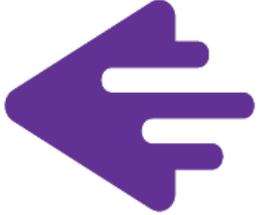


Webinar

التحول التقني



webinars.tts.sa

التحول التقني
TechTrans



مقدمة عن تحليل البيانات الضخمة Big Data Analysis

تقديم :

شذا بنت إبراهيم الخراع
مدرية في المجال التقني والإداري



مقدمة

في الوقت الراهن تُقدم حلول معالجة البيانات الضخمة نتائج ممتازة في معالجة البيانات بعد وقوع المشكلات وفي اكتشاف الاتجاهات التاريخية وتحديد مواطن الأزمات ، لكن تحليل البيانات بأثر رجعي لا يُلائم عالم اليوم بتحولاته السريعة ، وبات من الضروري الاستعداد بأدوات البيانات الضخمة التي تُركّز على احتياجات الأعمال، بالإضافة إلى توفير البنية المناسبة للبيانات والاستعانة بتوجيهات المتخصصين، الأمر الذي يسمح بإنتاج تقارير الفرضيات للتحليلات في الحالات الطارئة والمستقبلية ومن ثم استغلال الفرص الجديدة.

وتستلزم تحليلات الأعمال :

أولاً : جمع البيانات .

ثانياً : تنظيمها في قواعد للبيانات ، تُراعي العلاقات فيما بينها أو ما يُطلق عليه “تسوية قواعد البيانات” .

وبالنظر إلى الكميات الهائلة المُتاحة اليوم من البيانات ، يستحيل على فرق العاملين قراءتها وتحليلها لتحقيق قيمة ملموسة للعمل .

وهو ما يُثبت أهمية منصات التحليل في تقديم رؤى للأعمال تدعم تفوق المؤسسات .



مقدمة

كما يتجلى في هذا السياق دور منصات البيانات الجديدة التي يستخدم الكثير منها تقنيات (تعلم الآلة ، والتعلم الإدراكي) ولا تحتاج إلى استخدام لغات برمجة مُعقدة.

كما ينبغي الالتفات لدور تقنيات وضوابط مُتممة لا تقل أهمية عن البيانات الضخمة مثل:

(إنترنت الأشياء والذكاء الاصطناعي والحوسبة السحابية والأمن السيبراني) .



خمسة توصيات للاستفادة من البيانات الضخمة من شركة "جاردين وان سوليوشن (JOS)" المتخصصة في خدمات تكنولوجيا المعلومات :



خامساً :
البداية

ولو بخطوات صغيرة



رابعاً :

ما لا يُدرك كله لا
يُترك كله

(التوجه نحو التقنيات الحديثة مع
الاحتفاظ بالوسائل التقليدية)



ثالثاً :

تشكيل فريق
من علماء البيانات
المختصين



ثانياً :

التنقيب
عن البيانات



أولاً :

تصدر سباق
البيانات الضخمة
(تمكين)



أما بالنسبة لأدوات هذه المنصات
لتحليل البيانات الضخمة توجد العديد
منها وكل أداة لها مجال خاص فيها
وسنسلط الضوء على منصة معينة
مختصة بإنشاء المخططات لمشاركتها
عبر الإنترنت ، وهي :





Plotly عبارة عن أداة تحليلية تتيح للمستخدمين إنشاء المخططات ولوحات التحكم للمشاركة عبر الإنترنت.

مميزاته:

- سهولة تحويل أي بيانات إلى رسومات لافتة للنظر ومفيدة.
- يوفر للصناعات مراجعة المعلومات الدقيقة عن مصدر البيانات.
- تقدم Plotly استضافة غير محدودة للملفات العامة من خلال خطة المجتمع المجانية.



