



## قاعدة بيانات

قاعدة بيانات أرشيف أدب السجون - مصر 2011-2020

الدليل والمنهجية

مخزن البيانات الحرة

Open Data Tank Initiative e.V

## أولاً: تقديم مختصر للدليل:

تضم قاعدة بيانات أرشيف أدب السجناء المنتجات الإبداعية التي تم رصدها وصدرت من داخل السجناء للمحبوسين على خلفية سياسية وقضايا تخص المجال العام في مصر خلال الفترة من (عام 2011 حتى 31 مارس 2019)، حيث تم استبعاد الرسائل والمنتجات الإبداعية التي صدرت قبل أو بعد تلك الفترة ، بغض النظر عن توقيت القبض أو القضايا المحبوس على ذمتها صاحب المنتج الإبداعي التي قد تكون قبل تلك الفترة.

تعتبر قاعدة البيانات مفتوحة المصدر ومنشورة برخصة المشاع الإبداعي وبحيث تشمل بعض التفاصيل الشخصية والقانونية والإجرائية فيما لا يتعدى انتهاك الخصوصية، حيث تضم بيانات قانونية (الأرقام القضائية وبعض الاتهامات الموجهة والأحراز والتلفيات إن تم التوصل إليها) وبيانات واقعة القبض أو الأتهام (تاريخ الواقعة) وبيانات شخصية (الاسم، السن، الوظيفة).

من خلال قاعدة البيانات، تم إعداد إحصاءات وصفية بين عديد من المتغيرات مثل (الرسائل النصية والإقليم الجغرافي، نوع المنتج الإبداعي، نوع المنتج الإبداعي والنوع الاجتماعي، الوضع القانوني والفئة العمرية لصاحب المنتج الإبداعي).

المنهجية التي تم الاعتماد عليها في المصادر هي **Methodology of Triangulation and Data Verification** المعروفة في مجال العلوم الاجتماعية حيث تنوعت المصادر التي تم إستهدافها والتي تم الإعتماد بشكل أساسي على مواقع التواصل الاجتماعي (فيس بوك، تويتر) نظراً لكونهم منصات نشر يعتمد عليها المحبوسون وذويهم في نشر الإنتاج الإبداعية، إلى جانب البحث بشكل عبر محرك البحث جوجل **Google** لتغطيته لأكثر من مصدر.

### ثانياً: محتويات الإصدار:

- ملف قاعدة البيانات التفصيلي على هيئة xlsx والتي تحتوي على الإصدار المعلوماتي: نسخة كاملة تحتوي على قاعدة البيانات وورقة sheet لتمثيل البيانات إحصائياً.
- صور screenshots لجداول الإحصائيات.
- ملف pdf عن الدليل والمنهجية.
- ملف Pdf عن ورقة تقييم الفجوات المعلوماتية.

## ثالثاً: معايير إدراج واستبعاد الوقائع:

- **المعيار الزمني:** اشترط أن يكون المنتج الإبداعي متصل بخلفية سياسية أو بالمجال العام المصري، وأن يكون صدر خلال الفترة الزمنية من عام 2011 حتى 31 مارس 2019 بغض النظر عن التفاصيل الزمنية للواقعة المحبوس على إثرها صاحب المنتج الإبداعي الذي قد يكون قبل تلك الفترة المستهدفة.
- **المعيار الجغرافي:** يشترط أن يكون المنتج الإبداعي قد تم صدر من أماكن احتجاز داخل حدود جمهورية مصر العربية، بغض النظر عن سبب الاحتجاز الذي قد يكون متصل بمكان جغرافي خارج جمهورية مصر العربية.
- **المعيار الشرطي:** أن يكون المنتج الإبداعي صدر تم صناعته أو كتابته داخل مكان الاحتجاز خلال الفترة الزمنية للاحتجاز.
- **المنتجات المستبعدة:** تم استبعاد الأحاديث الصحفية والمنتجات المصنوعة من محبوسين وتم انتاجها بعد خروجهم من محبسهم، وكذلك تم استبعاد الإنتاجات التي كانت من أشخاص مُطلق سراحهم لأشخاص محبوسين حيث أن شرط إدراج المنتج كان أن يكون صادر من داخل أماكن احتجاز إلى خارجها وليس العكس، وكذلك تم استبعاد الرسائل الخاصة الغير منشورة للعامه.

