

الجهاز المركزي للإحصاء الفلسطيني
ورقة عمل مقدمة في المؤتمر الدولي العلمي الخامس لاتحاد الإحصائيين العرب
القاهرة، 9-10/2/2016

التحديات التي تواجه الأجهزة الإحصائية في ظل ثورة البيانات الكبيرة: دراسة حالة، فلسطين

إعداد: مروان بركات/ نائب مدير عام-الإدارة العامة للإحصاءات الجغرافية
الجهاز المركزي للإحصاء الفلسطيني
mbarakat@pcbs.gov.ps

الملخص

تعتبر البيانات الكبيرة Big data مُنتج من العصر الرقمي والإنفجار المعلوماتي الذي يتسم به عالمنا الحالي، وتتميز البيانات الكبيرة -بصفتها ظاهرة ثقافية وتكنولوجية وعلمية- بحجمها الهائل، وسرعة تكونها العالية وتنوع مصادرها وصيغها. إن هذا النوع من البيانات هو ثروة كامنة للمؤسسات الإحصائية الرسمية، يتطلب الاستفادة منها توفير أدوات وإمكانيات وطرق تفكير غير تقليدية.

تستعرض الورقة مفهوم البيانات الكبيرة، ووتيرة نموها، وكذلك تطرقت لفحص أهمية الاستفادة منها في سياق عمل المؤسسات الإحصائية الرسمية، وقد تم تطبيق دراسة حالة لنظام إدارة المعلومات المكانية في فلسطين (GeoMOLG).

تناول القسم المتعلق بالتحديات مجموعة من الجوانب التشريعية والمالية والإدارية والمنهجية والتكنولوجية ذات العلاقة.

وقد اشتملت الورقة على تحليل لوضع الجهاز المركزي للإحصاء الفلسطيني والبيئة الناظمة لعمله، مع التركيز على مجالات الاستفادة ومصادر البيانات والمتطلبات اللازمة لتنفيذ مشاريع البيانات الكبيرة.

تعتبر نتائج الورقة خطوة على طريق إستكمال مجموعة من البحوث التفصيلية تمهيدا لإتخاذ قرارات على المستوى الهيكلي للأجهزة الإحصائية المهمة بإستخدام البيانات الكبيرة، وما سيرفدها من بيئة قانونية داعمة، بهدف تحفيز الاستثمار الأمثل للبيانات الكبيرة من حيث تحويلها الى معلومات مفيدة في وقت قياسي، لخدمة المواطنين والباحثين وصانعي القرار.

Challenges facing statistical agencies under the big data revolution: Case Study, Palestine

Marwan Barakat/Deputy Director General –Area Statistics Directorate
Palestinian Central Bureau of Statistics PCBS-Palestine
mbarakat@pcbs.gov.ps

Abstract

As an educational, technological and cultural phenomena, big data considers as a freshly product of informational revolution due to its huge size, the high speed of formation and variety of its sources and formats. This type of data is a latent wealth of official statistical institutes in case of adaptation, but to do this, there is an urgent need to for specific tools, possibilities and non traditional ways of thinking.

This paper reviews the concept of the big data, examining the importance of benefiting of such data.

The paper also examine the importance of taking advantage of big data in the context of the work of NSO's.

It has been applied to the case study of spatial/geographical information management system in Palestine (GeoMOLG).

The challenges section dealt with a set of associated legislative, financial and administrative, methodological and technological aspects

The paper included an analysis of the current status of the Palestinian Central Bureau of Statistics and the environment governing its work, with a special focus on utilization of data sources and requisites needed to implement projects of large data

The paper results are considered as a step to continue with other detailed researches in preparation for structural level decisions for NSO's to provide a supportive legal environment to encourage the optimal investment in big data through transforming it to a valuable information in optimal time to serving researchers, decision makers and other users.

1. مقدمة

قبل حوالي ألفين وثلثمائة عام وقف بطليموس الثاني حائراً مهموماً متسائلاً عن الكيفية التي يمكن أن تتيح لجمهور المستفيدين الوصول الى ضالته المنشودة في الوقت المطلوب، أمام جبل المعلومات المُكدسة الذي قوامه تسعمائة ألف مخطوطة من ورق البردي تحويها مكتبة الاسكندرية التي كانت تمثل مركز معارف العالم في ذلك العصر، قام بطليموس الثاني بعرض المعضلة على العالم متعدد المواهب: (زينودتس)، الذي توصل بعد تفكير طويل وعمل مضنٍ الى الحل من خلال فهرسة جميع مقتنيات المكتبة، وبالتالي تسهيل الوصول الى الكتاب الملائم في الوقت الملائم¹.

في القرن السادس عشر إبتكر الفلكي الدانماركي (تايخو براهي) أجهزة كانت هي الأكثر تطوراً في عصره، إستخدمها لرصد حركة الكواكب، وكدس من خلال هذا الرصد -على مدار معظم سنوات عمره- كميات ضخمة من البيانات توثق حركة الكواكب²، وبعد موته المفاجئ قام تلميذه (يوهانس كبلر) بمعالجة بيانات تايخو، مما سمح له بوضع القوانين الدقيقة التي تحكم حركة الكواكب وإكتشاف طبيعة نموذج النظام الشمسي³.

في ثمانينات القرن التاسع عشر قدر مكتب التعداد الأمريكي زمن معالجة بيانات التعداد التالي (1890) بأكثر من عشر سنوات، وذلك بسبب تضخم عدد السكان والتفاصيل الإضافية التي تحويها استمارة التعداد. لحل هذه المعضلة إبتكر (هوليريث) الآلة التي عُرفت بإسمه، وأستمرت في العمل بعد ذلك لحوالي مائة عام: آلة الجدولة الالكتروميكانيكية التي صممها خصيصاً للمساعدة في تسريع معالجة البيانات، وكانت النتيجة هي الانتهاء من تجهيز البيانات خلال سنتين⁴.

في تموز/يوليو من العام 1997 ظهر مفهوم البيانات الكبيرة للمرة الأولى في مقالة أعدها باحثين من وكالة "ناسا" ورد فيها أن التمثيل البصري يثير تحديات لأنظمة الحاسوب بسبب ضخامة مجموعات البيانات التي لا تتسع لها الذاكرة الرئيسية، ولا القرص الصلب، ووصفت المقالة هذه الظاهرة بـ "مشكلة البيانات الكبيرة"⁵.

لقد تطور المفهوم الحديث للبيانات الكبيرة بعد ظهوره الأول في العام 1997؛ ليحمل دلالات مُركبة تصف هذه الظاهرة الثقافية-التكنولوجية، التي يمكن تعريفها على النحو: "بيانات مُكدسة، لها من الضخامة والتعقيد ما يحول دون القدرة على معالجتها بإستخدام أدوات⁶ برمجية تقليدية."

من المفيد أن نلاحظ الارتباط بين المفهوم الحديث للبيانات الكبيرة والسياق التاريخي لها:
- جبل المعلومات المُكدسة العvisية على المعالجة في مكتبة الاسكندرية القديمة والتي تم الوصول الى حل مُبتكر لمشكلتها من خلال الفهرسة
- البيانات الضخمة التي كدسها (براهي) فقام (كبلر) بمعالجتها وتحويلها الى معرفة نافعة على شكل قوانين ونموذج للنظام الشمسي
- بيانات التعداد الأمريكي الضخمة التي إبتكر (هوليريث) حلاً خلاقاً للتسريع في عملية معالجتها
- البيانات المُكدسة الضخمة المعقدة في عصرنا الحديث والتي تحتاج إلى وسائل غير تقليدية لاستخراج معلومات قيّمة منها في زمن قصير

يبقى التأكيد على أن ما ينطبق على البيانات العادية ينطبق على البيانات الكبيرة من حيث أنها الوسيلة وليست المبتغى، إنها أداة تكون ناجحة فقط في حال القدرة على استخلاص معلومات ومعرفة فعالة منها؛ لإتخاذ القرارات السليمة بناءً عليها، وبغير ذلك ستبقى مجرد أكوام من الأعداد لا يفهمها؛ وبالتالي لا يستفيد منها أحد.

2. خصائص البيانات الكبيرة

للبيانات الكبيرة العديد من الخصائص/الأبعاد، أُختصرت في معظم المراجع بكلمات تبدأ بحرف V باللغة الانجليزية، هي الحجم Volume والسرعة Velocity والتنوع Variety*

1.2 الحجم Volume

"تتمدد البيانات لتملأ المساحة المتاحة للتخزين"، إنطبق هذا القول المأثور على عالمنا الرقمي خلال السنوات الماضية.

أغرق العالم بطوفان من البيانات نتجت بشكل أساسي عن تسارع وتيرة الرقمنة (digitization) والانتقال من عصر تشبيك الحواسيب إلى عصر تشبيك الأشياء (Internet of Things IoT)† والتوليد الذاتي للبيانات من قبل الآلات -بعد برمجتها- دون تدخل البشر.

