del Censo 2002. Se observa la siguiente evolución de la distribución de los niveles socioeconómicos entre los años 2003 y 2009:

Mapa 14: “Distribución del Ingreso en el Gran Santiago”



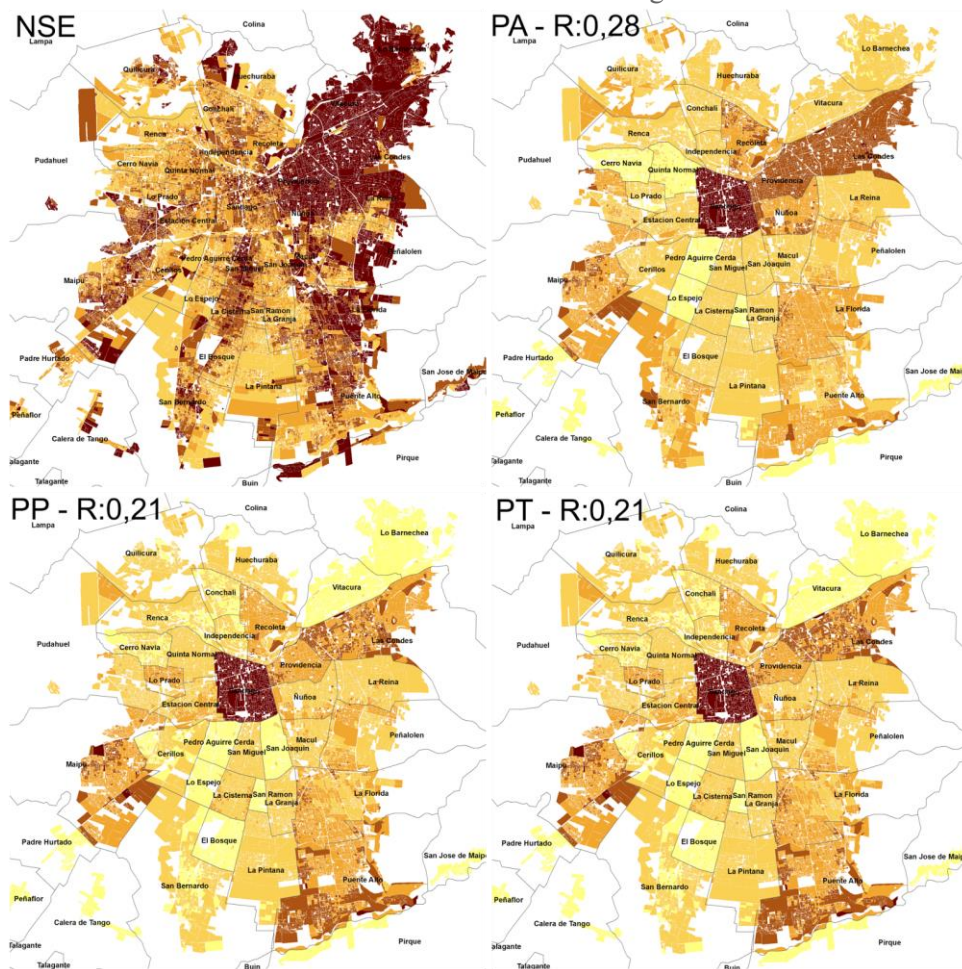
Elaboración propia en base a Geo Adimark 2003 y 2009.

El patrón territorial de la distribución del ingreso en el Gran Santiago, se muestra con una alta concentración hacia el oriente desde las manzanas orientales de la comuna de Providencia, extendiéndose dicha concentración hacia Las Condes, Vitacura, Lo Barnechea; además de las

manzanas localizadas al norte de Ñuñoa y La Reina.. Existe una concentración además, en un corredor entre La Florida y Peñalolén; otro corredor entre San Miguel y La Cisterna. Maipú presenta en su planta urbana una concentración hacia el Sur. Finalmente, una pequeña concentración en Huechuraba. Dicho patrón se mantiene entre los años 2003 y 2009.

Comparando la distribución territorial de los ingresos en el Gran Santiago, se realiza una comparación de los patrones de localización observados, a través del cálculo de la correlación espacial entre los ingresos y la producción de datos activa, pasiva y total. Dicho análisis se resume en la siguiente cartografía:

Mapa 15: “Correlación entre Ingreso Por Hogar y Producción de Datos en un día laboral, a nivel de Manzanas en el Gran Santiago”



Mapa 16: NSE: Nivel Socio Económico; PA: Producción Activa de Datos; PP: Producción Pasiva de Datos; PT: Producción Total de Datos; R: Coeficiente de Correlación.

Fuente: Elaboración propia, 2014.

