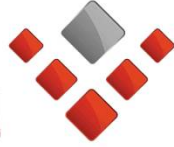




دار المنظومة  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



نسيج Naseej



# تقنيات البيانات الضخمة وأسس توظيفها لتطوير خدمات مؤسسات المعلومات

أ.د. أحمد فرج أحمد

أستاذ ورئيس قسم المكتبات والوثائق والمعلومات

[ahmed.farag@aun.edu.eg](mailto:ahmed.farag@aun.edu.eg)

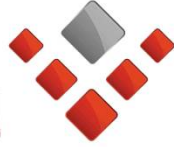
المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات

"نورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"

بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



دار المنظومة  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



نسيج Naseej



الأجندة

مصادر البيانات الضخمة في مؤسسات المعلومات

إشكاليات توظيف البيانات الضخمة في مؤسسات المعلومات

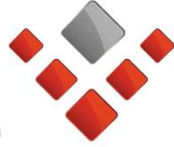
خصائص البيانات الضخمة والبنية الهيكلية لهاذوب

نماذج من مبادرات البيانات الضخمة في مؤسسات المعلومات

التوظيف لتحسين البحث وتخصيص المحتوى والخدمات.



دار المنظومة  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



نسيج Naseej



## مصادر البيانات الضخمة في مؤسسات المعلومات

النشاطات والأحداث والفعاليات والخدمات

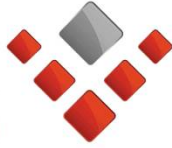
بوابات ومواقع مؤسسات المعلومات وصفحاتها على شبكات التواصل الاجتماعي

المعلومات الناتجة عن دراسات سلوكيات وممارسات المستخدمين المعلوماتية.





دار المنظومة  
DAR ALMANDUMAH  
البراد في قواعد المعلومات العربية



نسيج Naseej



## مصادر البيانات الضخمة في مؤسسات المعلومات

البيانات والمعلومات الناتجة عن خدمات المعلومات الرقمية بمؤسسات المعلومات

استخدام الهواتف الذكية، والمساعدات الرقمية الصوتية DVA، وخدمات تحديد المواقع GPS،  
والحوسبة السحابية، والروبوتات، وانترنت الأشياء IoT وغيرها...

المستشعرات "Sensors" في أنظمة التعريف بترددات الراديو "RFID" و "I beacon" وغيرها  
من تقنيات انترنت الأشياء

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"نورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



دار المنظومة  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية

نسيج Naseej

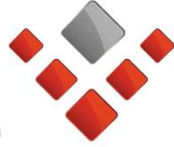


# إشكاليات توظيف البيانات الضخمة في مؤسسات المعلومات

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



**دار المنظومة**  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



**نسيج** Naseej



**القلق** من سرعة نمو المصادر والمقتنيات بشكل يؤثر على فعالية الخدمات

**الخوف** من عدم استيعاب تدفق البيانات الضخمة والسيطرة عليها

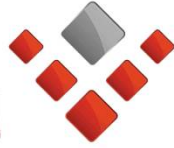
**تحديات** الأنظمة التقليدية المستخدمة في استخراج وتحليل البيانات

**معوقات** النسخ الاحتياطي لحجم ضخم جدا من البيانات

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



دار المنظومة  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



نسيج Naseej



قضايا البحث والاسترجاع

تهديدات تواجه خصوصية بيانات المستخدمين

ندرة العنصر القادر على إدارة متطلبات ومنصات البيانات الضخمة.