## رابعاً: معايير ووحدة التعداد الإحصائي:

- الأعداد الواردة في الجداول الإحصائية لا تُمثل إجمالي عدد المحبوسين، وذلك لتكرار الإنتاجات الإبداعية الصادرة من نفس الشخص، وتعتبر وحدة التعداد الإحصائي هي صدور "منتج إبداعي من شخص معين في تاريخ معين"، وفي حالة تغير أياً من المتغيرين السابقين (المنتج الإبداعي، أو تاريخ التحرير أو النشر) يُحتسب كمنتج جديد.
- الغرض الأساسي من البناء المعلوماتي والإحصائي هو رسم الأنماط والسلوكيات العامة الخاصة بالإصدارات والمنتجات الإبداعية الصادرة من المحبوسين على خلفية سياسية في أماكن احتجاز مصرية.
- الأعداد داخل الجداول الإحصائية في الإنتاجات الإبداعية قد تُمثل "نسبة" مما يتم إصداره على أرض الواقع، ولكن يمكن اعتبارها أنها نفس النسبة "تقريباً" في مختلف النطاقات الزمانية والمكانية والظرفية، وذلك بسبب الاستعانة بمختلف المصادر على مستوى الجمهورية مع مراعاة عامل "المركزية الشديدة للمعلومات" عبر التنوع والتركيز على المصادر الغير مركزية والمحلية المختلفة.
- العدد صفر لا يعني عدم وجود إنتاجات إبداعية مطلقاً أو واقعياً، ولكنها تعني عدم التوصل إلى إنتاجات حسب نفس آلية جمع المعلومات والمنهجية ومعايير إدراج الحالات.

## خامساً: آلية جمع المعلومات والمصادر:

### • آلية الاستعانة بالمصادر:

- تم الإعتماد بشكل أساسي في المصادر المستهدفة على مواقع التواصل الإجتماعي حيث تعتبر منصات إعلامية للمحبوسين على خلفية سياسية من كافة التيارات وكذلك لدويهم لنشر الإنتاجات الإبداعية التي تصدر من مختلف أماكن الاحتجاز ولذلك تم تحديد موقعي (فيس بوك، تويتر) بشكل أساسي، ثم التركيز على جمع الإنتاجات الإبداعية وتكملة تفاصيل الوقائع من خلال محرك البحث جوجل **Google** والمواقع الإخبارية والحقوقية التي تظهر فيه بشكل أساسي.

### • مصادر المعلومات:

- مصادر المعلومات التي تم الاستعانة بها في البناء المعلوماتي لقاعدة البيانات الأولية بشكل رئيسي هي "مصادر أولية" (معلومات تم التحقق منها لوجود المصدر الأساسي لنشأتها)، بينما تمت الاستعانة بالمصادر الثانوية كمصادر تكميلية لبيانات الواقعة وباقي التفاصيل الشخصية أو القانونية إن تم التوصل إليها.

- تم وضع رابط مباشر لتلك المصادر تفصيلاً حيث أنه لكل منتج إبداعي في قاعدة البيانات، مع مراعاة أنه يمكن أن تتلف بعض الروابط بتقادم الفترات الزمنية ومحو المحتوى المعلوماتي على الأنترنت.

### • البيانات الوصفية أو الفوقية Metadata لكل منتج إبداعي:

- هي بيانات جديدة تصف البيانات الأساسية مثل بناء متغيرات "الفئة الوظيفية" لعمود "الوظيفة"، وعواميد "سنة ونصف سنة والعهد الرئاسي لتحرير/نشر المنتج الإبداعي" لعمود "تاريخ تحرير/نشر المنتج الإبداعي"، وكذلك عمود "الإقليم الجغرافي" لعمود "محافظة الواقعة".

- البيانات الوصفية هي أدق ما تم التوصل إليه من جميع المعلومات المتاحة وفقاً لقواعد إحصائية في مجال العلوم الاجتماعية، ولكنها ليست بالضرورة كاملة الدقة (100% هي نسبة مطلقة غير واقعية في الأبحاث الاجتماعية) حيث أنه هناك معايير فرضيات واستنتاجات تم الاستعانة بها

- تم تمييز أعمدة البيانات الوصفية **Metadata** داخل قاعدة البيانات بالخلفية ذات اللون الأحمر.

## سادساً: منهجية التحقق من البيانات وعوامل تقييم المعلومات:

تم استخدام منهجية التحقق Methodology of Triangulation and Data Verification المعروفة في مجال العلوم الاجتماعية، والتي يتم فيها الاعتماد على مصادر مختلفة لنفس المعلومة أو مجموعة المعلومات، ويتم خلال تلك العملية تقييم المعلومات الواردة من مصادر مختلفة وفقاً لعدة عوامل للاستبعاد أو التعديل أو الدمج أو الإضافة.