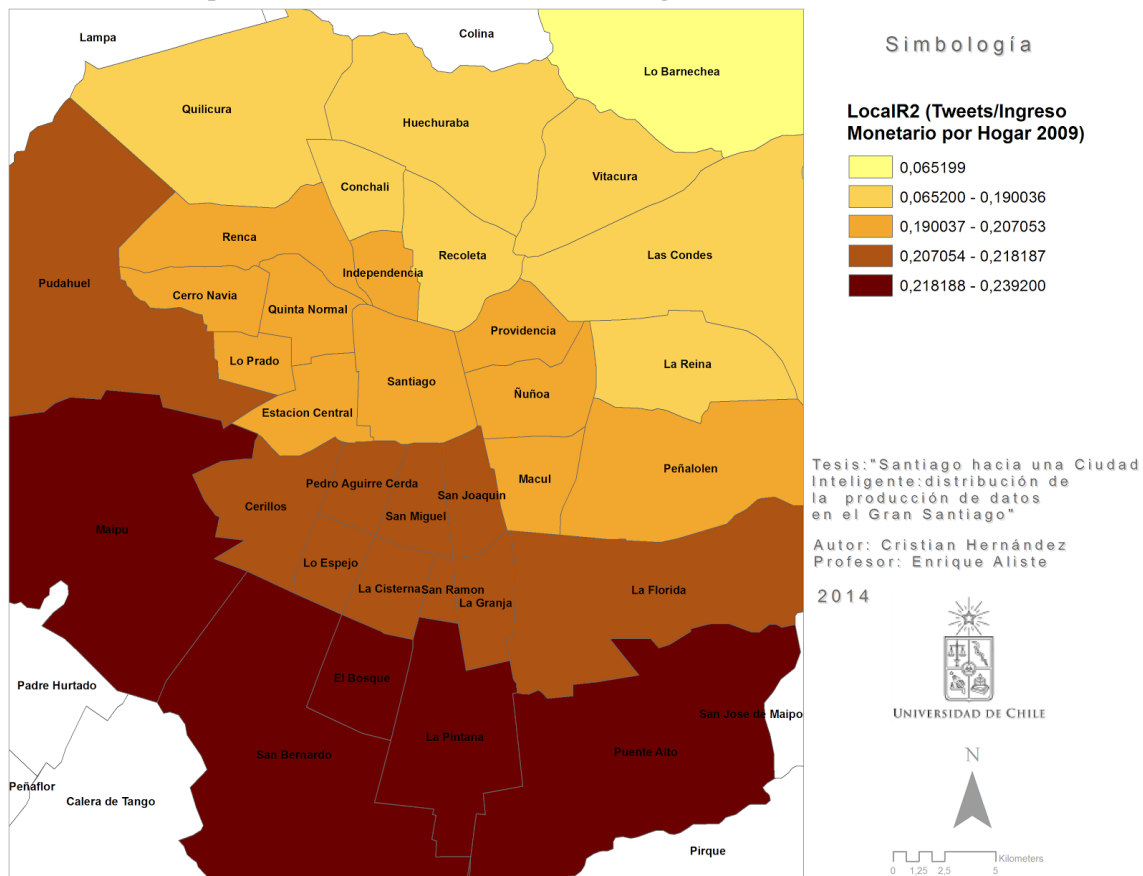
Al comparar la correlación entre el patrón territorial del ingreso y la producción activa de datos, vemos que existe una muy baja correlación, de tan solo “0,28”. Esto implica que los ingresos y la distribución que presenta explicarían de forma muy baja la distribución de la producción activa de

datos, teniendo que buscar en otras variables la causa del patrón que muestra la producción activa de datos.

Por su parte, la producción pasiva y total de datos, presentan una aún más baja correlación en relación al ingreso, alcanzando tan solo “0,21”. Es decir, que si bien el ingreso tiene una influencia en la distribución de la producción de datos, no es de mayor consideración. En especial en el caso de los datos producidos de forma pasiva, y totales.

A través del análisis de correlación entre el ingreso monetario promedio por hogar³⁹ y la producción de datos para las comunas del Gran Santiago, ahora mediante el análisis GWR, tenemos lo siguiente:

Mapa 16: “Análisis LocalR2: Tweets e Ingreso Monetario Promedio”



Fuente: Elaboración propia, 2014.