تقوم مؤسسة البيانات الدولية IDC بإصدار دراسات منذ العام 2007 حول كمية المعلومات الالكترونية التي تم إنشاؤها أو التقاطها أو تكرارها في جميع أنحاء العالم خلال كل سنة‡، وتطلق المؤسسة على هذه المعلومات إسم "العالم الرقمي" Digital Universe. يبين جدول 1 النمو في حجم هذه المعلومات خلال السنوات 2005-2020:

جدول 1: تطور نمو العالم الرقمي 2005-2020

السنة	حجم العالم الرقمي (زيتا بايت [§])
2005	0.130
2006	0.180
2007	0.281
*2008	0.470
2009	0.800
2010	1.227
*2011	1.800
2012	2.800
2013	4.400
*2014	5.900
*2015	7.910
*2020	44.000

* بيانات تقديرية.

[°] إقتصرت الكثير من المراجع على ذكر أول ثلاث كلمات باعتبارها سمات مميزة للبيانات الكبيرة، إلا أن مراجع أخرى أضافت خصائص إضافية مثل الدقة Veracity.

[†] يشير هذا المفهوم الى ربط طيف واسع من المعدات والحيوانات والبشر مع شبكة الإنترنت، وتواصلها هذه الأشياء معاً، ويؤرخ لولادة هذا العصر في لحظة وقعت بين العامين 2008-2009، وذلك عندما تخطى عدد "الأشياء" المربوطة بالإنترنت عدد سكان العالم.

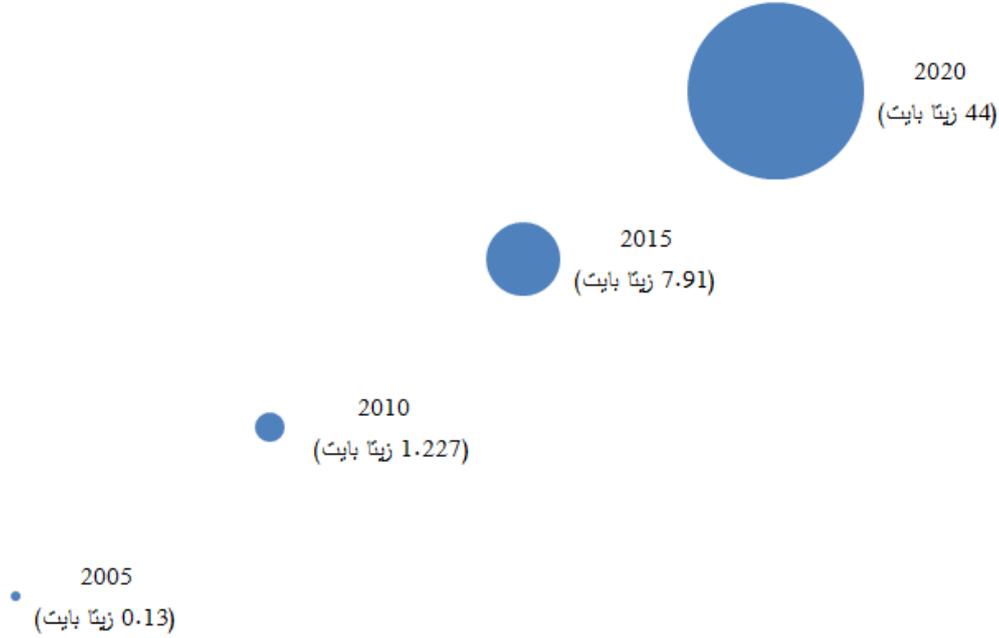
[‡] يشمل ذلك حساب أو تقدير المعلومات الرقمية سواء كانت مخزنة أو غير مخزنة.

[§] 1 زيتا بايت = ألف إكسابايت = مليون بيتابايت = مليار تيرابايت = تريليون جيجابايت = كوادريليون ميجابايت = كوونينيليون بايت.

ملاحظة: تم استخدام النظام العشري في البيانات الوارد ذكرها في هذه الدراسة وليس النظام الثنائي، وذلك بما يتوافق مع المنهجية المستخدمة من قبل مؤسسة IDC وجامعة كاليفورنيا-بيركلي، وشركة إنتل.

يمكن تجسيد هذا النمو بصرياً من خلال الشكل 1 التالي:

شكل 1: تطور نمو حجم العالم الرقمي خلال الاعوام 2005-2020:



عند ربط حجم العالم الرقمي بعدد السكان، يمكن حساب نصيب الفرد في العالم من المعلومات الرقمية: 1.08 تيرابايت* /فرد في العام 2015، للمقارنة: أنتج الجنس البشري ما مجموعه 0.005 زيبايت من المعلومات، وذلك منذ بزوغ فجر الحضارة حتى نهاية العام 2003.⁷

يستدعي هذا الانفجار المعلوماتي من حيث الحجم، الحاجة الى استخدام مقاربات وأدوات برمجية غير تقليدية، بهدف امتلاك القدرة على المعالجة والتحليل للبيانات، وبالتالي الحصول على إستخلاصات مفيدة وذات قيمة.

2.2 السرعة Velocity

هي المعدل الزمني الذي يتم فيه إنشاء وتخزين وتحليل البيانات وتمثيلها بصرياً⁸. فيما يلي إحصاءات متنوعة تعطي فكرة عن السرعة التي تتولد فيها البيانات:

- في كل يوم تُنتج أجهزة الهواتف الخليوية بمفردها بيانات حجمها 5000 تيرابايت⁹.
- في كل يوم يجمع موقع فيسبوك 500 تيرابايت من البيانات تحمل 2.5 مليار محتوى معلوماتي و 2.7 مليار نقرة إعجاب و 300 مليون صورة، وقد بلغ حجم الصور ولقطات الفيديو المخزنة لدى الموقع في العام 2012 أكثر من 100 ألف تيرابايت¹⁰.
- تتعامل شركة AT&T مع مستودعات هائلة من المعلومات ولدتها الشبكات والزبائن، وفي اليوم الواحد يتم توثيق مليارات الحركات المكانية-الزمانية والسجلات بمختلف أنواعها¹¹ في قاعدة بيانات تُعتبر من الأضخم في العالم، بحجم 323 تيرابايت، حيث ينضوي ضمن قاعدة البيانات هذه مجموعة من الجداول يحوي أكبرها 938 مليار سجل¹².

* يتيح التيرابايت الواحد تخزين مليون رواية حجم كل منها 100 صفحة، أو مليون صورة حجم كل منها 1 ميجابايت، أو 300 ساعة تسجيل فيديو ذو جودة عادية، أو 17 ألف ساعة من التسجيل الصوتي، أو 1,538 قرص مدمج بسعة 650 ميجابايت.

معظم البيانات تُنتج في الوقت الحقيقي أو في وقت قريب من الوقت الحقيقي، بمعنى أن الظاهرة يتم توثيقها أثناء حدوثها تقريباً.

إن التدفق السريع للبيانات يفوق في بعض الأحيان قدرة أجهزة المؤسسة على التخزين، مما استدعى الحاجة إلى ظهور تقنية جديدة يتم فيها تحليل البيانات قبل تخزينها، وهو ما يطلق عليه التحليل داخل الذاكرة
in-memory analytics^{13 14}.

يضاف لما سبق أن طبيعة الأعمال في كثير من المؤسسات تستدعي إستخلاص المعلومات المفيدة من البيانات في الوقت الحقيقي real-time أو في وقت قريب من الوقت الحقيقي لحدوث الظاهرة¹⁵.

3.2 التنوع Variety

تتضمن البيئة التقليدية لإدخال ومعالجة البيانات التعامل مع بنية structure محددة لها، مما يُسهل إدارتها ضمن قاعدة بيانات علائقية Relational Database Management.

أما البيانات الكبيرة فهي تتنوع من حيث احتوائها على بيانات تقليدية ذات بنية محددة Structured وبيانات لا بنيوية unstructured وبيانات شبه بنيوية semi structured¹⁶. إضافة إلى تنوع مصادر هذه البيانات.