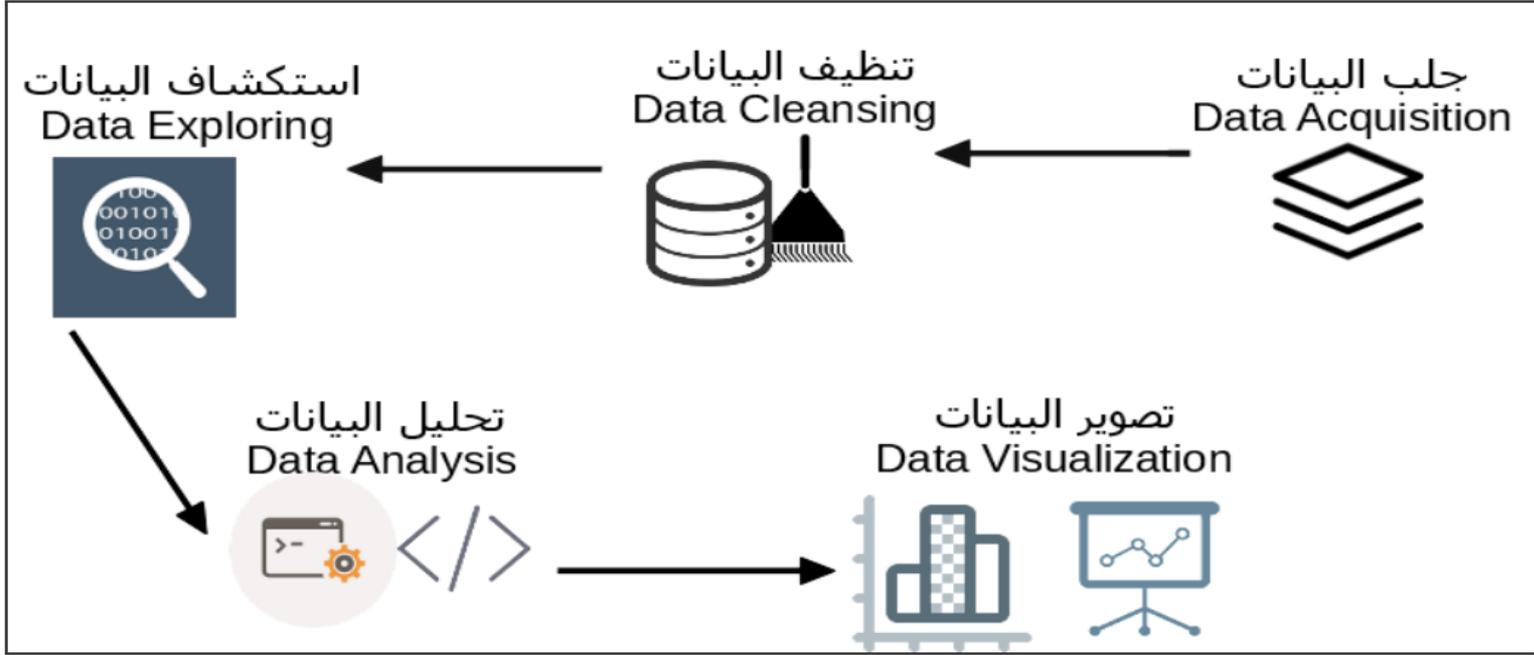
هي شركة حوسبة فنية مقرها في مونتريال ، كيبك ،
تعمل على تطوير تحليلات البيانات وأدوات التصور
عبر الإنترنت. يوفر Plotly أدوات الرسوم البيانية
والتحليلات والإحصاءات على الإنترنت للأفراد
والتعاون ، فضلاً عن مكتبات الرسوم البيانية العلمية لـ
Python و R و MATLAB و Perl و Julia و
REST و Arduino.

هي شركة لعلوم البيانات والذكاء الاصطناعي تركز
على إخراج علوم البيانات من المختبر إلى العمل.
يعمل على تسهيل إنشاء تطبيقات الويب التفاعلية
والرسوم البيانية والمرئيات ونشرها ومشاركتها بأي
لغة برمجة.



التأسيس 2012: ، مونتريال، كندا
تأسست على يد: كريس بارمر، جاك
بارمر، ماتيو سوندكويست، أليكس جونسون
نوع النشاط التجاري: شركة خاصة.