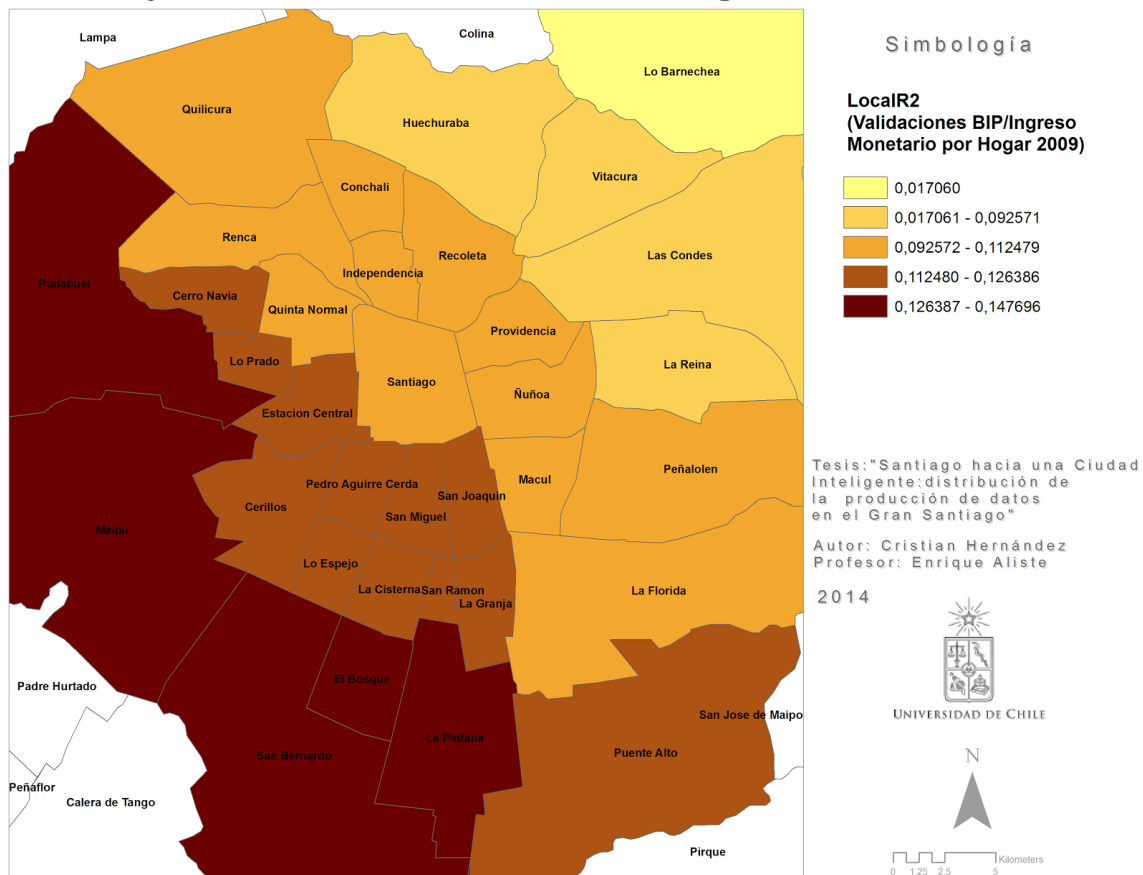
La correlación (LocalR2) entre el ingreso y la producción activa de datos, tal como se muestra para el caso del análisis por manzana, utilizando los datos de ingreso provistos por Adimark, tenemos que el ingreso explica de forma muy baja la producción de datos en el Gran Santiago a nivel comunal, con valores máximos de correlación LocalR2 de $\approx 0,24$ (0,239).

³⁹ Ingreso Monetario Promedio a nivel Comunal 2009, CASEN.

En promedio para las 34 comunas del Gran Santiago, el promedio de la correlación entre el ingreso monetario por hogar, y cantidad de datos, es de 0,2. En términos de la correlación ajustada, la cifra disminuye a 0,17.

El patrón espacial resultante de la cartografía, indica a través de dos mediciones de correlación entre el ingreso (Adimark 2003 y Casen 2009), la presencia de cinco anillos en tanto la producción de datos explicada por los ingresos. Donde existen las mayores correlaciones es hacia el Sur-Oeste, disminuyendo la relación del ingreso con la producción de datos, de forma continúa hacia el Nor-Oriente.

Mapa 17: “Análisis LocalR2: Validaciones BIP e Ingreso Monetario Promedio”