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"نورة البيانات وتأثيرها على مؤسسات المعلومات العربية : بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



**دار المنظومة**  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية

**نسيج** Naseej



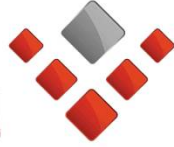
# خصائص البيانات الضخمة

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية : بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م





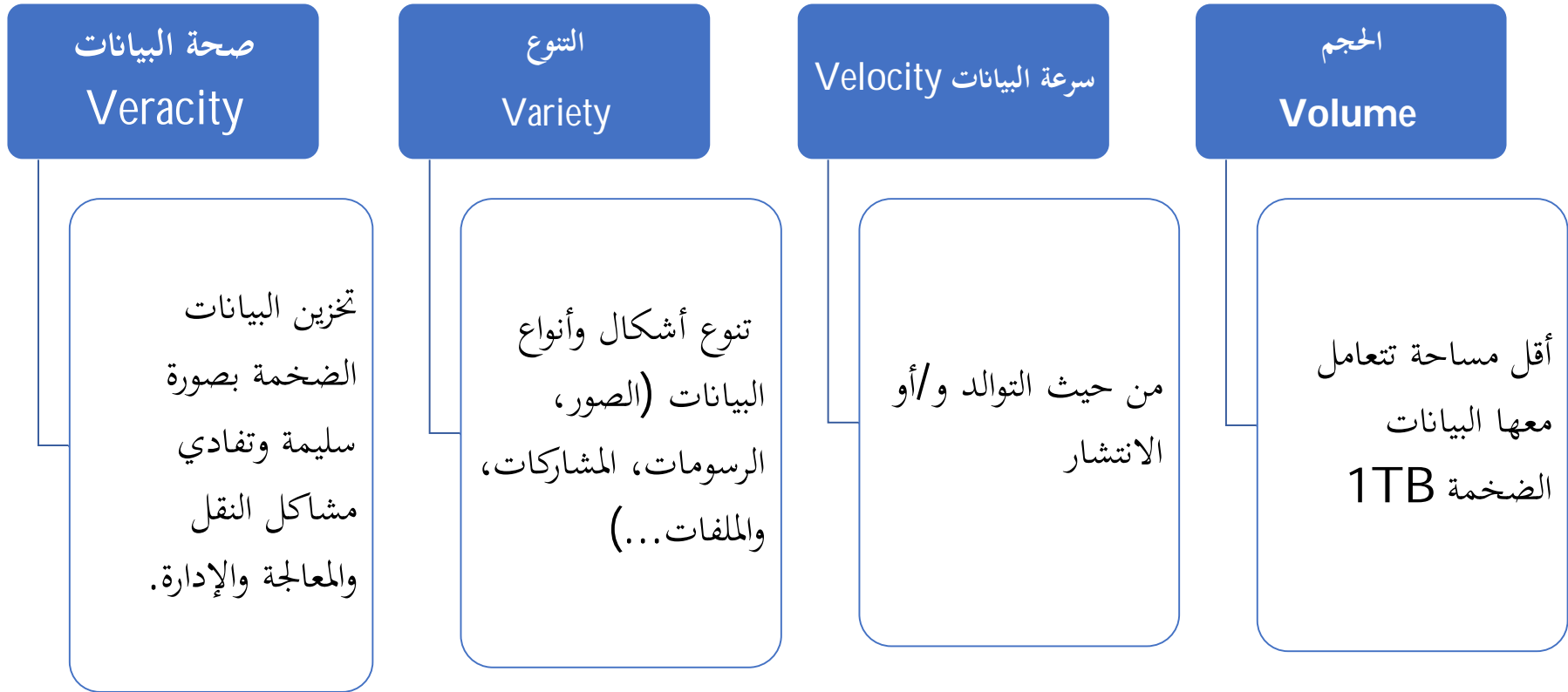
دار المنظومة  
DAR ALMANDUMAH  
المراد في قواعد المعلومات العربية



نسيج Naseej



## نموذج Vs 10 لخصائص البيانات الضخمة



المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات

"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"

بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م

## التأثير العالي Vulnerability

إثارة البيانات الضخمة  
لمخاوف أمنية واحتماليات  
الخرق، والتعدي على  
خصوصية البيانات  
الشخصية.

## العرض المرئي Visualization

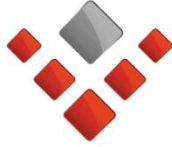
الاستفادة من أنماط عرض  
البيانات وفق أشكال وأنماط  
متنوعة.