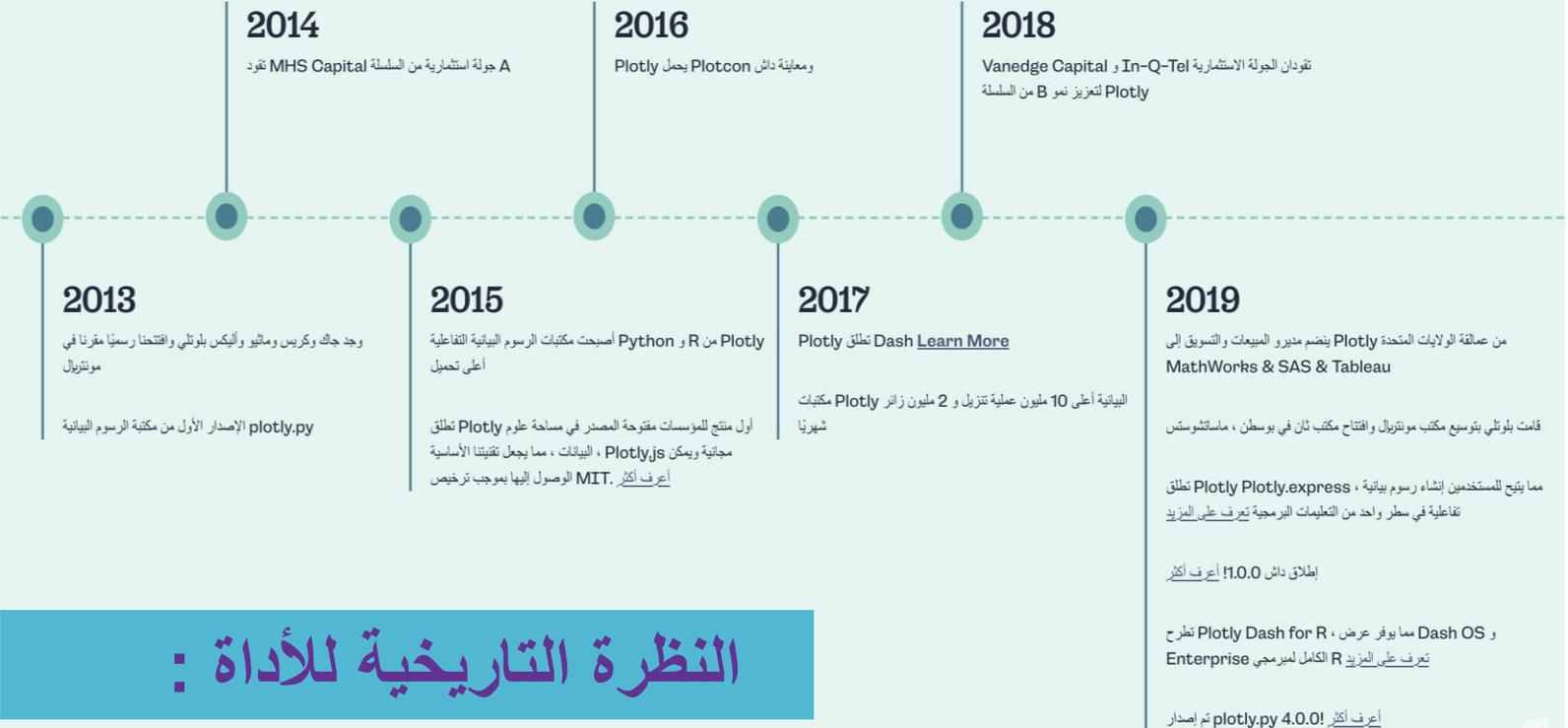




شكل 1 عملية تحليل البيانات وتصويرها تبدأ بجلب البيانات من مصادرها ثم العمل على تنظيفها وتصحيح البيانات المفقودة إن وجدت ثم باستكشاف البيانات وذلك بالتعرف على البيانات، ثم تحليل البيانات وتطبيق النماذج لاستخلاص المعلومات المفيدة منها، ثم في المرحلة الأخيرة يتم تصوير (عرض) البيانات.



مخطط زمني مؤامرة



النظرة التاريخية للأداة :



أما الشركات العالمية التي تتعامل
مع أداة منصة البيانات الضخمة (plotly) هي :

BASF

ComEd.

CISCO

Hydro
Québec

Invesco

IQT.
IN·Q·TEL

RioTinto



S&P Global

AZZ



أنواع العلاقات وأنواع التصوير البياني

:

عند استعراض البيانات ووضعها تحت مفهوم الرسم البياني أو التصوير البياني، يتم الأخذ بالحسبان معرفة نوع العلاقة (Relations) بين المتغيرات (Variables) وعلى ضوءها يتم اختيار نوع التصوير البياني .
في هذه الفقرة نستعرض بعض أشهر أنواع التصوير البياني مع التعرّيج حول بعض العلائق :



أنواع العلاقات وأنواع التصوير البياني

:

السلاسل الزمنية (SeriesTime)

وهي مجموعة من المشاهدات أو المتغيرات مرتبة وفق حدوثها في الزمن .