### - مراحل تفكيك المحتوى المعلوماتي والتعامل معه:

- **1. الفهم:** بقراءة وفهم السياق وطبيعة تركيب البيانات ومصدر المعلومات والمحتوى المعلوماتي الموجود كاملاً واستيعابه مع تجاوز المصطلحات المستخدمة أو أية أخطاء أو أغراض بشرية.
- **2. التقسيم والتصنيف:** بتفكيك وتحديد المحتوى المعلوماتي وتحديد وحدة التعداد الإحصائي وإدراج كل معلومة في خاناتها بقاعدة البيانات.
- **3. بناء تصنيفات للبيانات (البيانات الوصفية):** حيث يتم فهرسة البيانات وتكويدها بحيث تكون أدق ما تم التوصل إليه من جميع المعلومات المتاحة وفقاً لقواعد إحصائية في مجال العلوم الاجتماعية.
- **4. معالجة الفجوات المعلوماتية:** لأي من المتغيرات بإعادة بناء مستويات أخرى لمصادر المعلومات.
- **5. الاستنتاج:** باستنتاج معلومات جديدة صحيحة 100%، ولكنها غير موجودة في المحتوى المعلوماتي الحالي (مثل استنتاج نوع كل واقعة من خلال قراءة نص الخبر أو تحليل الواقعة).
- **6. الفرضية:** بافتراض معلومات بنسبة كبيرة قد تكون صحيحة، ولكنها غير موجودة في المحتوى المعلوماتي الحالي، وذلك وفقاً لمعايير محددة أكاديمياً وخبرات ورؤية عامة للباحث.

### عوامل تقييم المعلومات:

- حسب نوع الواقعة ومكانها وزمانها.
- حسب العناصر المتداخلة في الواقعة.
- حسب كم التفاصيل المتوفرة للمصدر.
- حسب تطابق التفاصيل الإطارية لمعلومات المصدر مع الحقيقة.
- حسب كم التضارب والأخطاء بين التفاصيل الداخلية لمعلومات المصدر.
- حسب منطقية ومنهجية التفاصيل الداخلية لمعلومات المصدر.
- حسب تقادم المعلومات.

## سابعًا: الإطار المفاهيمي وتصنيف البيانات:

- تعريف مصطلح "إنتاج إبداعي داخل مكان احتجاز" في قاعدة البيانات:
  - مصطلح إنتاج إبداعي داخل مكان احتجاز يمثل أي شكل من أشكال الجهود الذهنية أو الأشكال الإبتكارية سواء كانت بشكل أدبي أو مادي وتم صناعتها أو كتابتها أو إنشائها داخل مكان من أماكن الاحتجاز سواء كانت رسمية أو غير رسمية ويمكن أن تكون (رسماً أو منتج مصنوع باليد أو رسالة نصية،....).
  - تضم قاعدة البيانات كل منتج إبداعي تم رصده خلال الفترة الزمنية ( عام 2011 حتى 31 مارس 2018 ) .
  - مصطلح "عدد الإنتاج الإبداعية" المُستخدم في الجداول الإحصائية يُمثل بالفعل عدد الإنتاج التي تم رصدها مع مراعاة احتساب كل منتج إبداعي على حدة حتى إذا تكررت الإنتاج الإبداعية لنفس الشخص.

### • معايير تقسيم المعلومات:

#### أ. فيما يخص بيانات المنتج الإبداعي:

- \*\* تم تقسيم الإنتاج الإبداعية كما يلي:
1. رسالة نصية: وهي المنتجات النصية التي خرجت على شكل رسائل عادية أو خواطر أو مقالات أو كتابات محدودة بغرض النشر عبر منصات إعلامية أو وسائل تواصل إجتماعي.
  2. رسم: وهي الرسومات التي كانت تهدف إلى توصيل مغزى معين عبر الرسم.
  3. تصميم وإهداء: وهي الرسومات التي كان يتم إهدائها إلى أشخاص معينة.
  4. صنّع في السجن: المنتجات التي كانت عبارة إنتاجات صنّعت بشكل يدوي داخل أماكن احتجاز كالحفظاظات والسيح والحقائب.
  5. رواية: وهي رواية يتم تأليفها وكتابتها بالكامل داخل مكان الاحتجاز.
  6. قصيدة: وهي القصائد والأشعار من تأليف المحبوس.
  7. مجلة: مجلة كاملة تصدرها مجموعة من المحبوسين أو المسجونين بشكل دوري داخل مكان الاحتجاز.