## المتغيرات Variability

التوافق مع التغيرات التي قد  
تطرأ على النظام، فقد تحدث  
مشكلات يتطلب تخطيطها  
مجهودات برمجية تستلزم وقت  
وجهد وتكلفة.



دار المنظومة  
DAR ALMANDUMAH  
البراد في قواعد المعلومات العربية



Naseej نسيج



## التقلب Volatility

الفترة الزمنية للاحتفاظ  
بالبيانات، وتوقيتات  
تقادمها، ووضع قواعد  
حاكمة لتحديث البيانات  
ومدى توفرها، والاسترجاع  
السرير عند الحاجة.

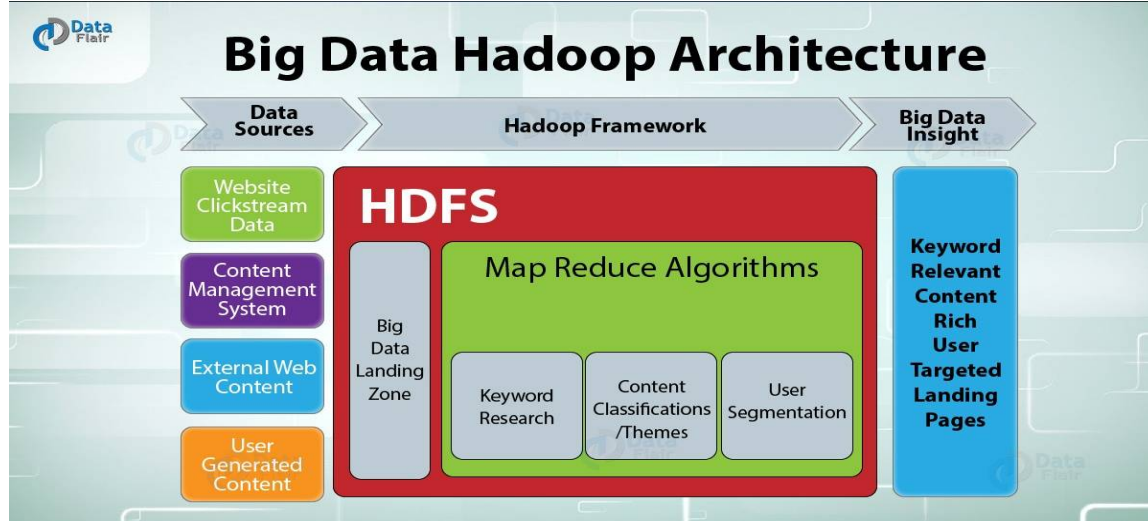
## الصلاحية Validity

دقة التحليل والاستفادة  
من تحليلات البيانات  
الضخمة واعتماد  
ممارسات فعالة لضمان  
الجودة والاتساق

## القيمة Value

استثمار البيانات  
الضخمة لتحقيق الأرباح  
وتعزيز الاقتصاد المعرفي  
و ضمان استمرارية تطوير  
الخدمات والمنتجات.

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"نورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



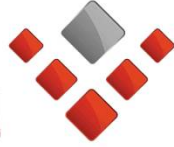
# البنية الهيكلية لهادوب

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق  
"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية : بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 1-5U 3 مارس 2022م