التصنيف والترتيب (Ranking)

يستخدم هذا النوع عند الرغبة بترتيب المتغيرات، مثل من الأعلى إلى الأدنى.

العلاقة المترابطة (Correlation)

مقارنة بين متغيرين أو أكثر، لمعرفة ما إذا كان هناك تغير سلبي أو إيجابي بينهم

الانحراف (Deviation)

مقارنة بين المتغير المنشود والمعدل الطبيعي .



أنواع العلاقات وأنواع التصوير البياني :

علاقة جزء من كل (Part-to-whole)

وهي علاقة توضح مقدار المتغير مقارنة ببقية المتغيرات .

التوزيع التكراري

(Frequency distribution)

يستخدم عادة نوعية الأعمدة الرئيسية المدرج التكراري (Histogram) .

العلاقة الجغرافية والجغرافية مكانية

(Geographic or geospatial) .



المتقدم

مثل برامج التابلو وبنتابو (Tableau/Pentaho) أو استخدام لغات برمجية مفتوحة المصدر مثل بايثون (Python) وآر (R) وتقدم مكتبات (libraries) متخصصة في تصوير البيانات أو باستخدام تقنيات الويب جافا سكريبت .

البسيط

الذي يستخدمه الغير متخصصين مثل برامج الجداول الحسابية (Calc/MS Excel/Google Sheets) .

أدوات التصوير البياني

يوجد العديد من أدوات تصوير البيانات ، منها :



أدوات التصوير البياني:

ولأن لغة بايثون مفتوحة المصدر ومدعومة من المجتمع البرمجي بتخصصات مختلفة، سنخصص هذه الفقرة بالحديث عن بعض مكتبات لغة بايثون الخاصة بالتصوير البياني :

02 **سيبورن (Seaborn)** بنيت على ماتبلوتليب لتحسين وتسهيل مهمة التصوير البياني .

01 **ماتبلوتليب (matplotlib)** وهي أشهر وأول مكتبة رسومات في لغة البايثون، والعديد من المكتبات التصويرية بنيت عليها.

03 **بلوتلي (plotly)** تمكن من إنتاج واجهات رسومية للبيانات بطريقة تفاعلية.

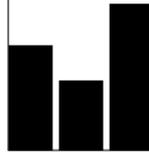


أشكال التصوير البياني (Diagrams) :

مخطط بياني (مخطط الأعمدة (Bar Chart)) :

يعتبر من أشهر الأشكال المستخدمة في عالم التصوير البياني قديماً وحديثاً، وهي عبارة عن أعمدة أفقية أو عمودية توضح التباين والاختلاف بين المتغيرات، يوضح الشكل التالي مخطط بياني (أعمدة).

كنا قد تعرضنا لبعض أنواع العلاقات في الفقرة السابقة، وسنتحدث هنا بمزيد من التفصيل حول الأشكال التي تستخدم في عرض البيانات بعد تحليلها. إن اختيار الشكل المناسب لنوعية البيانات له بالغ الأثر والأهمية في توصيل الرسالة بأفضل طريقة ممكنة :



أشكال التصوير البياني (Diagrams) :

-المخطط المتناثر Scatter Plot



يوضح القيم على شكل نقاط في داخل الرسم البياني.

-المخطط الشبكي Network chart (.



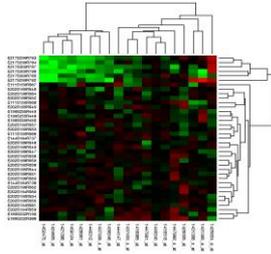
أشكال التصوير البياني (Diagrams) :

-مخطط غانت (Gantt chart)



من أقدم أنواع التصوير البياني، وتستخدم غالباً في إدارة المشاريع وتوزيع المهام على الزمن.