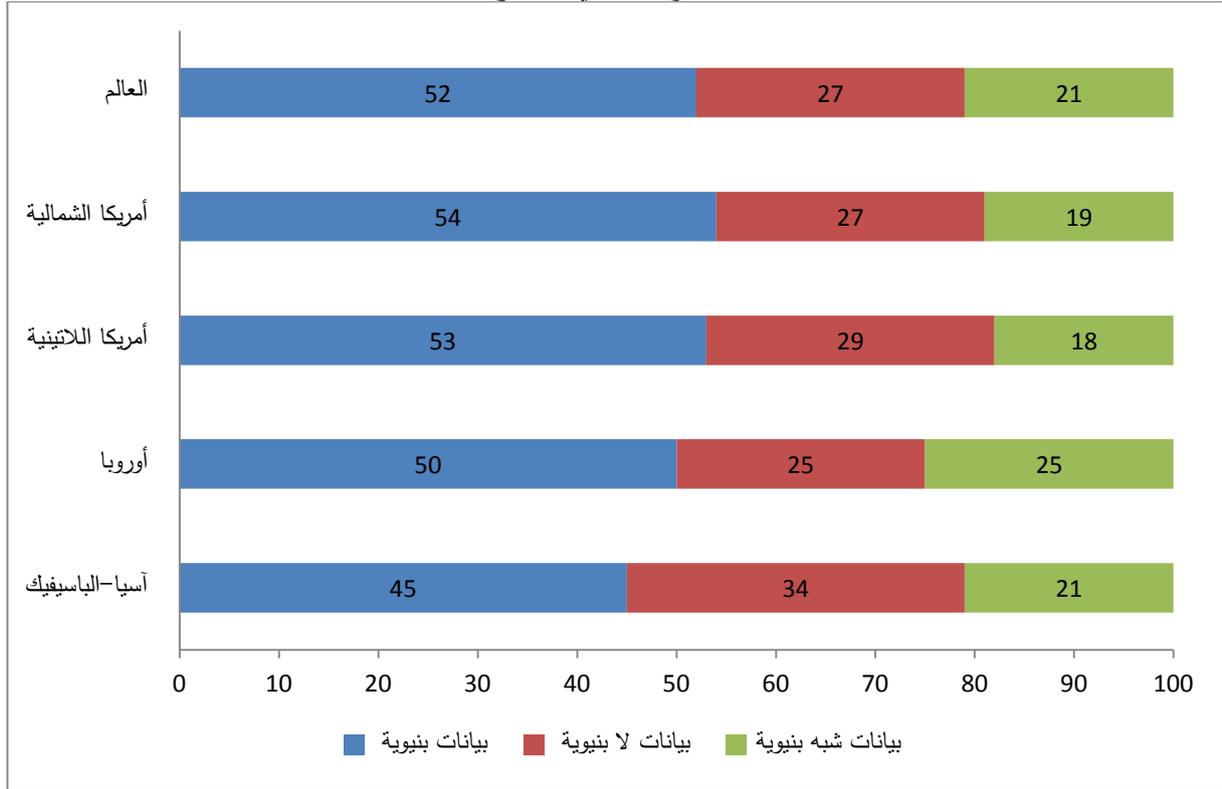
فالبيانات البنوية Structured Data هي تلك المنظمة والمعرفة في هيكلية محددة باستخدام قاعدة بيانات علائقية أو جداول ممتدة ومحددة في حقول ثابتة وأعمدة وسطور^{17 18}، وقد طغى استخدامها في عصر ما قبل البيانات الكبيرة، وما زال هذا الاستخدام قائماً. ينطبق هذا التعريف على بيانات المؤسسات الإحصائية من مسوح وتعدادات وسجلات إدارية.

يشير مصطلح البيانات اللابنيوية Unstructured data إلى البيانات التي قلما تمتلك هيكلية محددة مسبقاً مثل الصور والتسجيلات الصوتية والمرئية ورسائل البريد الإلكتروني والوثائق والنصوص الحرة¹⁹ وغالباً ما يقوم المستخدم بتوليدها. من المتوقع أن تزيد نسبة هذا النوع من البيانات في السنوات القادمة، مقارنة مع النوع الأول.

أما البيانات شبه البنوية Semi structured data فهي بيانات مماثلة للابنيوية بإستثناء احتوائها على معلومات توصيفية تساعد على تحليلها مثل البطاقات التعريفية Tags²⁰.

لقد تفوقت نسبة البيانات اللابنيوية في العالم على مجموع النوعين الآخرين، كما هو موضح في شكل 2.

شكل 2: التوزيع النسبي لأنواع البيانات*



مما سبق يتضح بأن البيانات الكبيرة قد تتواجد بشكل منظم سهل التفسير حسب قاموس بيانات محدد وقد تتواجد بشكل شبه منظم بسبب تنوع البيانات وقد تكون غير منظمة نهائياً مما يصعب فهمه والتعامل معه. بناء على هذه الصفات فإن البيانات الكبيرة تحتاج الى وسائل وأدوات جديدة ومختلفة وغير تقليدية لجمعها وتنظيفها وإدارتها ومعالجتها وتحليلها بطريقة فعالة.

يثار نتيجة للخصائص السابقة سؤال حول متى تُصنف البيانات على أنها كبيرة/ضخمة؟

الإجابة هنا نسبية وفقاً لعوامل عديدة منها البنية التحتية التي تمتلكها المؤسسة من معدات وبرمجيات، مع ذلك فإن البيانات تعتبر كبيرة/ضخمة إذا فاقت سعتها التخزينية إمكانيات جهاز الحاسوب المنفرد، بحيث يكون هناك حاجة لتوزيعها على عدة أجهزة، وإذا كان هناك حاجة لإستخدام برمجيات متطورة/غير تقليدية مثل Hadoop لإدارتها بكفاءة وفعالية.

عند استقصاء الوضع في فلسطين وربطه مع التعريفات السابقة، يتضح بأن هناك مؤسسات عديدة لديها بيانات كبيرة –بعض النظر عن استخدامها- منها: وزارة الحكم المحلي، شركات الهواتف النقالة، شركة الكهرباء، مصلحة المياه، وزارة النقل والمواصلات، وزارة الصحة والمستشفيات المركزية.

3. أهداف الاستخدام والفائدة والاهمية

لا تزال تداعيات ظاهرة البيانات الكبيرة في بداياتها، من حيث الاستفادة منها بالطرق الامثل، حيث تشير الدراسات إلى أن نصف بالمائة فقط من حجم العالم الرقمي في العام 2012 قد وجد طريقه للتحليل بأي

* التقريب الى مئة من عمل الباحث.

صورة من الصور، في حين أن ربع هذا العالم الرقمي فيه إمكانيات كامنة لإستخراج معلومات واستخلاصات مفيدة منه²¹.

إن الهدف الاساسي من استخدام البيانات الكبيرة هو إتاحة قدرة عالية على تحليل [وربط] طيف واسع ومتنوع من البيانات من مختلف المصادر²² بسرعة عالية مما يساعد على إتخاذ قرارات سليمة في الوقت المناسب، وتقليل التكلفة، وتحسين المنتجات والخدمات²³، وزيادة العائدات، بما يعنيه كل ذلك من زيادة في كفاءة عمل المؤسسة وتوفر أفضلية تنافسية لديها.

بتطبيق ما سبق في السياق الإحصائي فإن الهدف من استخدام البيانات الكبيرة يتمثل في كونها مصدر كبير لبيانات الاحصاءات الرسمية وغير الرسمية يمكن الاستفادة منها لما تتمتع به من سرعة إنتاج عالية (لحظية) وتنوعها وتوفرها بحجم كبير والتي يمكن استثمارها لإنتاج احصائيات فعالة تصف الواقع بصورة لحظية وسريعة تساعد صانعي القرار على التعديل الفوري والتطوير السريع للعمليات والفعاليات الحياتية المرتبطة بتلك البيانات الكبيرة.

يمكن تلخيص فوائد الاستخدام بإنتاج إحصاءات أفضل من حيث الوقتية والصلة بالواقع، بكفاءة سعرية عالية، بالإضافة الى تخفيف العبء على المبحوثين²⁴.

على صعيد المؤسسات الإحصائية فإن أهمية البيانات الكبيرة تعود الى كونها مصدر كبير وجديد للإحصاءات الرسمية وغير الرسمية يمكن الاستفادة منها بالإضافة للمصادر التقليدية من سجلات ادارية (بعضها بيانات كبيرة) وبحث ميداني بالإضافة الى أن البيانات الكبيرة يمكن أن توفر احصائيات جديدة لم يكن بالإمكان الحصول عليها بالمصادر التقليدية.

يشمل استخدام البيانات الكبيرة في المؤسسات الإحصائية قائمة متنوعة من التطبيقات منها:
- إطار المعاينة أو إنشاء سجل: تعريف وحدات المجتمع أو توفير معلومات مساعدة مثل المتغيرات الطبقيّة

- تعويض كلي للبيانات: استبدال بيانات بالكامل بديلا عن مسح معين
- تعويض جزئي للبيانات لمجموعة من المجتمع: تقليص حجم عينة المسح
- تعويض جزئي لبعض المتغيرات: تقليص طول استمارة المسح أو إغناء المسح ببيانات أكثر دقة
- تعويض متغيرات مفقودة
- المعالجة: المساعدة في إكتشاف وعلاج الاخطاء
- الربط مع بيانات أخرى: إغناء مجموعات البيانات بمتغيرات أخرى
- توكيد البيانات: فحص دقة وإتساق بيانات المسح
- توليد رؤى تحليلية جديدة: تحسين قياس ووصف ظواهر اقتصادية واجتماعية
- تحسين الكفاءة التشغيلية والفعالية²⁵

4. فحص جدوى مشاريع البيانات الكبيرة

إن فحص جدوى استخدام البيانات الكبيرة في مشاريع معينة يحتل أهمية حيوية قبل إتخاذ قرار نهائي بالتنفيذ، خصوصا في حال توفر بيانات تقليدية.

من الضروري أن يستند تحليل الجدوى الى احتياجات المؤسسة وطبيعة أعمالها والمُحددات الداخلية والخارجية المتعلقة بهذه الأعمال.

فيما يتعلق بالمؤسسات الإحصائية فإن قرار استخدام مصدر معين من مصادر البيانات الكبيرة، يجب أن يأتي نتيجة دراسة متأنية لإحتياجاتها الفعلية والآفاق المستقبلية التي يؤسس لها هذا الاستخدام، وإمكانية تحسين المخرجات الإحصائية ضمن معادلة التكلفة مقابل المنافع، وإمكانية الدمج مع بيانات أخرى، ومساهمة المصدر الجديد في تخفيف العبء عن المستجوبين، ومدى استدامته، بالإضافة الى تلبية لمعايير الجودة مثل الدقة والصلة بالواقع والوقئية²⁶.