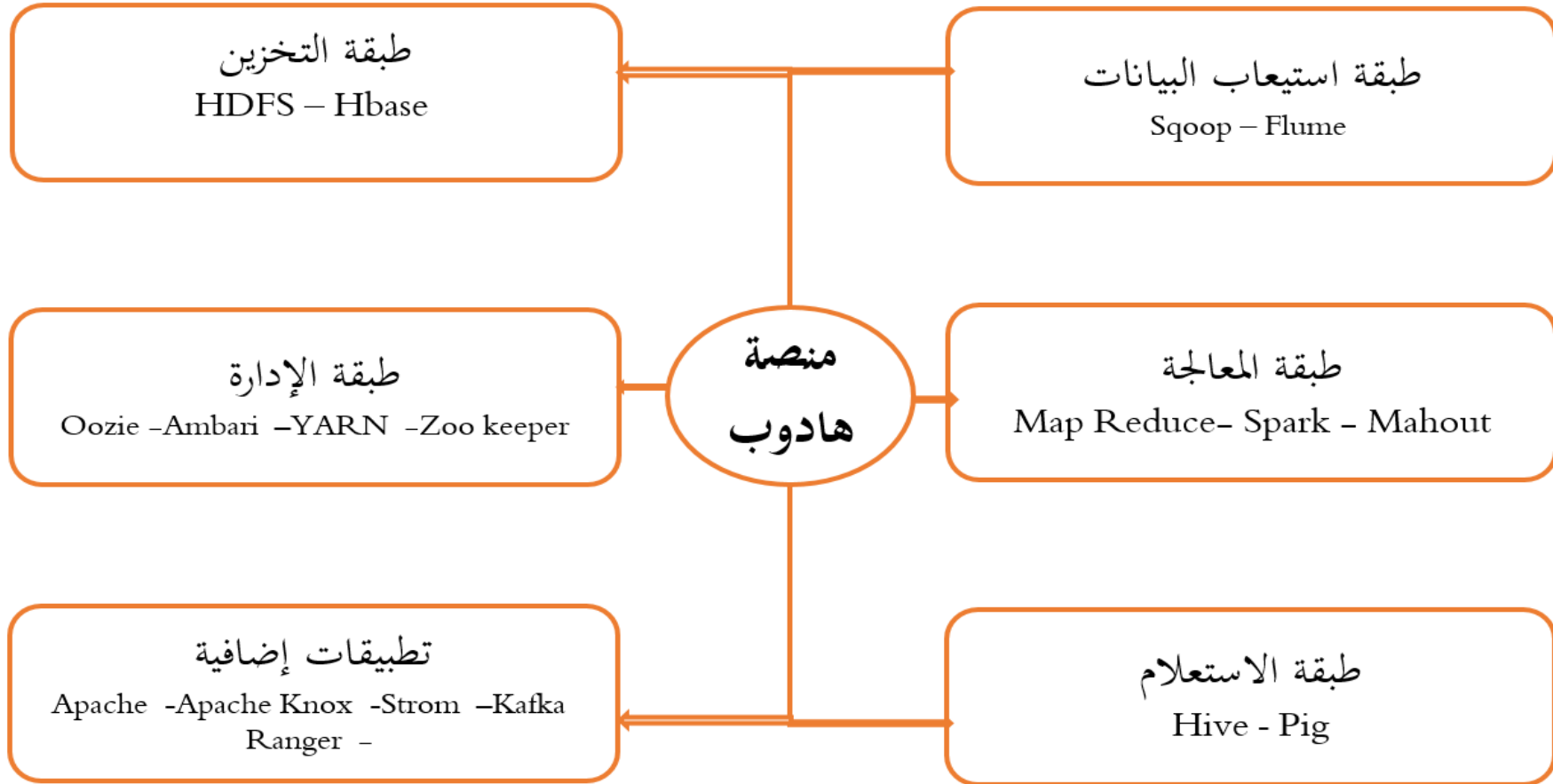




دار المنظومة  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



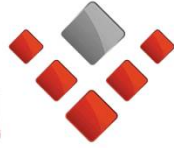
Naseej نسيج



المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"نورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



**دار المنظومة**  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



**نسيج** Naseej



# نماذج من مبادرات البيانات الضخمة في مؤسسات المعلومات

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية : بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



University of Michigan Library  
Open Access Bibliographic  
Records

Canada's Libraries and  
Archives

National Science Digital  
Library NSDL

UCL Research Data Repository

Twitter Archive at the Library  
of Congress

Harvard University Library: Big  
Data Applications for Books

Harvard Library. "Harvard Library  
Open Metadata

Creation of a Metadatabase for  
Geophysical Data in Australia

JISC & HESA Library Data Labs  
Project

Brooklyn Public Library (BPL): Big  
Data for the Visualization of User  
Data

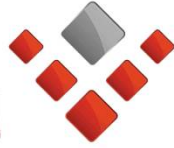
المؤتمر العلمي الثاني عشر لقسم المكتبات والو