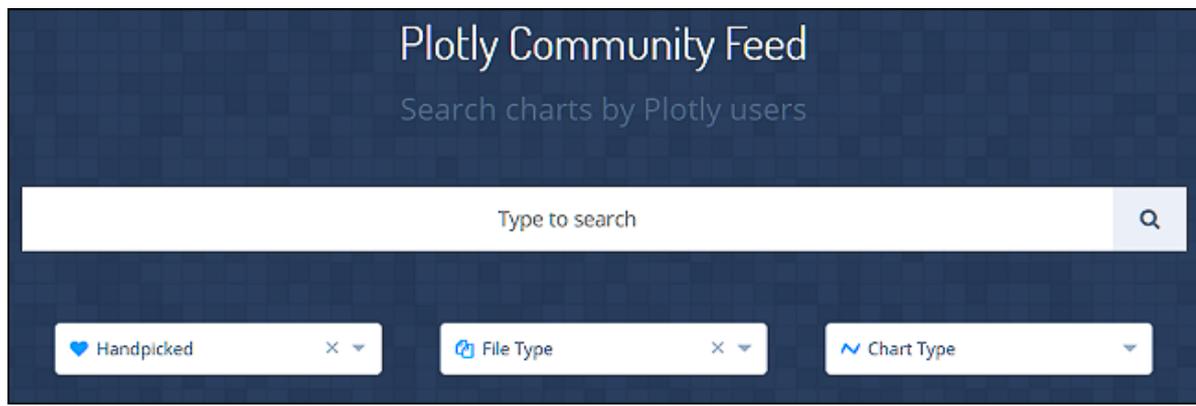
-الخريطة الحرارية (Heat map)



وهي طريقة لتمثيل مصفوفة البيانات عن طريق تمييز خلايا الرسم بالألوان و/أو كثافة اللون،

في المثال التالي استخدمت (DNA microarray) لرسم مصفوفة.





الدخول على واجهة الأداة :

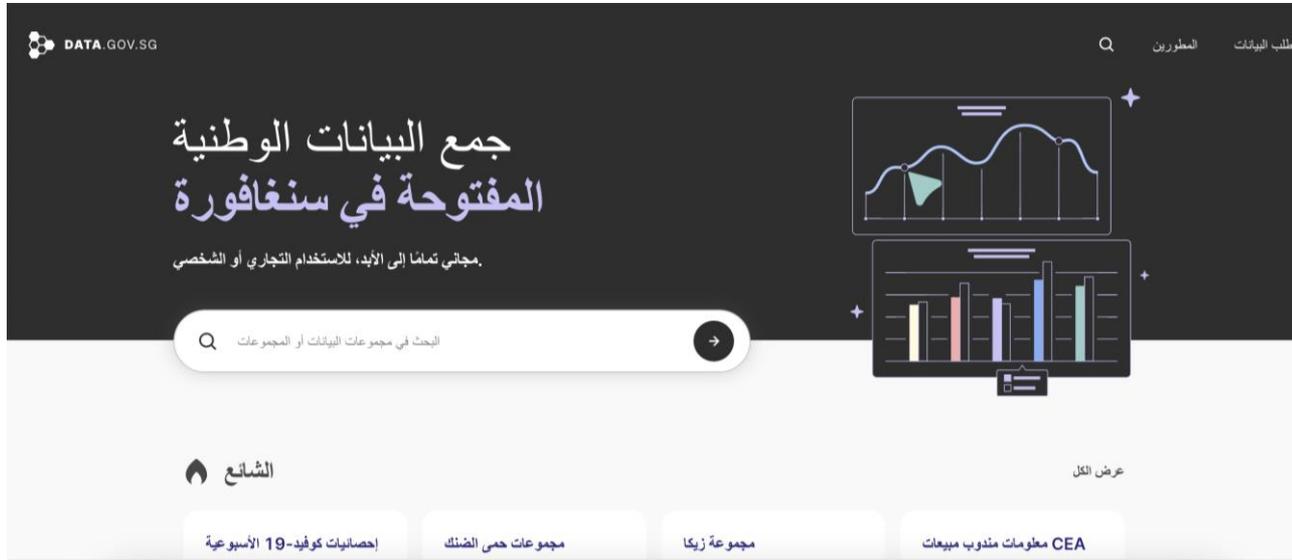
يحتاج المستخدم إنشاء حساب مجاني لتصدير العمل الذي أنجزته على الموقع، ولكن من السهل استخدام وسيط لتسجيل الدخول مثل **غوغل Google** أو **فيسبوك Facebook**. يتيح لك الحساب المجاني إنشاء ما يصل إلى **25** رسماً بيانياً ومشاركتها مع الآخرين وتصديرها كصور ذات لواحق **PNG** و **JPG** ، تضيف المستويات المدفوعة أشياء إضافية تمنحك القدرة على المشاركة الخاصة مع الفرق **Teams**، وتتيح الكثير من الرسوم البيانية، والتصدير إلى المزيد من التنسيقات ،

هذه المستويات المدفوعة مخصصة حقاً للأشخاص الذين يحتاجون إلى إنشاء الرسوم البيانية ومشاركتها بشكل احترافي.