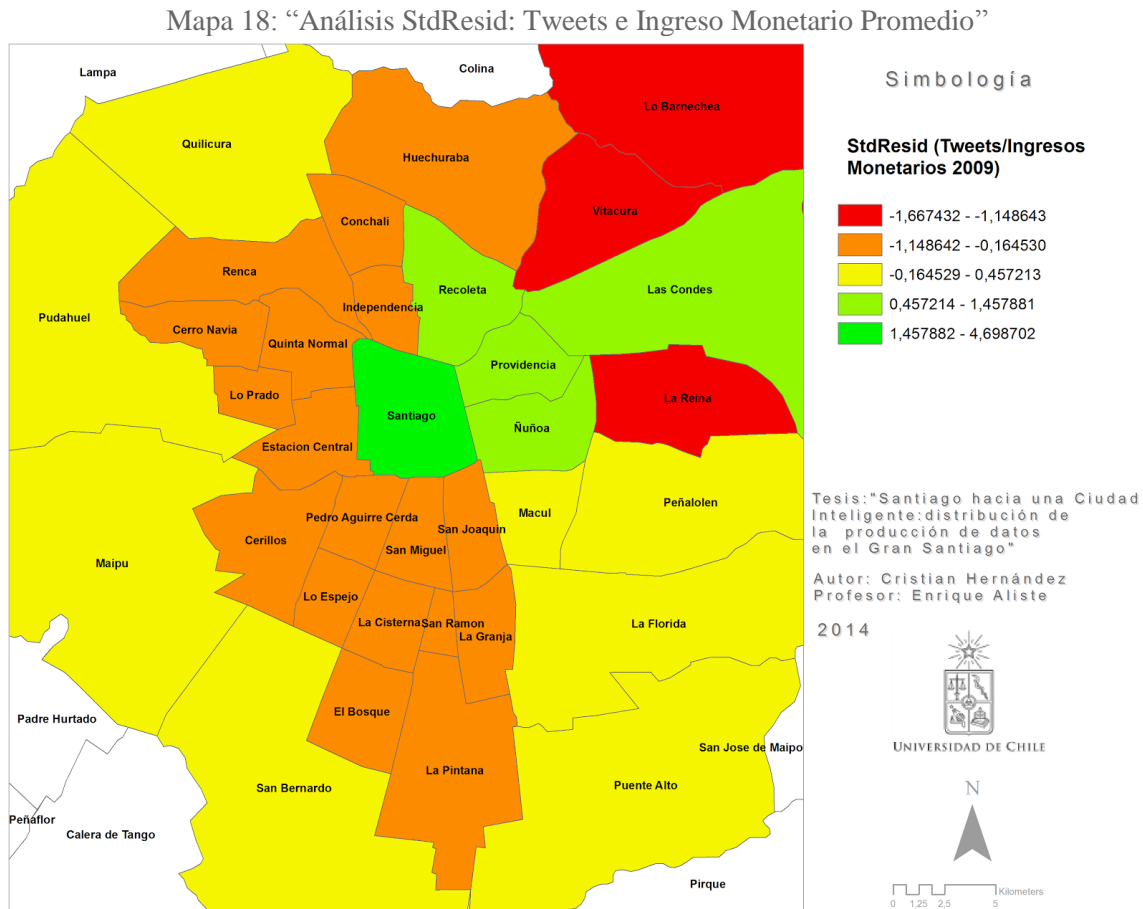


Fuente: elaboración propia, 2014.

Para el caso del análisis de la correlación entre el ingreso y la producción de datos pasivos (Validaciones BIP), a nivel comunal para las 34 comunas del Gran Santiago, tenemos que nuevamente el ingreso explica escasamente la producción de datos, alcanzando solamente la cifra de $\approx 0,15$ (0,1476) donde 0 explica nula correlación y 1 determina una correlación directa o proporcional entre las variables. Se identifica nuevamente el patrón de anillos Sur-Oeste, Nor-Oriente, que muestran como el ingreso explica gradualmente en menor medida la producción de datos pasivos, a medida que nos acercamos al Nor-Oriente. En términos del promedio de

correlación entre el ingreso y la producción de datos pasivos, tenemos que R^2 es igual a 0.14, mientras que el promedio ajustado corresponde a 0.08.

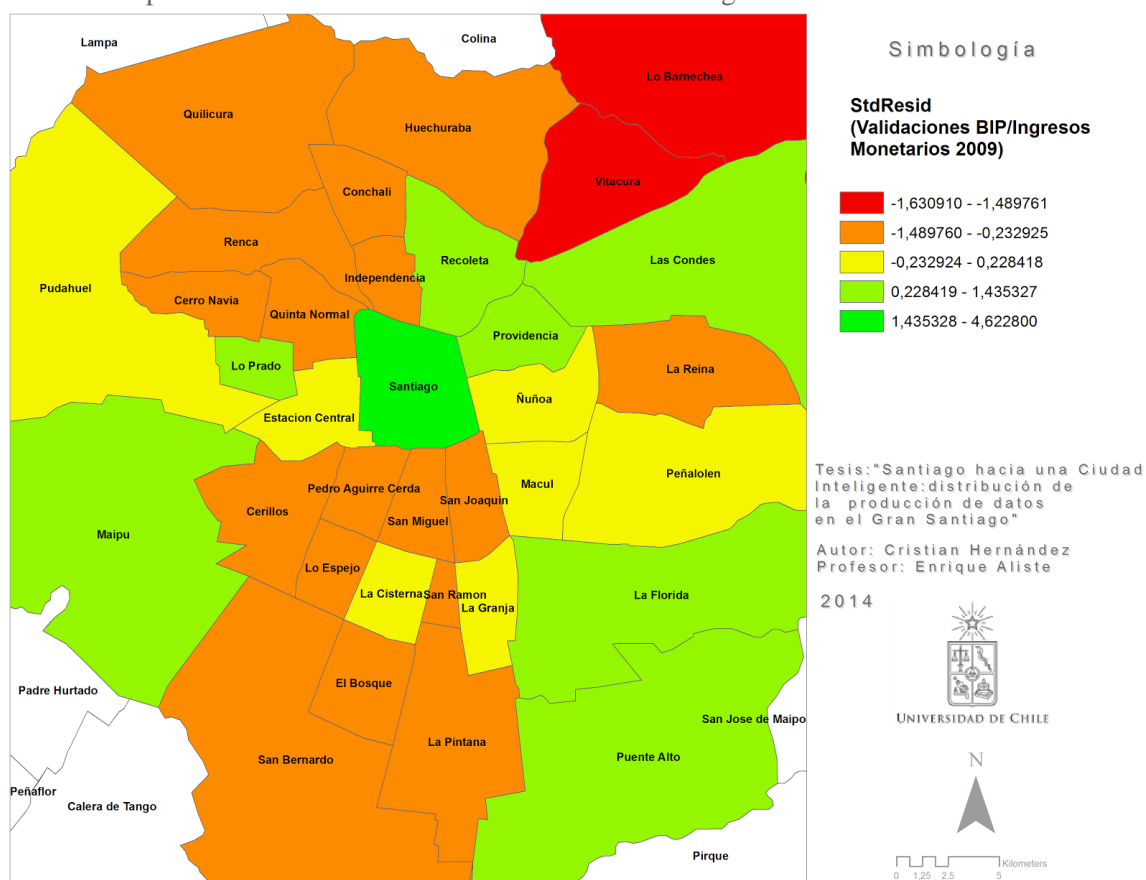
En términos del análisis de “valores residuales”, se observa que para la producción activa de datos comunas de mayor producción de datos desde *Twitter* como Santiago, Providencia, Ñuñoa, Las Condes, y en menor medida Recoleta, presentan una tendencia a presentar mayores valores residuales, es decir, que presenta una mayor cantidad de “variable inexplicada”, en relación al ingreso. Tal como se observa en el Mapa 18:



Fuente: Elaboración propia, 2014.

Lo mismo ocurre al comparar el ingreso con la producción pasiva de datos a través de validaciones BIP, donde un mayor número de comunas, principalmente las de mayor producción de datos pasivos, presenta los mayores valores residuales:

Mapa 19: “Análisis StdResid: Validaciones BIP e Ingreso Monetario Promedio”



Fuente: Elaboración propia, 2014

En términos resumidos para el análisis de la correlación del ingreso con la producción de datos, los valores que entrega Adimark (2003) indican que, si bien existe una correlación positiva entre dichas variables, es decir, que el ingreso explicaría la producción de datos, dicha correlación es escasa, con tendencia a ser nula. En términos comunales, y realizando el análisis de correlación entre el ingreso monetario promedio por hogar en las comunas del Gran Santiago, y la producción de datos activa y pasiva, se aprecia que la correlación sigue siendo positiva, pero con tendencia a ser nula. Finalmente, observando los valores residuales, que indica en términos de desviaciones estándar, cuánto de la variable permanece sin explicación, es decir, cuánto de la producción de datos a nivel comunal, permanece sin explicación en términos del ingreso, se observa que las comunas que representan la mayor producción activa, y aquellas que representan las mayores producciones de datos pasivos, son aquellas en las que se encuentran mayormente “sin explicación”, observándose los valores *StdResid*, más altos.

Todo esto indica, que se debe buscar nuevas variables que expliquen la distribución de la producción de datos en el Gran Santiago, en términos que el ingreso no explica el por qué los datos se producen en ciertas áreas de la Ciudad, en detrimento de otras.

Capítulo IV: Conclusiones

IV.a.- Sobre los patrones espaciales de la producción de datos y su relación con el ingreso.

Llama la atención que el 88% de la producción de datos activa (tweets) se realiza en los límites de la planta urbana consolidada, con una tendencia a la concentración en plantas urbanas dispersas en la RM. En cierta forma se comprueba lo postulado por Jiang & Miao (2014), en relación a los límites de la “Ciudad Natural”, en tanto la información proveniente desde *Twitter*, muestra el movimiento de los habitantes urbanos, y cómo viven la ciudad en el día a día, dibujando las zonas que son utilizadas de la ciudad, en relación a otras que no. Resulta relevante el uso de dicho aporte al dibujo de la ciudad, en investigaciones futuras, sobre la dinámica de las plantas urbanas de Santiago, no solamente del centro urbano consolidado del Gran Santiago. Comprender por ejemplo potenciales nuevas zonas de expansión de la ciudad.

En términos de la producción activa de datos la concentración genera tres concentraciones mayores entre Santiago, Providencia, Ñuñoa y Las Condes, con presencia de algunos centros dispersos en otras comunas, pero de menor tamaño. En este caso, se podría ligar dicho resultado obtenido desde *Twitter* con los estudios sobre los “policentros” en Santiago (Escolano & Ortiz, 2005), en tanto se producen concentraciones en comunas que no pertenecen al *Central Business District* de Santiago, como La Florida, Puente Alto, Maipú o Recoleta/Huechuraba.

Santiago representa una hiperconcentración en la producción de datos activos, en tanto por sí misma dicha comuna abarca el 21% de la producción. Si se suma Las Condes, Providencia y Ñuñoa, se alcanza el 49%. Estudios sobre movilidad urbana diaria, como los datos de población flotante podría dar luces sobre la alta concentración de producción de datos activos la comuna de Santiago. Por otro lado, cobra relevancia que la fuente de datos escogida para éste estudio pueda develar o servir como muestra para la corroboración de dichos fenómenos, es decir, que si analizáramos todos los días la información que *Twitter* nos entrega de *tweets* geolocalizados, podríamos tener una visión más acabada de la movilidad urbana diaria.