5. التجارب الدولية

لعب القطاع الخاص وشركات الاعمال دورا رياديا في التجارب التي حدثت للاستفادة من البيانات الكبيرة وذلك لتعزيز تسويق المنتجات وزيادة الأرباح وتحسين الخدمات.

في القطاع الرسمي والحكومي يندر وجود تجارب ناضجة ومكتملة في مجال الاستفادة من البيانات الكبيرة وإنما هي محاولات لاستكشاف مجالات وأفاق الاستفادة العملية منها كمصدر اضافي لإنتاج الاحصاءات الرسمية. لتوحيد هذه الجهود وتعزيز الاستفادة منها ومشاركة المعرفة والخبرة قامت اللجنة الاحصائية في الأمم المتحدة (بصفتها المسؤولة عن الاحصاءات الرسمية في العالم) بالأنشطة التالية:

- خلال اجتماعها في الدورة 45 سنة 2014 اقرت اللجنة الاحصائية بأهمية البيانات الضخمة كمصدر حيوي للبيانات الاحصائية
- في شهر أيار 2014 قامت اللجنة الاحصائية بتشكيل لجنة فنية متخصصة للتعامل مع مواضيع البيانات الضخمة بحيث تكون هي العنوان الرئيسي للجنة الاحصائية في هذا المجال
- اجتمعت لجنة الخبراء في شهر تشرين أول / 2014 في بكين-الصين مباشرة بعد انعقاد المؤتمر الدولي الذي ناقش استخدام البيانات الضخمة في الاحصاءات الرسمية
- قامت الشعبة الاحصائية بالأمم المتحدة (UNSD) بالتعاون مع اللجنة الاقتصادية للأمم المتحدة في اوربا (UNECE) بإجراء مسح عن البيانات الضخمة والمشاريع في هذا الاطار.

فيما يلي عرض لمشاريع* ذات علاقة مصدرها المسح المشار إليه:

- تطوير مناهج ونماذج التدريب على استخدام البيانات الكبيرة للإحصاءات الرسمية/UNESCAP[†]:
- يهدف هذا المشروع الى تطوير مناهج تحدد احتياجات بناء القدرات المتعلقة بفهم وتقييم استخدام البيانات الكبيرة في الإحصاءات الرسمية، خصوصا في بلدان آسيا والمحيط الهادئ، وذلك استنادا الى تقييم معارف ومهارات الموارد البشرية لهذه البلدان.
- مشاريع متعددة في مجال بناء القدرات
- التمثيل البصري للبيانات الكبيرة Data Visualization
- البرمجة الآلية في استخراج المحتوى من نصوص البيانات الكبيرة
- إنتاج إحصاءات الحصاد بالاعتماد على صور الاقمار الصناعية
- إنتاج إحصاءات زراعية باستخدام بيانات الاقمار الصناعية والاستشعار الأرضي
- الكشف عن المساكن غير المأهولة/الاحصاء البريطاني: يستخدم هذا المشروع عدادات قياس كمية الكهرباء لتقدير عدد المساكن غير المأهولة
- بيانات الهاتف المحمول الإجمالية لتحديد أنماط التنقل/الاحصاء البريطاني
- تحليل بيانات تويتز/الاحصاء المكسيكي: يتم في هذا المشروع قياس الرفاه الشخصي من خلال تحليل المحتوى، والحصول على معلومات ذات علاقة بالتنقل والسياحة من خلال تحليل الموقع
- استخدام خاصية نظام تحديد الموقع العالمي GPS المضمنة في الهواتف النقالة لإنتاج تقديرات سكانية
- إنتاج مؤشرات قصيرة المدى ذات علاقة بميزانية الاسرة

* القائمة تشمل المشاريع المخطط لها أو المنفذة فعليا.

† اللجنة الاقتصادية والاجتماعية لآسيا والمحيط الهادئ.

- التسجيل النقدي عبر الانترنت من متاجر البيع بالتجزئة/الاحصاء الصيني
- مؤشرات ذات علاقة بميول الافراد/الاحصاء الصيني: يستند هذا المشروع الى بيانات مصدرها رسائل التواصل الاجتماعي
- احصاءات بيئية بما في ذلك حساب المنبعثات
- جرد المباني غير السكنية (دراسة جدوى)/ الاحصاء الكندي
- الانترنت كمصدر بيانات لاستخدام تكنولوجيا المعلومات والاتصالات من قبل الشركات والمؤسسات العامة/الاحصاء الايطالي
- تحليل المنهجيات المتعلقة باستخدام الانترنت/يوروستات
- احصاءات اسعار المستهلك/يوروستات
- استخدام الانترنت في تقدير الوظائف الشاغرة
- مشاريع إنتاج إحصاءات سياحة استنادا الى تحديد مواقع الهواتف النقالة، والكاميرات على الحدود
- مؤشرات كثافة المرور استنادا الى بيانات مجسات الاستشعار²⁷

6. التحديات

يثير استخدام البيانات الكبيرة تحديات عديدة تشمل الجوانب التشريعية والخصوصية والمالية والادارية والمنهجية والتكنولوجية²⁸.

1.6 الجوانب التشريعية

تمنح التشريعات في بعض البلدان الحق في الوصول إلى بيانات المنظمات الحكومية وغير الحكومية، وفي بلدان أخرى يقتصر هذا الحق على الوصول إلى بيانات المنظمات الحكومية فقط، مما يؤدي إلى تقييد هذا الوصول²⁹.

بغض النظر عن هذا التباين، فإن مواكبة عصر البيانات الكبيرة يحتاج الى تطوير جذري في التشريعات التي تحكم عمل المؤسسات الإحصائية ومختلف المنظمات العاملة في بلدان العالم، بحيث تعكس هذه التشريعات طبيعة نشوء وتدفق واستخدام البيانات.

فيما يتعلق بفلسطين فإن الجهاز المركزي للإحصاء الفلسطيني هو السلطة الرسمية المخولة بموجب أحكام القانون بطلب البيانات الخاصة بعمل الإحصاءات الرسمية من المؤسسات الحكومية وغير الحكومية، والمصالح الاقتصادية الخاصة³⁰.

2.6 الجوانب المتعلقة بالخصوصية (السرية)

تُعرف الخصوصية عادة باعتبارها حق الأفراد في السيطرة والتأثير على كشف المعلومات المتعلقة بهم. بالتوازي مع المؤسسات التي ترغب في حماية قدرتها التنافسية وحماية زبائنهم.

تثير البيانات الكبيرة إشكالية مفادها أن مستخدمي البيانات والأجهزة الذين يعتبروا مصدر أساسي لتوليد البيانات، هم على الأرجح غير مُدركين للإستخدامات التي يمكن أن ينتج عنها توليد بياناتهم³¹.

تضمن المادة 17 من قانون الإحصاءات العامة الفلسطيني لسنة 2000 سرية البيانات الفردية بجميع أنواعها³². كذلك اعتمد مجلس الوزراء الفلسطيني في العام 2005 تبني فلسطين رسميا للمبادئ الأساسية للإحصاءات الرسمية الصادرة عن اللجنة الإحصائية للأمم المتحدة، يتعلق المبدأ السادس بالسرية (الخصوصية)، وينص على أنه "يتعين إضفاء السرية التامة على البيانات المتعلقة بالأفراد التي تجمعها الوكالات الإحصائية لأغراض إعداد الإحصاءات، سواء تعلق ذلك بأشخاص طبيعيين أو معنويين،

ويتعين استخدامها بالضرورة في الأغراض الإحصائية³³ وفي سياق تطبيق القانون والمبدأ السابقين، حدد الجهاز المركزي للإحصاء الفلسطيني في وثيقة "ميثاق الممارسات للإحصاءات الرسمية الفلسطينية" الكيفية التي يتم فيها تحقيق مبدأ سرية البيانات.

3.6 الجوانب المالية

من المرجح أن تلعب التكلفة دوراً أساسياً فيما يتعلق بالحصول على البيانات الكبيرة من القطاع الخاص، خصوصاً في ظل غياب تشريعات ملزمة في هذا الشأن.

لذا فإن حسابات الكفاءة السعرية يجب أن تكون هي العامل الأساسي في الاختيار الصحيح، من حيث تحقيق التوازن بين الجدوى وضمان الجودة (مثل الوقتية والدقة والاتساق وسهولة الوصول) وتخفيف العبء عن المستجوبين في مقابل تكاليف الحصول على البيانات ومتطلبات معالجتها³⁴. إن مبدأ الشراكة والمنفعة المتبادلة بين الأجهزة الإحصائية والقطاع الخاص من الممكن أن يساهم في تقليص هذه النفقات جذرياً.

4.6 الجوانب الإدارية

البيانات الكبيرة للإحصاءات الرسمية تعني مزيد من المعلومات الواردة التي تخضع لسياسات وتوجيهات بشأن إدارة هذه المعلومات وحمايتها.