"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع والتحديات"

بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



**دار المنظومة**  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



**نسيج** Naseej



# سبل توظيف البيانات الضخمة لتحسين البحث وتخصيص المحتوى والخدمات

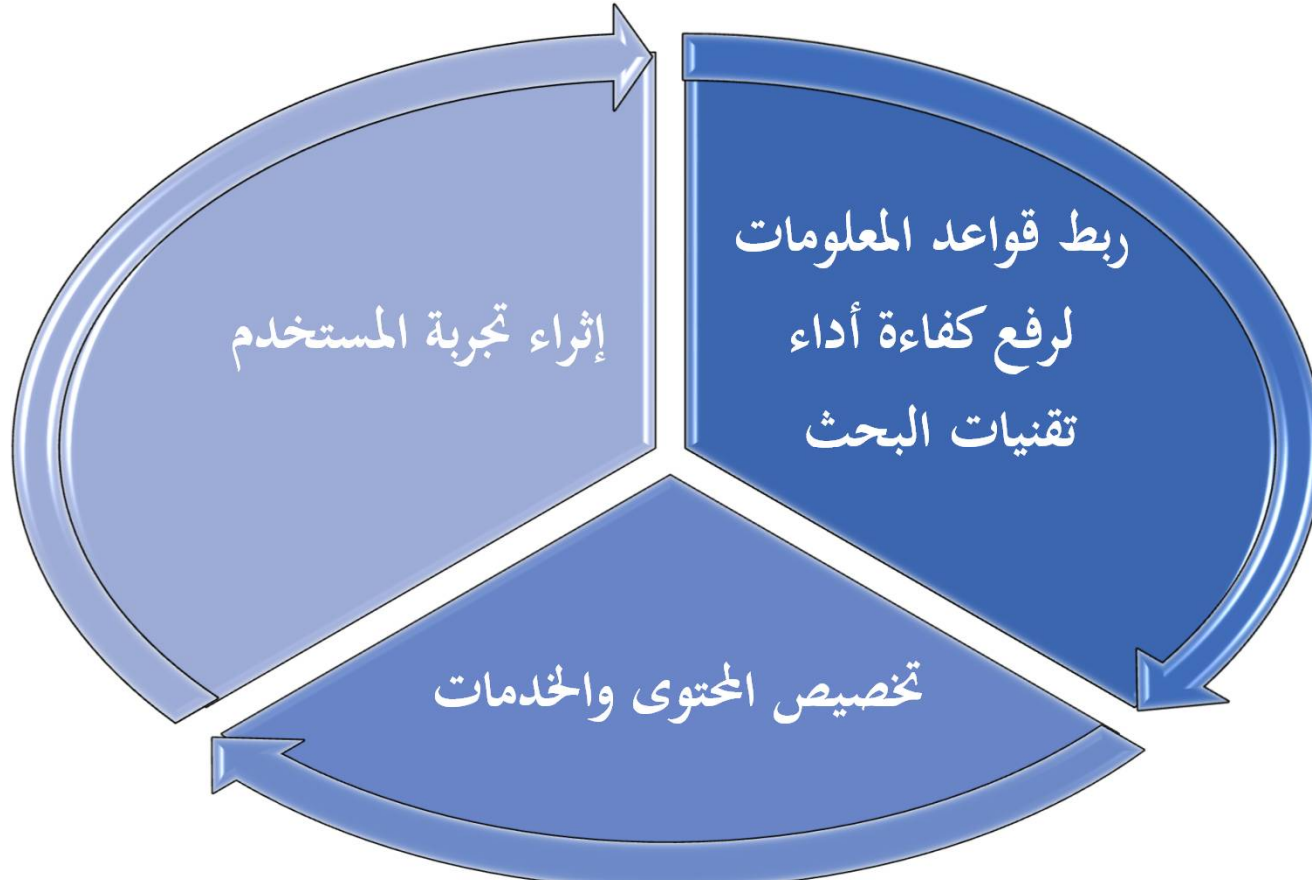
المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية : بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022 م