En términos de la producción pasiva de datos, provenientes desde Validaciones BIP, el Metro posee un peso muy importante en las comunas en las que está presente, es decir, en comunas en las cuales se localiza una o más estación de Metro. Sin embargo, los datos producidos desde las Validaciones en Buses son mayores en cantidad. En efecto, mientras en Metro para un día laboral promedio se realizan 2.158.898 de validaciones BIP, lo que representa el 39% del total de validaciones (5.471.380), en Buses se realizan 3.312.482 validaciones, lo que representa el 61% del total. Esto se relaciona con la mayor dispersión en la ciudad de las paradas de buses, en relación a las estaciones de Metro, que hace disminuir el peso que las paradas de buses poseen, y que se demuestra en los resultados finales, por ende el peso de una estación de Metro en tanto la producción de datos a través de Validaciones BIP, determina la forma de la concentración pasiva de datos, que suma las validaciones de Metro y Bus.

El gran peso de los datos producidos de forma pasiva en relación a los datos producidos de forma activa, determinan la forma de la concentración de los datos totales. Resulta importante recalcar dicho punto, ya que la mayor cantidad de datos producidos por la sensorización del territorio, proviene de un sensor pasivo dispuesto en el espacio, y no de la producción activa de los habitantes urbanos en Santiago, como en este caso, a través de *Smartphones*, sin restarle relevancia a la granularidad y a la diversidad de fenómenos que pueden ser explorados a través de datos producidos por teléfonos inteligentes.

En términos de la relación del ingreso y la producción de datos, se concluye que el ingreso no explica ni la producción de datos activa, ni la producción de datos pasiva. De hecho, en las mediciones realizadas con datos de ingreso provenientes de Adimark y los datos de ingreso provistos por Casen, ambos a nivel de hogar para un nivel espacial de manzanas, y hogares, respectivamente, para las 34 comunas del Gran Santiago, se observan coeficientes similares de correlación, que van desde el 0,28 (Adimark) hasta el 0,24 (Casen). Más aún, en términos de valores residuales, se observa que precisamente las comunas que presentan las mayores concentraciones en la producción de datos activos y pasivos, poseen también los valores más altos residuales, es decir, que se mantienen con una mayor tendencia a no ser explicados por el ingreso.

El presente estudio tuvo como resultado una hipótesis negativa, en tanto el ingreso posee una escasa relación con los niveles de producción de datos tanto activos como pasivos en el Gran Santiago. Es por ende necesario, en estudios posteriores, buscar nuevas variables que expliquen dicho patrón espacial de la producción de datos en el Gran Santiago. En conclusión se necesitan nuevas variables para explicar el por qué del patrón espacial de la producción de datos en el Gran Santiago.

IV.b.- Potencial en el Gran Santiago de favorecer la Sociedad Habilitada Espacialmente

En este punto se debe realizar un análisis distinto en relación a la utilización de datos activos y pasivos, en tanto que los datos activos presentan una serie de servicios privados que abren sus datos a través de diversas APIs, mientras que los datos pasivos, por lo general requieren de autorizaciones por parte de empresas privadas y públicas, que requieren vencer una burocracia mucho mayor para ser abiertos al público. A modo de ejemplo, podemos comparar las fuentes de datos usadas en el presente estudio. Mientras *Twitter* abre sus datos para el uso del público en general, Transantiago, a través de la empresa operadora del sistema de transporte, y el Ministerio de Transporte del Gobierno de Chile, sólo disponen para el uso libre la muestra de un día laboral promedio del año 2012.

En este marco, resulta interesante la oportunidad que presentan comunas como Santiago, Las Condes, Providencia y Ñuñoa en cuanto al aprovechamiento de datos activos, en tanto en dichas comunas se concentra la mayor producción. Sin duda una política de aprovechamiento de dichos datos, incentivando a la ciudadanía a utilizarlos, podría dar como resultado una iniciativa de

Ciudad Inteligente que favorezca que los vecinos de dichas comunas tiendan a ser “habilitados espaciales”, al utilizar la información libre geolocalizada de su comuna.

Por otro lado, y continuando con el análisis de la producción activa de datos, comunas como San Ramón, Lo Espejo, Pedro Aguirre Cerda o Cerro Navia, que no alcanzan a representar siquiera el 1% de la producción de datos activos en el Gran Santiago, deberían desarrollar estrategias que les permitan recolectar datos activos, como el *Crowdsourcing* o recolección de datos de forma colaborativa. Dicho proceso que permite apoyarse en un grupo numeroso de colaboradores, permite tanto recolectar datos (Willett et al, 2012), como analizar los datos (Welbourne, 2014), sin recurrir a costosos sistemas de sensores, ni a costosos *softwares* para el análisis.