يتعلق التحدي الإداري الآخر بالموارد البشرية³⁵، فعلم البيانات الكبيرة أكثر تطوراً في القطاع الخاص منه في الأجهزة الإحصائية، مما يستدعي قيام هذه الأجهزة بالاستثمار في بناء قدرات العاملين لديها، إضافة إلى بناء قاعدة معلومات القدرات المتعلقة بالبيانات الكبيرة على المستوى الوطني.

5.6 الجوانب المتعلقة بالمنهجيات

يقوم الإحصائيون في عالم البيانات التقليدية بتعريف مجتمع الهدف ومجتمع المسح، وبناء إطار المعاينة للوصول إلى هذا المجتمع، وسحب العينة، وجمع البيانات، إلى آخر الخطوات اللازمة لتنفيذ المسح ومعالجة ونشر البيانات.

وكانهم بذلك ينشئون صندوق ويملئونه بالبيانات بطريقة منظمة جداً ومسيطر عليها تماماً.

في عالم البيانات الكبيرة فإن هذه البيانات هي التي تأتي أولاً بسماتها الضخمة والسريعة والمتنوعة، وعلى الإحصائي أن يتكيف مع طبيعتها ويستخدم أدوات جديدة تتلائم معها، أي أنه يجب أن يفكر خارج الصندوق.

تستدعي الجوانب المتعلقة بالمنهجيات ضرورة التحليل السريع (إن لم يكن في الوقت الحقيقي) لبيانات ضخمة جداً، وإجتراح أساليب قادرة على دمج المعلومات التي استخرجت في العملية الإحصائية، مثل الربط على نطاق واسع بين مجموعات البيانات الكبيرة المختلفة، والوصول إلى نتائج ذات موثوقية عالية عند تطبيق هذه الأساليب على مجموعات البيانات الكبيرة جداً.

إن استخدام البيانات الكبيرة في المؤسسات الإحصائية يوجب استخدام تقنيات جديدة تستجيب لمجموعة من التحديات مثل:

- قياس جودة المخرجات الناتجة عن بيانات خارجية يصعب إدارتها
- قيمة البيانات الواردة من المصادر الخارجية
- صعوبة دمج المعلومات متعددة المصادر³⁶

6.6 الجوانب التكنولوجية

إن تحسين سرعة الوصول إلى البيانات الإدارية يتطلب استخدام مكثف لواجهات برمجية معيارية؛ مما يتيح إمكانية ربط تطبيقات النقاط البيانات ومعالجتها مباشرة من مصادر البيانات الإدارية

إن جمع البيانات في الوقت الحقيقي (أو في زمن قريب من الوقت الحقيقي) يُعظم الاستفادة ويفتح فرص جديدة للجمع بين البيانات الإدارية مع البيانات عالية السرعة القادمة من مصادر مختلفة أخرى، مثل:

- البيانات التجارية (معاملات بطاقات الائتمان، حركة المعاملات، المبيعات، وما إلى ذلك)
- أجهزة التعقب (الهواتف النقالة، نظم تحديد الموقع العالمي GPS، كاميرات المراقبة، تطبيقات) وأجهزة الاستشعار (المروور، الأرصاد الجوية، التلوث، الطاقة، ...)
- وسائل الإعلام الاجتماعي (تويتر، الفيسبوك، ...) ومحركات البحث على الإنترنت.

في عصر البيانات الكبيرة، فإن النموذج الجديد يستند إلى إمكانية جمع ودمج البيانات من مختلف الأنواع والمصادر وتجميعها لإنتاج معلومات جديدة، يؤثر هذا النموذج تحدي إضافي في المستقبل، وفي حال تحققه من خلال الربط بين المصادر التقليدية للبيانات مثل المسوح والسجلات الإدارية مع مصادر البيانات الكبيرة، فإنه يوفر فرصة كبيرة للوصف الذكي لسلوكيات المجتمعات³⁷.

7. الجهاز المركزي للإحصاء الفلسطيني

تتمثل رسالة الجهاز المركزي للإحصاء الفلسطيني في "إنتاج ونشر الإحصاءات الرسمية المتسقة، الموضوعية، ذات الجودة العالية وفي الوقت المناسب، لتلبية الحاجات الحالية والمستقبلية للمستخدمين على المستوى الوطني والدولي في إطار من الشفافية وبطريقة مناسبة بالاعتماد على أفضل الممارسات الإحصائية"³⁸.

ترتبط أهمية استخدام البيانات الكبيرة لدى الجهاز بتحقيق ما ورد في رسالته من حيث تلبية حاجات المستخدمين في وقت أقصر مما تحتاجه الطرق الإحصائية التقليدية (تحسين الوقتية)، وتوقع الاحتياجات المستقبلية لهم وتوفيرها باستخدام أحدث الطرق.

1.7 مجالات الاستفادة

يمكن تلخيص مجالات الاستفادة وأسباب التوجه نحو استخدام البيانات الكبيرة من وجهة نظر الجهاز بما يلي:

- تحديث عملية الإنتاج الإحصائي
 - المشاركة في تطبيق سياسة الحكومة المتعلقة بمجتمع المعلومات
 - تلبية متطلبات توفير مؤشرات جديدة مثل مؤشرات الأهداف الإنمائية المستدامة
- Sustainable Development Goals (SDG)

- توفير منتجات وخدمات جديدة
- تخفيض التكلفة [على المدى المتوسط]
- تخفيف العبء عن المستجوبين
- توفير إحصاءات بسرعة عالية

2.7 مصادر البيانات

يدرس الجهاز في المرحلة الحالية استخدام البيانات الكبيرة المتعلقة بمجموعة من المواضيع (مرتبة حسب الأولوية):

1. بيانات الصور الجوية والاقمار الصناعية
2. بيانات الهواتف النقالة

3. بيانات المجسات الكهربائية
4. بيانات استشعار كثافة المرور
5. السجلات الطبية
6. الوسائط الاجتماعية
7. سجلات الانترنت

من المتوقع أن تساعد هذه المصادر في توفير مؤشرات مجالها استعمالات الاراضي، الصحة، الطاقة، النقل، المصارف والتأمين والمالية، التجارة العالمية، الاسعار، البيئة، مجتمع المعلومات، الأهداف الإنمائية المستدامة.

3.7 المتطلبات

يرى الجهاز المركزي للإحصاء الفلسطيني بأن تحقيق رؤيته لإستخدام البيانات الكبيرة يحتاج الى توفير مجموعة من المتطلبات والموارد:

1. بناء علاقات تعاون وشراكات استراتيجية مع الجهات ذات العلاقة مثل: المؤسسات الإحصائية الاقليمية والدولية، المؤسسات الاكاديمية، مؤسسات البحث العلمي، مؤسسات القطاع الحكومي، شركات الهواتف النقالة، مُزودي الصور الجوية وصور الاقمار الصناعية، مُزودي خدمات الوسائط الاجتماعية، شركات خدمات تكنولوجيا المعلومات
2. توفير مصادر لتمويل مشاريع البيانات الكبيرة
3. التوعية والإعلام لجمهور المستخدمين بمختلف فئاتهم فيما يتعلق بالتحول نحو استخدام البيانات الكبيرة
4. الاستفادة من التجارب الناجحة في مجال استخدام البيانات الكبيرة في الإحصاءات الرسمية
5. تنظيم اجتماعات وندوات ومحاضرات لصناع القرار والقادة، بهدف زيادة حلقة المؤيدين لاستخدام البيانات الكبيرة في الإحصاءات الرسمية
6. العمل على وضع لوائح واعتماد لوائح منظمة للتعامل مع خصوصية وسرية البيانات الكبيرة
7. بناء القدرات المتعلقة بالتعامل مع البيانات الكبيرة من خلال تدريب وتأهيل العاملين في الجهاز، والاستعانة بكفاءات خارجية إن لزم الامر، خصوصا في المجالات التالية: أمن المعلومات، بنية تكنولوجيا المعلومات، علم البيانات، التمثيل البصري للبيانات، الرياضيات والإحصاء والنمذجة، إدارة مشاريع البيانات الكبيرة، هندسة البرمجيات
8. تعميم التوجيهات والإرشادات ذات العلاقة، على مختلف مكونات النظام الإحصائي الرسمية، وذلك فيما يخص: حشد التأييد والتواصل الفعال، الخصوصية والسرية، الوصول للبيانات الكبيرة، تنمية المهارات، إطار جودة البيانات الكبيرة، طرق التقدير، تصنيف المصادر، ربط البيانات الكبيرة مع المؤشرات المتعلقة بالأهداف الإنمائية المستدامة، استخدام بيانات الهواتف النقالة في الإحصاءات الرسمية، استخدام صور الاقمار الصناعية في الاحصاءات الرسمية، استخدام وسائل التواصل الاجتماعي في الاحصاءات الرسمية³⁹

8. دراسة حالة: نظام إدارة المعلومات المكانية في فلسطين

اتخذت وزارة الحكم المحلي الفلسطينية قرارا استراتيجيا بتطوير نظام معلومات جغرافية متكامل GeoMOLG* يُسهل الوصول للمعلومات المكانية، وذلك بهدف تقليص الوقت والجهد والتكلفة فيما يتعلق بالإجراءات الرسمية التي تتطلب مسبقا توفر المعلومات الأساسية التي يتم بناء القرارات والتوصيات استنادا إليها، بما يعنيه ذلك من تقديم خدمة أفضل للمواطنين فيما يتعلق بالمعلومات المكانية

*GeoMOLG: The Integrated Spatial Information System of the Ministry of Local Governance.