نموذج من النماذج التي طبقت البيانات الضخمة

في منظماتها على أهم القطاعات الاستراتيجية ك (المواصلات ، التقنية ، المجتمع ، البنية التحتية ، الصحة ، المالية ، البيئة ، التعليم ، الاقتصاد) ، وهي دولة سنغافوره :



تسجيل الدخول



OPEN DATA PLATFORM
منصة البيانات المفتوحة

المجموعة المنظمات مجموعات البيانات المنصة

المملكة العربية السعودية

منصة البيانات المفتوحة

تم توفير الآلاف من مجموعات البيانات للجمهور بهدف تعزيز الوصول إلى المعلومات والتعاون والابتكار.



التحول التقني
TechTrans



وبذلك يتبين لنا ١٠ اتجاهات رئيسية ستغير صناعة تحليلات البيانات الكبيرة :

١. مقدمي الخدمات السحابية العامة يوسعون نفوذهم
٢. تستمر مزايا السحابة العامة على السحب الخاصة في الاتساع .
٣. تنامي السحابة المختلطة أو الهجينة .
٤. تسريع وقت المؤسسة إلى القيمة .
٥. جلب تطبيقات متطورة إلى السوق .
٦. أصبحت نهج البيانات الضخمة الكبيرة بدائل قابلة للتطبيق لمنصات ثابتة .
٧. أصبحت تقنية Hadoop مجرد قطعة في لغز البيانات الكبيرة .
٨. يقوم المستخدمون على نحو متزايد بخلط ومطابقة عمليات نشر البيانات الكبيرة متعددة البائعين في النظم الإيكولوجية المفتوحة .
٩. يجري تفكيك قواعد البيانات وإعادة تجميعها في نهج مبتكر .
١٠. أصبحت تطبيقات تحليل البيانات الكبيرة المعبأة متاحة على نطاق أوسع .



الخاتمة

- تمثل البيانات الضخمة مرحلة هامة من مراحل تطور نظم المعلومات والإتصالات
- تعبر في مفهومها المبسط عن كمية هائلة من البيانات المعقدة التي يفوق حجمها قدرة البرمجيات والآليات الحاسوبية التقليدية على تخزينها ومعالجتها وتوزيعها، الأمر الذي أدى إلى وضع حلول بديلة متطورة تمكن من التحكم في تدفقها والسيطرة عليها.
- تمتلك تقنية البيانات الضخمة إمكانية تحليل بيانات مواقع الإنترنت وأجهزة الإستشعار وبيانات شبكات التواصل الإجتماعي.
- تحليل هذه البيانات يسمح بوجود ارتباطات بين مجموعة من البيانات المستقلة لكشف جوانب عديدة، ومنها على سبيل المثال التنبؤ للإتجاهات التجارية للشركات ومُكافحة الجريمة في المجال الأمني وغيرها.
- توفر هذه التنبؤات لصانعي القرار أدوات مبتكرة لفهم أفضل للظروف وبالتالي اتخاذ قرارات صحيحة تحقق الأهداف المطلوبة.



المراجع

- <https://data.gov.sg/> .١
- <https://01gov.com/change-big-data-into-real-data/> .٢
- <https://units.imamu.edu.sa/rcentres/en/asa/Pages/Introduction-to-Data-visualization.aspx> .٣
- <https://plotly.com/products/oem/> .٤
- <https://syriantech.com/2018/08/create-graph-web/> .٥
- <https://open.data.gov.sa/ar/home> .٦



ان أصبت فمن الله
وإن أخطأت فأسأل الله لي العفو والغفران

شكراً لكم
Thank you

شذا إبراهيم الخراع



Webinar

التحول التقني

سلسلة من الندوات المباشرة عبر الإنترنت، يقدمها نخبة من الخبراء والمتخصصين. بهدف المساهمة في رفع الوعي التقني لدى كافة أفراد المجتمع.



لمشاهدة محاضرات
ويبينار التحول التقني