#### ب.: فيما يخص بيانات قانونية وإجراءات متعلقة بالواقعة:

- \*\* تم تقسيم الإنتاج الإبداعية وفقاً للوضع القانوني لصاحب المنتج الإبداعي وقت كتابة الرسالة كما يلي:
1. إعدام أو إحالة إلى المُفتي.
  2. محكوم.
  3. قيد التحقيق أو محال للمحاكمة.
  4. غير معلوم.

### ج. فيما يخص بيانات شخصية للمرسل:

\*\* تم تقسيم الإنتاجات الإبداعية وفقاً للفئة العمرية لصاحب المنتج الإبداعي كما يلي:

1. بين 10-18 سنة.
2. بين 18-30 سنة.
3. بين 30-40 سنة.
4. بين 40-50 سنة.
5. أكبر من 50 سنة.
6. غير معلوم.

\*\* تم تقسيم الإنتاجات الإبداعية وفقاً للنوع الاجتماعي لصاحب المنتج الإبداعي كما يلي:

1. ذكر.
2. أنثى.

\*\* تم تقسيم الإنتاجات الإبداعية وفقاً للمرحلة العمرية لصاحب المنتج الإبداعي كما يلي:

1. قاصر.
2. بالغ.

\*\* تم تقسيم الإنتاجات الإبداعية وفقاً للفئة الوظيفية لصاحب المنتج الإبداعي كما يلي:

1. عضو هيئة تدريس: تمثل الدكاترة والأساتذة الجامعيين والمعيدون بالجامعات.
2. طالب تعليم عالي: تمثل الطلاب الجامعيين.
3. طالب تعليم أساسي وثانوي: تمثل الطلاب الذين لم يتخطوا بعد مرحلة التعليم الثانوي والأساسي.
4. قطاع حكومي.
5. قطاع خاص وأعمال حرة.
6. نقابات مهنية: تمثل الأشخاص التابعون لنقابات مهنية مثل المهندسين والصيدلة والعلميين والأطباء والمحامين والمعلمين والصحفيين.
7. الحرفيون والفلاحون والعاملون باليومية.
8. خريج.
9. صحافة وإعلام.
10. محامة.
11. ربة منزل.
12. غير معلوم.

## د. فيما يخص بيانات قانونية وإجراءات متعلقة بالواقعة:

- \*\* تم تقسيم الإنتاجات الإبداعية وفقاً للإقليم الجغرافي لواقعة إتهام صاحب المنتج الإبداعي كما يلي:**
- تم اعتبار التقسيم الإداري للجمهورية وهي 27 محافظة، وتم تقسيمهم وفقاً للإقليم الجغرافي كالتالي:
- 1. المحافظات المركزية:** وتشمل (القاهرة، الجيزة، الإسكندرية).
  - 2. محافظات الدلتا:** وتشمل (القليوبية، الدقهلية، الشرقية، الغربية، المنوفية، البحيرة، كفر الشيخ، دمياط).
  - 3. مدن القناة:** وتشمل (بورسعيد، الإسماعيلية، السويس).
  - 4. محافظات الصعيد:** وتشمل (الفيوم، بني سويف، المنيا، أسيوط، سوهاج، قنا، الأقصر، أسوان).
  - 5. المحافظات الحدودية:** وتشمل (شمال وجنوب سيناء، مرسى مطروح، البحر الأحمر، الوادي الجديد).

## ثامناً: نظام الفهرسة والفرضيات:

- تم ترتيب الحالات/الصفوف بالكامل داخل قاعدة البيانات حسب النظام المُفهرس التالي للأعمدة:
  1. تاريخ الواقعة.
  2. محافظة الواقعة **Governorate Level** : حسب ترتيب المحافظات المركزية ثم محافظات الدلتا ثم مدن القناة ثم محافظات الصعيد ثم المحافظات الحدودية.
  3. نوع المنتج الإبداعي.

- تم فصل اسم مميز للواقعة أو مدى ارتباطها بحدث معين أو وجود اسم إعلامي لها في خانة جديدة تحت عنوان "حدث عام أو اسم إعلامي للواقعة".
- في خانة رقم المحضر أو القضية، تمت كتابة جميع الأرقام الرسمية حسب تراتبيتها لدى الجهات الرسمية وتمت فهرستها جميعاً كما يلي:
  1. كلمة البداية "رقم" ثم رقم الورق الرسمي مكتوباً بالأرقام وليست حروفاً.
  2. ثم كلمة "لسنة" ثم السنة الميلادية وأحياناً القضائية كما في النقض مكتوبة بالأرقام وليست حروفاً.
  3. ثم كلمات "إداري أو جنايات أو جنح أو جنائيات أو كلي أو عرائض أو حصر تحقيق".
  4. ثم دائرة الجهة الرسمية التابع لها الرقم.
  5. ثم كلمة "المقيدة ب" ثم نفس الترتيب السابقة إن كانت لنفس المؤسسة القضائية.