وكذلك التسهيل على الموظفين الحكوميين من حيث الوصول الى هذه المعلومات وتزويد المواطنين بها بأعلى جودة ممكنة وبأقصر وقت.

1.8 الفوائد المباشرة للنظام

تتمثل الفوائد المباشرة لنظام GeoMOLG بتقليص الوقت والجهد والتكلفة في الوصول للمعلومات المكانية، وذلك نتيجة مباشرة لما يلي:

1. انخفاض عدد الزيارات الميدانية بشكل ملحوظ نظرا لتوفر جميع المعلومات المكانية بالإضافة الى صور جوية حديثة بدرجة وضوح تصل الى 25 سم، مما يمكن من قراءة الواقع بشكل دقيق
 2. تقليص الجهد اللازم للوصول للمعلومات المكانية نظرا لتوفير جميع المعلومات بصيغة إلكترونية بديلا عن النسخ الورقية
 3. تقليص وقت الوصول للمعلومات نتيجة لإمكانية الوصول إليها عبر تطبيق شبكة إلكترونية يحتوي على جميع الأدوات والمهام اللازمة لإتمام العمل
- تشير التقديرات الأولية الى توفير سنوي مقداره 1800 يوم عمل وحوالي 112 ألف دولار.

2.8 الاعتبارات الأساسية في تصميم النظام

تم اتخاذ عدد من الاعتبارات الأساسية في تصميم النظام منها:

1. أن يشتمل على جميع المعلومات المكانية التي تتطلبها طبيعة عمل وزارة الحكم المحلي على اختلاف إداراتها ودوائرها ومديرياتها
2. الدخول الى النظام عبر شبكة الانترنت دون أن يتطلب الوصول إليه تنصيب أي من التطبيقات أو البرامج المتخصصة على جهاز المستخدم النهائي
3. أن يكون سهل الاستخدام بحيث يتمكن المستخدم النهائي من التدريب عليه بوقت قصير، دون أن يكون المستخدم النهائي خبيرا في نظم المعلومات المكانية أو تكنولوجيا المعلومات
4. أن يتم تغذية النظام بالبيانات الأكثر دقة وتحديثا، بحيث يضمن النظام إمكانية تحديث البيانات عبر شبكة الانترنت
5. أن يكون قابلا للربط مع النظم وقواعد البيانات المختلفة

3.8 مميزات النظام

1. تعتبر تجربة الحكم المحلي هي الأضخم من نوعها على المستوى الوطني بحيث تم توفير الكم الأكبر والأهم من المعلومات المكانية في موقع واحد وبأعلى قدر من الدقة
2. تم استخدام أكثر تطبيقات الانترنت المكانية ملائمة لتسهيل الوصول للمعلومات المكانية من خلاله
3. تم استخدام تكنولوجيا (ArcSDE*) لتمكين تحديث المعلومات في نفس الوقت من قبل عدة مستخدمين
4. تم تصميم وتطوير النظام بالكامل من قبل فريق وزارة الحكم المحلي
5. يعتبر نواة حقيقية لإيجاد نظام معلومات مكاني موحد تشارك فيه كافة المؤسسات بحيث تتولى كل مؤسسة من ذات الاختصاص مسؤولية صيانة وتحديث البيانات المتعلقة بها

4.8 منهجية بناء النظام

تم اعتماد منهجية تتكون من خمسة عناصر لبناء نظام معلومات مكانية يستجيب بدقة لاحتياجات العمل ويحقق النتائج المتوقعة مع مراعاة التسلسل في تنفيذها ويمكن توضيحها على النحو التالي:

* ArcSDE : Spatial Database Engine.

أولاً: متطلبات العمل

تم حصر المهام التي من المفترض أن يقوم بها النظام بحيث يتم استبدال الأساليب التقليدية بأخرى محوسبة ومثال ذلك تزويد النظام برقم القطعة والحوض والبلدة لأرض معينة ومن ثم الحصول على جميع المعلومات التنظيمية الخاصة بها من عدد الطوابق المسموح بها والارتدادات والرسوم وغيرها.

ثانياً: جمع البيانات

تم تحديد المعلومات المكانية والبيانات اللازم توفرها في النظام من أجل الاستجابة للمهام والوظائف المحددة سلفاً.

أ. المعلومات التي يوفرها النظام

تم توفير المعلومات الأساسية التي تتطلبها طبيعة عمل وزارة الحكم المحلي والتي يمكن إجمال أهمها على النحو التالي:

1. المخططات الهيكلية (Urban Master Plan) وتشمل: المخططات الهيكلية المصدقة في مناطق (أ) و (ب)، المخططات الهيكلية في مرحلة الاعتراض في مناطق (أ) و (ب)، المخططات الهيكلية المصدقة في مناطق (ج)، المخططات الهيكلية في مرحلة الاعتراض في مناطق (ج)، المخططات الهيكلية التفصيلية، المخططات الهيكلية الجزئية، مخططات التعديل التنظيمي ومخططات التنظيم التفصيلي، المخططات الكادستراية والتي تشمل القطع والأحواض
2. مناطق (أ) و (ب) و (ج)
3. خرائط تصنيف الأراضي حسب القيمة الزراعية (عالية، متوسطة أو منخفضة)
4. مناطق التنوع الحيوي، مناطق المحميات الطبيعية، مناطق المشاهد الطبيعية، المواقع الأثرية، الأودية، خطوط الكنتور، أراضي الدولة (المسجلة والمعلنة والممسوحة)، مواقع التجمعات والضواحي، مواقع المباني الرسمية والخدمات العامة، حدود التجمعات الفلسطينية، حدود المحافظات
5. صورة جوية حديثة (سنة 2014) بدرجة وضوح تصل إلى 25 سم وبدقة مكانية تصل إلى 50 سم في بعض المناطق

ب. أهمية الحصول على المعلومات المكانية في الواقع الفلسطيني
يوضح جدول 2 بعض القرارات التي تعقب الحصول على المعلومات المكانية:

جدول 2: القرارات/المتطلبات الناتجة عن الحصول على المعلومات المكانية

المعلومات المكانية	المتطلبات
المخططات الهيكلية	عند الحصول على تصنيف قطعة أرض معينة حسب الاستعمالات المبنية في المخطط الهيكلي فإنه يتم تحديد عدد الطوابق المسموح بناؤها والارتدادات والإفراغات والرسوم والمخالفات وغيرها
مناطق (أ) و (ب) و (ج)	جميع الأراضي الواقعة في مناطق (أ) تتبع إداريا وأمنيا دولة فلسطين بينما الأراضي الواقعة في مناطق (ب) فإنها تتبع إداريا دولة فلسطين وأمنيا الجانب الإسرائيلي، في حين أن الأراضي الواقعة في مناطق (ج) تقع تحت السيطرة الاسرائيلية كليا. وبناء عليه فإنه عند طلب ترخيص لإقامة مشروع أو بناء في مناطق (أ) أو (ب) فإن الحصول على التراخيص يتم من الجانب الفلسطيني، بينما في مناطق (ج) فإن ذلك يتم عن طريق الجانب الإسرائيلي. تجدر الإشارة الى أن خدمات المياه والكهرباء يتم تزويدها في مناطق (أ) و (ب) في حال إثبات أن البناء مرخص، أما المباني المقامة في مناطق (ج) فإنه يتم إيصال هذه الخدمات إليها مقابل تعهد بالترخيص دون أن تتحمل الجهة المانحة للترخيص أي تبعات لذلك
التصنيف الزراعي	لا يسمح البناء في جميع الأراضي المصنفة بأنها عالية القيمة الزراعية بينما يتم دراسة كل حالة على حدة بخصوص منح تراخيص للبناء في الأراضي المصنفة بأنها متوسطة القيمة الزراعية، أما فيما يتعلق بالأراضي منخفضة القيمة الزراعية فلا يوجد مانع من وجهة النظر الزراعية لإنشاء المباني عليها
القطع والاحواض	يتم الحصول على مواقع جميع الأراضي التي تتوفر لديها خزانات تسوية والحصول على مساحاتها ومقدار الإفراز فيها
المحميات الطبيعية	لا يسمح البناء ضمن جميع الأراضي المصنفة على أنها محميات طبيعية
التنوع الحيوي	لا يسمح البناء ضمن جميع الأراضي المصنفة على أنها ذات تنوع حيوي
المواقع الأثرية	لا يسمح البناء ضمن جميع الأراضي المصنفة على أنها مواقع أثرية
الأودية	لا يسمح البناء ضمن حرم الوادي المعروف بالنظام
الصورة الجوية	يمكن قراءة الواقع من خلالها دون الحاجة للزيارات الميدانية فمثلا في جميع الأراضي الواقعة خارج المخططات الهيكلية فإنه يسمح ببناء واحد فقط فلا يسمح بتعدد الأبنية. عند تقديم طلب للترخيص في هذه الأراضي يمكن التأكد من خلال صورة الوضع القائم في هذه الأراضي فلا تصدر رخصة في حال وجود بناء عليها.