Desde el punto de vista de los datos pasivos, la discusión queda abierta en relación a la apertura de datos del transporte, como Transantiago. Si bien el proyecto de Ciudad Inteligente elaborado por la Subsecretaría de Transporte menciona la apertura de datos, el plazo autoimpuesto para dicha apertura de datos, es lo que hace dudar del real impacto de dicha medida, en tanto se propone el año 2020 como meta. Otros datos que desde el punto de vista del autor debieran ser abiertos, son aquellos datos producidos en las autopistas urbanas de Santiago, teniendo en cuenta las cifras del parque automotriz de Santiago (1.721.328 de vehículos⁴⁰). Conocer el flujo de vehículos en las autopistas urbanas, podría permitir por ejemplo, crear una aplicación de la congestión en tiempo real de las autopistas, sin necesidad de estar ya en el atochamiento para saber que efectivamente dicho evento está ocurriendo. Sin duda ese espacio híbrido “público/privado”, genera distorsiones desde el punto de vista del derecho al acceso a los datos.

Ésta podría ser una de las conclusiones más importantes de este estudio, y que sirve de pie a futuras investigaciones, además de una crítica al escaso acceso a los datos del Transantiago, ya que gracias al análisis de un día de datos del transporte, a través de las validaciones BIP, es posible comprobar patrones de origen destino, al nivel de estudios como el de MIDEPLAN-SECTRA (2008). En teoría, si los datos estuvieran abiertos al acceso público, podríamos realizar cada día, un estudio “origen-destino” en el Gran Santiago, lo que le agregaría la componente de tiempo real al análisis de la Ciudad, en este caso Santiago.

En términos de Sociedad Habilitada Espacialmente a nivel comunal, poco se puede concluir en el caso de datos pasivos como los del Transantiago, que permanecen cerrados, sin embargo, si la información provista como datos públicos como un día laboral promedio para el año 2012, se mantiene, deberíamos decir que Santiago vuelve a ser la comuna que posee las mejores condiciones, en cuanto a la cantidad de datos que en dicha comuna se generan, para desarrollar iniciativas de “habilitamiento espacial” de sus vecinos.

Finalmente mencionar que la posición del autor tras esta investigación, es que la inteligencia de una ciudad no radica en la suma de artefactos (sensores) que introducimos en el territorio urbano. Al día de hoy, sin siquiera poseer datos ya organizados a nivel comunal podemos generar una comuna inteligente, solo con el uso de *APIs* de servicios como *Twitter*, *Foursquare*, *Facebook* o

⁴⁰ Anuario Parque de Vehículos en Circulación, 2013.

Instagram (por nombrar algunos). Sumando datos estáticos posteriormente (como Censos de Población o encuestas propias), podríamos crear proyectos que incentiven una ciudadanía empoderada a través del acceso a los datos, a mucho menor costo, y sin tener que esperar hasta el 2020 como es el objetivo de la Subsecretaría de Transporte.

Desde el punto de vista de la producción de datos que se realiza en cada comuna, existe una clara vocación potencial de comunas como Santiago, Las Condes, Providencia, Puente Alto y Maipú, a utilizar la gran cantidad de datos producidos en sus comunas, con el fin de mejorar sus políticas de planificación y diseño urbano. ARUP (2011), recopila una serie de ejemplos sobre cómo se utilizan los datos producidos de forma activa y pasiva (o una mezcla de ambos), para la mejora del proceso de planificación y diseño de la Ciudad, tales como mejorar las condiciones de tránsito, conocer qué piensan las personas acerca de ciertos espacios (como plazas o esquinas “conflictivas”), realizar mapas colaborativos, o conocer la dinámica de micro espacios a través de la sensorización del piso, para conocer la dirección de los pasos de los transeúntes en un espacio público, dejando ver que la sensorización puede ser posible para las comunas, sin necesidad de generar grandes y grandilocuentes proyectos sobre Ciudad Inteligente.