**دار المنظومة**  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية

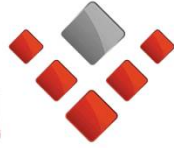
**نسيج** Naseej



المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



دار المنظومة  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



نسيج Naseej



## ربط قواعد المعلومات لرفع كفاءة أداء تقنيات البحث

نموذج : توظيف البيانات الضخمة في آلية عمل أنظمة البحث التجميعي

Innovative  
Interfaces  
Encore  
Synergy

Ex Libris  
Primo  
Central

OCLC  
WorldCat  
Local

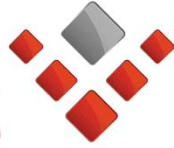
Serials  
Solutions  
Summon

EBSCO  
Discovery  
Service

المؤتمر العلمي الثاني عشر لتسليم المكتبات والوثائق وطنية المعلومات  
"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية : بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



دار المنظومة  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



نسيج Naseej



## ربط قواعد المعلومات لرفع كفاءة أداء تقنيات البحث

يمكن لمؤسسات المعلومات تصور آلية الربط مع قواعد المعلومات من خلال:

تجميع نتائج الاستعلامات،  
وتنسيق عرضها من خلال القرص  
وحذف المكررات...

بروتوكولات وتراخيص استخدام مع  
الناشرين

إتاحة نتائج صادرة من عدة  
قواعد في نفس الوقت

إجراء البحث بشكل تزامني  
في كل قاعدة معلومات

وتتولى أنظمة البيانات الضخمة على عاتقها:

إدارة وصيانة التسجيلات الواصفة

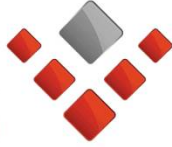
إدارة الحقوق والصلاحيات

إدارة خيارات التخصيص لتحكم  
مؤسسة المعلومات في النظام.

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"نورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



**دار المنظومة**  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



**نسيج** Naseej



## مميزات توظيف أنظمة البيانات الضخمة لربط قواعد معلومات الناشرين

**إتاحة الوصول إلى قواعد معلومات موثوقة وعالية الأداء.**

**تخصيص توصيف البيانات،** يمكن للمستخدمين ومؤسسات إدارة المحتوى تحديد واستخدام مخططات البيانات لوصف مجموعات المصادر.

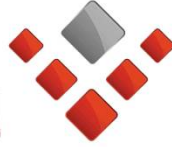
**التحكم في أساليب الوصول إلى المصادر الرقمية.** (مصادر مقيدة الوصول والاتاحة، مصادر متاحة حصرياً لفئة معينة من المستخدمين أو مشاركتها بشكل عام).

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"نورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م





**دار المنظومة**  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



**نسيج** Naseej



**المرونة** في تدفقات البيانات الموجهة للمستخدم

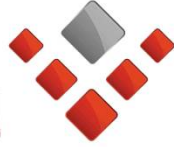
**اختيار** المعرف الثابت "Digital Object Identifier" DOI ، حيث ترتبط المصادر بمعرف ثابت وفريد.

**رفع** أداء إجراءات البحث والاسترجاع بفضل وصف المصادر بمجموعة من البيانات الوصفية القياسية والمخصصة والمحددة المجال

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات  
"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية : بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



دار المنظومة  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



نسيج Naseej



## البيانات الضخمة ودورها في تخصيص المحتوى والخدمات

يتطلب التخصيص بصفة عامة:

تطبيق آليات تعلم  
الآلة للتعرف على  
مفضلات  
المستخدمين  
واستخلاص المحتوى

تحليل سلوكيات  
وتوجهات  
المستخدمين

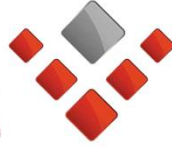
تحليل سجلات  
البحث " Search  
"Logs"

الحصول على  
البيانات الشخصية  
للمستخدمين وتجميع  
أكبر قدر ممكن منها

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات

"نورة البيانات وتأثيرها على مؤسسات المعلومات العربية : بين الواقع وطموحات المستقبل"

بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022 م



مصادر تجميع البيانات حول المستخدم والمحتوى

الملفات الشخصية للمستخدمين

البيانات الشخصية (الموقع الحالي والمؤهل الدراسي والتخصص العام والدقيق والعمر والجنس...)