ج. آلية تحديث البيانات

حيث أن وزارة الحكم المحلي هي جهة الاختصاص فيما يتعلق بالمخططات الهيكلية فقد تم صدور قرار رسمي يقضي بتزويد دائرة نظم المعلومات الجغرافية GIS بجميع التحديثات على المخططات الهيكلية سواء تلك التي تم إقرارها وتصديقها على مستوى اللجان الإقليمية أو على مستوى مجلس التنظيم الأعلى، مما يضمن توفير المعلومات بأقصى درجة من التحديث. من المخطط له تدريب كادر المديرية ليتمكن من تحديث البيانات التي تتم على مستوى المحافظة ليتم إدخالها على النظام دون الحاجة لإرسالها إلى دائرة نظم المعلومات الجغرافية المركزية في الوزارة.

د. نظام تحديد الموقع العالمي GPS لتدقيق التصوير الجوي

يعتمد تصميم المخططات الهيكلية بالأساس على رفع جميع العناصر المكانية (الأبنية، الشوارع، الأشجار، الأسوار، الخ) لمنطقة المراد تخطيطها. وبهذا فإنه كلما زادت الدقة فيما يتعلق برفع هذه العناصر تزداد الدقة في التصميم. بشكل عام، يمكن إجمال الخطوات التي تتم من أجل توفير العناصر المكانية للقيام بتصميم المخططات الهيكلية على النحو التالي:

1. التقاط الصور الجوية بمقياس رسم ودرجة وضوح ودقة مكانية معينة
2. يتم معالجة الصور للوصول الى اقصى دقة ممكنة
3. القيام بترسيم جميع العناصر المكانية من الصورة. من الجدير ذكره أنه في حالات كثيرة تم اكتشاف انخفاض مستوى دقة العناصر المكانية المرسومة، ولهذا عمدت الوزارة إلى تبني تكنولوجيا نظام تحديد الموقع العالمي GPS للتأكد من مدى مطابقة إحداثيات العناصر المكانية المرسومة من الصورة الجوية مع تلك المرصودة بواسطة الـ GPS. بشكل عام هناك عدد من المواصفات الفنية للصورة الجوية منها: الدقة المكانية للصورة، درجة وضوح الصورة، مقياس رسم الصورة الملتقطة، درجة الراديو متر، درجة السبيكترال، تاريخ التقاط الصورة، وقت التقاط الصورة، نظام الإحداثيات للصورة، نوع الصورة

5.8 مستخدمو النظام الحاليون

كانت الرؤية لدى الوزارة أن يتم تشغيل النظام على مستوى الوزارة وإداراتها ومديرياتها في المرحلة الأولى، ومن ثم إتاحة المجال لعدد من المؤسسات من الدخول إليه واستخدامه، وفي مرحلة لاحقة إتاحتها للجمهور عبر شبكة الانترنت. تجدر الإشارة إلى أنه تم تشغيل النظام على المستويين الأول والثاني (الوزارة وبعض المؤسسات) ويتم دراسة إتاحة النظام للجمهور عبر شبكة الانترنت في المرحلة القادمة.

6.8 تفعيل مزايا متقدمة في النظام

تم العمل على ضرورة اشتمال النظام على مزايا متقدمة منها ما هو مفعل وأخرى لازالت طور التجهيز وهي كما يلي:

1. Geocoding: والذي يؤمن الوصول إلى الأماكن عبر تزويد النظام بالعنوان (رقم البناية وإسم الشارع وإسم البلدة) وذلك في التجمعات التي لديها قواعد بيانات لأنظمة العناوين
2. Fabric: لضمان الحفاظ على التسلسل التاريخي عند تقسيم وإفراز الأراضي أو توحيدها
3. Geometric Network: لبناء أنظمة تساعد في إدارة شبكات المياه والكهرباء والهاتف والصرف الصحي
4. Network Analyst: لإدارة شبكات الطرق والقدرة على تحديد أفضل المسارات بأفضل زمن أو مسافة
5. Geoevent: لإمكانية نقل البيانات الحية من الميدان عبر أجهزة مخصصة لهذا الغرض مرتبطة بالنظام

7.8 الخطط المستقبلية

هناك عدد من المهام التي يتم العمل على إنجازها للاستمرار في تحسين مستوى الخدمة المقدمة للجمهور من قبل دوائر ومديريات الحكم المحلي فيما يتعلق بالمعلومات المكانية ومنها:

1. ترسيم خط الارتداد
2. إصدار رخص الأبنية الكترونياً
3. إيداع المخططات الهيكلية وجمع الاعتراضات عبر النظام من خلال شبكة الانترنت
4. جعل النظام متاحاً عبر الهواتف الذكية بما يتناسب مع شاشات العرض الخاصة بها

إن حجم العمل المناط بوزارة الحكم المحلي يملئ عليها بالضرورة تبني أحدث التقنيات والوسائل التكنولوجية المتطورة مما يمكنها بالنهاية من توفير الخدمة للمواطنين بأفضل أسلوب ممكن من جانب، وزيادة حافية الموظفين وضمان تأديتهم لأعمالهم بأسلوب عصري. من جانب آخر فإن تطوير الكادر

ومواكبته لكل ما يستجد في هذا المضمار عبر المشاركات المحلية والإقليمية والدولية وبرامج بناء القدرات المختلفة هو السبيل الأنجع لعكس الخبرات المكتسبة في تطوير آلية العمل وضمان التميز في تقديمه⁴⁰.

8.8 مقترحات لتطوير المشروع

يتضح من الفقرات السابقة أن نظام إدارة المعلومات المكانية في فلسطين، قد قطع شوطاً كبيراً لتحقيق الهدف من إنشائه.

- يمكن البناء على إنجازات النظام وتطويرها من خلال تنفيذ مجموعة من المهام يقترحها الباحث:
1. بناء شراكة تضم وزارة الحكم المحلي والجهاز المركزي للإحصاء الفلسطيني والجهات ذات العلاقة لتطوير نظام (GeoMOLG)
 2. تبادل الخبرات فيما يتعلق بتصنيف استعمالات الأراضي بما يتوافق مع التوصيات الدولية
 3. إجراء تخطيط شامل للأراضي الفلسطينية بالتعاون مع كافة المؤسسات ذات العلاقة
 4. تحديد المتطلبات والمواصفات المتعلقة بالصور الجوية مثل درجة الوضوح ودورية التحديث، بحيث تحقق إمكانية تنفيذ المشاريع المشتركة بمستوى جودة مرتفع
 5. بناء القدرات لمختلف الشركاء فيما يتعلق باستخدام النظام وتحديثه وتطويره
 6. ربط النظام بقواعد بيانات التعداد والمسوح التي ينفذها الجهاز
 7. تنفيذ مسوح متعددة الأغراض والتي تشمل المسوح الزراعية والبيئية ومسوح التجمعات السكانية ومسوح النقل، بحيث يتم في هذه المسوح استخدام الصور الجوية والطرق التقليدية معاً
 8. تطوير البنية التحتية للمؤسسات الشريكة بما يتلاءم ومتطلبات النظام من حيث الاستخدام والتحديث

9.8 جدوى استخدام نظام GeoMOLG

يتضمن جدول 3 ملخص لنتائج فحص جدوى استخدام النظام (بعد تنفيذ المقترحات الواردة في البند السابق) مقارنة بالطرق التقليدية:

جدول 3: ملخص نتائج فحص جدوى نظام GeoMOLG

ملاحظات	المسوح التقليدية	استخدام نظام GeoMOLG	البند
	تعطي البيانات المستخرجة من المسوح التقليدية بيانات أقل دقة بسبب عدم وجود سجلات أراضي وحيازات زراعية موثقة في فلسطين	تعطي البيانات المستخرجة من الصور الجوية بيانات أدق حول المساحات الزراعية في ظل عدم وجود سجلات أراضي وحيازات زراعية موثقة في فلسطين	الدقة
	طاقم عمل كبير جداً من النواحي الإدارية والإشرافية والميدانية	طاقم عمل قليل نسبياً	طاقم العمل
بداية بناء النظام تحتاج إلى وقت طويل نسبياً، لكنه يقل فيما بعد	بحاجة لوقت وجهد كبير	الوقت والجهد المبذول قليل في الميدان	الوقت والجهد
	وجود تحيز وأخطاء معابنة خاصة في حال عدم وجود إطار دقيق، وهو عرضة للأخطاء البشرية	تحيز وأخطاء احصائية قليلة وغير عرضة للأخطاء البشرية	الجودة

ملاحظات	المسوح التقليدية	استخدام نظام GeoMOLG	البند
تتخفص التكلفة عند استخدام النظام ما بين 8% الى 35% من تكلفة استخدام المسوح التقليدية	تكلفة عالية	تكلفة أقل (لصيانة النظام) تكلفة اعلى عند بداية بناء النظام	التكلفة
في حالة وجود قاعدة بيانات جغرافية للسكان يمكن أن تغطي البيانات المستخرجة الخصائص الديموغرافية للسكان مع البيانات المستخرجة من الصور الجوية	- هناك امكانية للتكرار والاسقاط لقطع الاراضي - تغطي البيانات المستخرجة الخصائص الديموغرافية للسكان	- تلافي تكرار قطع الاراضي او اسقاط أي منها (تغطي كافة اراضي الدولة) - لا تغطي البيانات المستخرجة العديد من المتغيرات مثل الخصائص الديموغرافية للسكان، وبعض التطبيقات الزراعية	الشمولية
	بحاجة لوقت وجهد كبير لتمكين صانعي السياسات من اتخاذ القرار	يمكن ربط بيانات جغرافية كبيرة في نفس الوقت مما يمكن من اتخاذ القرارات بسرعة فائقة	تعدد الأغراض واتخاذ القرار
	يتم توفير البيانات ونشرها بعد وقت طويل من الاسناد الزمني للمسوح، مما يقلل من أهميتها أحيانا	يتم توفير البيانات ونشرها في الوقت المحدد وبسرعة فائقة	نشر البيانات

يتضح من النتائج السابقة بأن هناك جدوى فعلية لاستخدام نظام GeoMOLG الذي يعتمد على آليات حديثة هي الصور الجوية وذلك بما يتكامل مع تنفيذ المسوح بالطريقة التقليدية.