Bibliografía

1. Aliste, E. "Territorio y ciencias sociales: trayectorias espaciales y ambientales en debate", 2010.
2. Almonacid. "Estudio de Factibilidad para la Implementar Tecnología RFID en Biblioteca Miraflores", 2007.
3. ARUP. "Designing with data: Shaping our future cities", 2013.
4. Batty. "Big data, smart cities and city planning", 2013.
5. Batty. "Smart Cities of the Future. Working Paper 188", 2012.
6. Batty, "Building a Science of Cities. Paper 170", 2011.
7. Castells, Manuel. "The impact of Internet on Society: a global perspective", 2014.
8. Castells, Manuel. "La Era de la Información (Volumen 1): Economía, Sociedad y Cultura. La Sociedad Red", 2002.
9. Charlton, Fotheringham, Brunston. "Geographically Weighted Regression", 2006.
10. CISCO (Dave Evans). "The Internet of Things How the Next Evolution of the Internet Is Changing Everything", 2011.
11. Cohen & Obediente. "Ranking de Ciudades Inteligentes en Chile", 2014.
12. Comisión Europea (Nota de Prensa). "Commission launches innovation partnership for Smart Cities and Communities", 2012.
13. CSC. "Data Revolution", 2011.
14. De Stefani. "Reflexiones sobre los conceptos de espacio y lugar en la arquitectura del siglo XX", 2009.
15. Elgueta, Ortiz. "Gran Santiago: Transformación del Patrón Espacial de Localización del Sector Residencial de la Elite", 2006.
16. Escolano, Ortiz. "La formación de un modelo policéntrico de la actividad comercial en el Gran Santiago. Revista de Geografía Norte Grande", 2005.
17. Fischer, Daniel. "Obtención de información para la gestión del Transantiago a través datos operacionales", 2010.
18. Florida, Richard. "Cities and the Creative Class", 2003.
19. Giffinger. "Smart cities: Ranking of European medium-sized cities", 2007.
20. Greco & Bancardino. "The Paradigm of the Modern City: SMART and SENSEable Cities for Smart, Inclusive and Sustainable Growth", 2012.
21. Greenfield, Adam. "Against The Smart City", 2013.
22. Gruen, Armin. "SMART Cities: The need for spatial intelligence", 2013.
23. Henri Lefebvre. "El Derecho a la Ciudad", 1968.
24. Hernández, Cristian. "Estado de la Brecha Digital en el Gran Santiago", 2006.
25. Hewlett-Packard. "Living Progress Report", 2013.
26. IBM (Dirks & Keeling). "A vision of smarter cities", 2009.
27. IDC-EMC. "The Diverse and Exploding Digital Universe", 2008.
28. IESE. "Cities in Motion", 2014.
29. ITU. "Facts and Figures", 2014.

30. Jiang & Miao. "The Evolution of Natural Cities from the Perspective of Location-Based Social Media", 2014.
31. Komminos. "Developing a Policy Roadmap for Smart Cities and the Future Internet", 2011.
32. Kung et al. "Exploring Universal Patterns in Human Home-Work Commuting from Mobile Phone Data", 2014.
33. Le Corbusier. "Hacia una Nueva Arquitectura", 1923.
34. McEvoy Sharlene A. "Brave New World: The Economic Impact of the Driverless Car: A Bumpy Road Ahead?", 2014.
35. Mideplan-Sectra. "Informe Final estudio Actualización de Encuestas Origen Destino de Viajes, III Etapa", 2008.
36. Miranda, Daniel. "Técnicas de Imputación para viajes con información incompleta a partir de datos transaccionales de Transantiago", 2011.
37. ONU (Abbas Rajabifard). "Realizing Spatially Enabled Societies – A Global Perspective in Response to Millennium Development Goals", 2009.
38. ONU. "Reporte Especial de la Promoción y Protección de la Libertad de Opinión y Expresión", 2011.
39. Pallot. "Future Internet and Living Lab Research Domain Landscapes: Filling the Gap between Technology Push and Application Pull in the Context of Smart Cities", 2011.
40. Rajabifard. "Spatially Enabling Government, Industry and Citizens", 2012.
41. Roche. "Are 'Smart Cities' Smart Enough?", 2012.
42. Rodriguez, Jorge. "Segregación residencial socioeconómica: ¿qué es?, ¿cómo se mide?, ¿qué está pasando?, ¿importa?", 2001.
43. Stefanidis et al. "Social Media and The Emergence of Open Source Geospatial Intelligence", 2011.
44. Steudler, Daniel & Rajabifard Abbas, "Spatially Enabled Society", 2012
45. Subsecretaría de Telecomunicaciones, Gobierno de Chile. "Ciudad Inteligente y con sentido Ciudadano", 2014.
46. Subsecretaría de Transporte, Gobierno de Chile. "Estrategia Ciudad Inteligente para el Transporte", 2014.
47. Townsend, Anthony. "Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia", 2013.
48. Welbourne et al. "Crowdsourced Mobile Data Collection: Lessons Learned from a New Study Methodology", 2014.
49. Willett et al. "Strategies for Crowdsourcing Social Data Analysis Wesley", 2012.
50. Williams et al. "Measures of Human Mobility using mobile phone records enhanced with GIS data", 2014.
51. Williamson. "Spatially Enabled Society", 2011.

Anexos

A.- Mapas utilizados en esta investigación.

Para descargar todos los mapas utilizados en esta investigación, dirigirse a la siguiente dirección en Internet:

<https://drive.google.com/folderview?id=0B8gcJYvjkHJqMXhiQ3FfdXJpLWM&usp=sharing>

B.- Código Javascript.

Para descargar el código Javascript utilizado en esta investigación, dirigirse a la siguiente dirección en Internet:

<https://drive.google.com/file/d/0B8gcJYvjkHJqSkdRclhmaG9mSk0/view?usp=sharing>

C.- Tabla Total de Coordenadas Twitter y Validaciones BIP.

Para descargar la totalidad de puntos correspondientes a *Tweets* y Validaciones BIP utilizados en esta investigación, dirigirse a la siguiente dirección en Internet:

<https://drive.google.com/file/d/0B8gcJYvjkHJqei1YLUplUFAXUDQ/view?usp=sharing>