البيانات المهنية مثل: المؤسسة التي يعمل بها، سنوات الخبرة، والإدارة والوحدة والقسم الإداري التابع له، والمسمى الوظيفي، وما إلى ذلك.

طلبات البحث

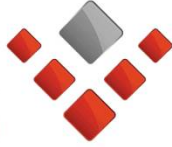
الملفات التي يتم تنزيلها

الكلمات الدالة والمفتاحية المستخدمة من قبل المستخدمين

البيانات الوصفية وإعداد التصنيف Taxonomy وغيرها.

"ثورة البيانات وتأثيرها على مؤسسات المعلومات العربية: بين الواقع وطموحات المستقبل"

بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



مصادر تجميع البيانات حول المستخدم والمحتوى

معالجة وتحليل بيانات سجل المستخدم والمحتوى داخل هادوب.

تغذية محركات البحث بالمحتوى وتوصيف البيانات وآليات تقديم النتائج عبر واجهة سواء كانت موقع ويب أو تطبيق هاتف ذكي.

خوارزميات تعلم الآلة لإجراء تحليلات متقاطعة Cross analytics بين عادات تصفح المستخدمين (البيانات غير المهيكلة) والمحتوى (البيانات المهيكلة) لتقديم تجربة بحث وتصفح مخصصة.

تسجيل نشاطات المستخدمين على بوابة مؤسسة المعلومات

تسجيل نشاطات المستخدمين على مواقع وصفحات الويب وشبكات التواصل الاجتماعي

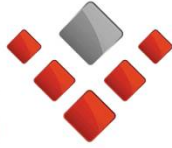
معرفة عضوية المستخدم في مجموعات الاهتمامات الموضوعية إن وجدت، والتعرف على اهتمامات ونشاطات هذه المجموعات

استقبل

بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



دار المنظومة  
DAR ALMANDUMAH  
الرواد في قواعد المعلومات العربية



نسيج Naseej



## البيانات الضخمة وتحسين تجربة المستخدم

مع انتشار استخدام تقنيات الكتب الرقمية وأجهزة القراءة الإلكترونية وتطبيقات الهواتف الذكية، عملت العديد من مؤسسات المعلومات على تبني آليات لدعم تجارب المستخدمين " UE User Experiences

المؤتمر العلمي الثاني عشر لقسم المكتبات والوثائق وتقنية المعلومات

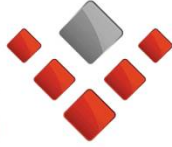
"نورة البيانات وتأثيرها على مؤسسات المعلومات العربية : بين الواقع وطموحات المستقبل"

بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م





دار المنظومة  
DAR ALMANDUMAH  
البراد في قواعد المعلومات العربية



نسيج Naseej



## تمتع البيانات الضخمة بـ:

استخدام نتائج التحليل  
لأغراض تحسين تجربة  
المستخدم وتحقيق الرضا  
وتعزيز العائد المادي من  
خدمات المعلومات  
المقدمة.

تخزين البيانات في  
مستودع على نطاق  
واسع ثم إجراء تحليل  
البيانات

إمكانية تتبع وتسجيل  
نشاطات مستخدمي  
المصادر والخدمات

القدرة على استخراج  
معلومات حول  
مجموعات مصادر  
المكتبات وغيرها من  
مؤسسات المعلومات.

"نورة البيانات وتأثيرها على مؤسسات المعلومات العربية : بين الواقع وطموحات المستقبل"  
بالمكتبة المركزية الجديدة جامعة القاهرة 30-31 مارس 2022م



دار المنظومة  
DAR ALMANDUMAH  
السرور في قواعد المعلومات العربية

نسيج Naseej



THANK YOU