9. الاستنتاجات

1. البيانات الكبيرة ظاهرة ثقافية-تكنولوجية، مُركبة، تمتاز بالضخامة والتعقيد، ما يحول دون القدرة على معالجتها باستخدام أدوات برمجية تقليدية
2. تنمو البيانات الكبيرة بوتيرة عالية جداً، وستستمر بالنمو في المستقبل المنظور
3. الاستثمار في استخدام البيانات الكبيرة مفيد بشكل عام لمختلف أنواع المؤسسات من حيث إتاحة إستخراج القيمة العالية من البيانات بما يُحسن من فرص إتخاذ القرارات السليمة في الوقت المناسب
4. استخدام البيانات الكبيرة من قبل المؤسسات الإحصائية يتيح إنتاج إحصاءات أفضل من حيث الوقئية والصلة بالواقع، بكفاءة سعرية عالية، بالإضافة الى تخفيف العبء على المبحوثين.
5. يثير استخدام البيانات الكبيرة تحديات تتعلق بالجوانب التشريعية والخصوصية والمالية والادارية والمنهجية والتكنولوجية
6. يوجد في فلسطين مؤسسات عديدة لديها بيانات كبيرة، مما يعطي فرصة واعدة للإستفادة من هذه البيانات
7. هناك جدوى فعلية لاستخدام نظام إدارة المعلومات الجغرافية في فلسطين GeoMOLG

10. التوصيات

1. إتخاذ قرار على مستوى المؤسسات الإحصائية العربية بإدراج موضوع استخدام البيانات الكبيرة ضمن خططها الوطنية نظرا للأهمية العالية التي يحتلها هذا الموضوع حاليا وفي المستقبل
2. تنفيذ مشاريع البيانات الكبيرة في المؤسسات الإحصائية استنادا الى دراسة للاحتياجات الفعلية والآفاق التي يفتحها تنفيذ هذه المشاريع والميزات المترتبة على ذلك خصوصا تحسين المخرجات الإحصائية، وإمكانية الدمج مع بيانات أخرى، والمساهمة في تخفيف العبء عن المستجوبين، بالإضافة الى تلبية معايير الجودة مثل الدقة والصلة بالواقع والوقتية
3. قيام المعهد العربي للتدريب والبحوث الإحصائية والمؤسسات الاكاديمية العربية بإستحداث تخصص "علم البيانات" ضمن الخطط التدريبية والتعليمية في مجال جمع وإدارة ومعالجة البيانات الكبيرة
4. الاستعداد والتحضير للتغييرات الهيكلية والمنهجية التي ستحدث في المؤسسات الإحصائية نتيجة لإستخدام البيانات الكبيرة
5. العمل على بناء علاقات تعاون وشراكات استراتيجية مع مختلف المؤسسات ذات العلاقة
6. البحث في توفير مصادر لتمويل مشاريع البيانات الكبيرة
7. التوعية والإعلام لجمهور المستخدمين بمختلف فئاتهم فيما يتعلق بالتحول نحو استخدام البيانات الكبيرة
8. الاستفادة من التجارب الناجحة في مجال استخدام البيانات الكبيرة في الإحصاءات الرسمية
9. تنظيم اجتماعات وندوات ومحاضرات لصناع القرار والقادة، بهدف زيادة حلقة المؤيدين لاستخدام البيانات الكبيرة في الإحصاءات الرسمية
10. العمل على وضع واعتماد لوائح للتعامل مع خصوصية وسرية البيانات الكبيرة
11. قيام الجهاز المركزي للإحصاء الفلسطيني ببناء شراكة مع وزارة الحكم المحلي لتطوير مشروع نظام إدارة المعلومات الجغرافية GeoMOLG وربطه مع البيانات السكانية المتوفرة لدى الجهاز

المراجع

- ¹ Baquero, José Carlos (n.d.). Introduction to Big Data without Getting Bugged Down in Details. GMV Innovation Solutions.
- ² Brahe, Tycho. (2011). Encyclopædia Britannica. Encyclopædia Britannica Ultimate Reference Suite. Chicago: Encyclopædia Britannica.
- ³ Kepler, Johannes. (2011). Encyclopædia Britannica. Encyclopædia Britannica Ultimate Reference Suite. Chicago: Encyclopædia Britannica.
- ⁴ Big Data and the History of Information Storage. Winshuttle.
- ⁵ Cox, Michael and Ellsworth, David. (July 1997). Application-Controlled Demand Paging for Out-of-Core Visualization.
- ⁶ Big Data. Merriam-Webster Online. In Merriam-Webster
- ⁷ Intel (n.d.). Intel's Big Data Policy: A Vision for Big Data.
- ⁸ ISO/IEC JTC. (2015). Big Data Preliminary Report 2014. ⁹ Zikopoulos, Paul, and others. (2015). Big Data Beyond the Hype: A Guide to Conversations for Today's Data Center. McGraw-Hill Education
- ¹⁰ Price, Daniel. (2015). Facts and Stats About The Big Data Industry. Cloud Tweaks.
- ¹¹ AT&T. (2015). AT&T Researchers: Inventing the Science Behind the Service.
- ¹² Vishen, Neeraj. (2013). Largest Databases of the World.
- ¹³ Marr Bernard (n.d.). Big Data: The 5 V's.

-
- ¹⁴ Barnatt, Christopher. (2012). Explaining Big Data.
- ¹⁵ ISO/IEC JTC. (2015). op.cit.
- ¹⁶ Ibid.
- ¹⁷ Australian Government. (2015). Australian Public Service Better Practice Guide for Big Data.
- ¹⁸ Nedelcu, Bogdan. (2015). About Big Data and it's Challenges and Benefits in Manufacturing.
- ¹⁹ Australian Government. (2015). Op.cit.
- ²⁰ Ibid.
- ²¹ Harvard Business Review. (2013). Big Data: The Future of Information and Business.
- ²² Zikopoulos, Paul and others. (2013). Harness the Power of Big Data: The IBM Big Data Platform.
- ²³ Harvard Business Review. (2013). Op.cit.
- ²⁴ الامم المتحدة-اللجنة الإحصائية. (2014). البيانات الضخمة وتحديث النظم الإحصائية: تقرير الأمين العام-الدورة الخامسة والأربعون.
- ²⁵ Tam, Siu-Ming, and Clarke, Frederic. (n.d.). Big Data, Official Statistics and Some Initiatives by the Australian Bureau of Statistics. ²⁶ Ibid.
- ²⁷ Statistics Divisions of UN/DESA and UN Economic Commission for Europe. (2015). Results of the UNSD/UNECE Survey on Organizational Context and Individual Projects of Big Data.
- ²⁸ الامم المتحدة-اللجنة الإحصائية. (2014). مصدر سبق ذكره.
- ²⁹ United Nations Economic Commission for Europe. (2013). What Does “Big Data” Mean for Official Statistics?
- ³⁰ قانون الإحصاءات رقم 4 لسنة 2000. فلسطين.
- ³¹ United Nations Economic Commission for Europe. (2013). Op.cit.
- ³² قانون الإحصاءات رقم 4 لسنة 2000. فلسطين.
- ³³ الجهاز المركزي للإحصاء الفلسطيني. (2006). ميثاق الممارسات للإحصاءات الرسمية الفلسطينية. رام الله – فلسطين.
- ³⁴ United Nations Economic Commission for Europe. (2013). Op.cit.
- ³⁵ Ibid.
- ³⁶ Ibid.
- ³⁷ Ibid.
- ³⁸ الجهاز المركزي للإحصاء الفلسطيني. (2011). سياسة وأهداف الجودة. رام الله – فلسطين.
- ³⁹ Palestinian Central Bureau of Statistics. (2015). Big Data Current Situation. Unpublished Information.
- ⁴⁰ وزارة الحكم المحلي. (2014). نظام وزارة الحكم المحلي المتكامل لإدارة المعلومات المكانية (GeoMOLG). فلسطين