- مثال ذلك؛ (رقم 302 لسنة 2015 إداري قصر النيل) أو (رقم 14322 لسنة 2015 جنايات قصر النيل ورقم 543 لسنة 2015 كلي وسط القاهرة).
- في خانات التفاصيل الإسمية الخاصة بالمحوسين في قاعدة البيانات، تمت مراعاة أن تكون الكلمات في ملف الإكسيل وفقاً للمعايير الأرشفية للغة العربية، مما يُسهّل من عملية البحث بعد الضغط على CTRL+F أو إعادة استخدامها أكاديمياً بواسطة برامج إدارة البيانات:
  1. الألف "أ" تُكتب "ا" بدون همزات.
  2. التاء المربوطة "ة" تُكتب "ه".
  3. الألف المقصورة "ى" تُكتب "ي".

- تمت مراجعة قاعدة البيانات بالكامل، من أجل تقليل نسبة "الازدواج الحسائي" بتجنب التكرار، وذلك عن طريق عقد مقارنات تحليلية وواقعية بالنسبة للبيانات الوصفية والجغرافية والزمنية ، من أجل الوصول إلى أكبر دقة ممكنة.

## تاسعاً: معايير ورخصة نشر قاعدة البيانات:

- قاعدة البيانات مفتوحة ومتاحة بالكامل ومنشورة برخصة قواعد البيانات المفتوحة **Open Database License (ODbL) v1.0**.
- تمت مراعاة مبدئي عدم انتهاك الخصوصية وعدم جلب الضرر.
- تم نشر قاعدة البيانات بنفس البيانات الوصفية **Metadata** المستخدمة داخل ملف الإكسل دون حذف، من أجل المراجعات والتدقيق، مع إدراج المعادلات الإحصائية جميعها مفتوحة، كمزيد من الشفافية وإتاحة التعلم.
- تمت مراعاة الشفافية والحياد والدقة والالتزام الكامل بتحديد المنهجية وذكر المصادر لكل معلومة بشكل تفصيلي.
- تم إدراج المعادلات الإحصائية الخاصة بالجدول الإحصائية جميعها مفتوحة، كمزيد من الشفافية وإتاحة التعلم.
- تم حجب عديد من التفاصيل الشخصية مثل العناوين بالتفصيل، حيث تمت مراعاة احترام الخصوصية وعدم تمييز هؤلاء الأشخاص في محيط إقامتهم أو إقامة ذويهم بعلامات دلالية مباشرة مثل رقم الشقة أو المنزل أو الشارع وخلافه.
- الاصطلاحات والتوصيفات المستخدمة بشكل عام في التقرير تكون على خلفية أكاديمية أو علمية أو توصيفات مجردة للوقائع بعيدة عن أية صبغة سياسية أو أيديولوجية أو سياق حقوقي أو أهواء شخصية أو تعميم أو نتائج استباقية، قدر الإمكان.

## عاشرًا: التحديات وإشكاليات المتعلقة بعمليات جمع المعلومات والأرشفة:

- مركزية المعلومات بشكل عام في مصر.
- يعمل محرك بحث **Google** وفق آليات قد تهمل بعض نتائج عملية البحث، وخصوصًا مع تقادم الزمن.
- حجب المواقع الإخبارية ومحو الأرشيف الصحفي لبعض المواقع على الأنترنت مثل (البديل، البداية، ....).
- صعوبة تحديد التفاصيل الزمنية والمكانية للواقعة.
- عدم توفر المعلومات في بعض النطاقات الزمنية والمكانية للوقائع وندرته بشكل كبير.
- المخاطر الأمنية التي تواجه العاملين في المجال العام.
- استخدام الصحافة المصرية لقوالب ومصطلحات عامة لا تصف الواقعة بشكل محدد.
- عدم الوعي المجتمعي أو الفتوي بأهمية وجدوى التوثيق.
- التناقضات والمبالغات والتضليل في المعلومات.
- عدم قدرة مصدر المعلومة على التعبير وإيصال المعلومات بسهولة.
- الحفاظ على الخصوصية وعدم جلب الضرر.