دار نشر جامعة قطر
Qatar University Press

مركز ابن خلدون للعلوم الإنسانية والاجتماعية
Ibn Khaldon Center for Humanities and Social Sciences

OPEN ACCESS

تاريخ الاستلام: 6 ابريل 2023
تاريخ القبول: 24 يوليو 2023

مقالة بحثية

الدعوة إلى استخدام الوصف الشبكاتي المحوسب لمعجم الدوحة التاريخي للغة العربية، أنموذج تطبيقي على معجم "كشّاف اصطلاحات الفنون والعلوم"

حسن مظفر الرزو

رئيس باحثين أقدم، هندسة محاكاة الحاسب الآلي، المعهد العالمي لحوسبة القرآن الكريم والعلوم الإسلامية، العراق

halrizzo@gmail.com

ملخص

استهدف هذا البحث مراجعة المحتوى الرقمي لمعجم الدوحة التاريخي للغة العربية، في محاولة لاستثمار جزء من مدوّنته اللغوية ـ العملاقة كمادة أولية لممارسة سلسلة من المعالجات المحوسبة التي نروم من خلالها إلى توطين نهج شبكة تحليل النص(Text Network Analysis) ؛ لإغناء مادة المعجم بفضاء رقمي شبكاتي من نمط جديد يسهم في تثوير الكنوز المعرفية المودعة في مدونته اللغوية الخصبة. بيد أن استعصاء المادة الرقمية لمعجم الدوحة على الانصياع لعمليات المعالجة المحوسبة قد دفعنا إلى انتخاب أحد المعاجم الاصطلاحية المهمة، وهو كشاف اصطلاحات الفنون والعلوم مادة لعمليات التحليل المحوسب وإنتاج مجموعة متنوعة من أنماط التمثيل الرسومي التي أُنشِئت من مفردات ثلاث فقرات أُنتخبت من كتاب الكشاف لهذا الغرض، وهي فقرة حرف الباء، وفقرة العلم، وفقرة المعرفة. وقد أنتجت عمليات الحوسبة الذكية لمحتوى الفقرات الثلاث عددًا كبيرًا من شبكات النصوص تمكّنا من خلالها تقطير المادة المعرفية فيها، مع توفير فضاء من التمثيل الشبكاتي المرئي الذي يمكن أن يمنح القراء والباحثين فرصة مراجعة المحتوى قبل مطالعتها، واستثمار الوقت في عملية التنقير في مزيد من النصوص بدلًا من إضاعة الوقت في موارد قد لا تلبي حاجات القارئ أو الباحث.

الكلمات المفتاحية: معجم الدوحة التاريخي للغة العربية، تحليل شبكة النص، كشاف اصطلاحات الفنون، التنقيب في المعلومات، الحوسبة الذكية، شبكات المفاهيم

للاقتباس: الرزو، حسن. "الدعوة إلى استخدام الوصف الشبكاتي المحوسب لمعجم الدوحة التاريخي للغة العربية، أنموذج تطبيقي على معجم كشاف اصطلاحات الفنون والعلوم"، **مجلة تجسير**، المجلد الخامس، العدد 2 (2023)

https://doi.org/10.29117/tis.2023.0139

© 2023، الرزو، الجهة المرخص لها: دار نشر جامعة قطر. نُشرت هذه المقالة البحثية وفقًا لشروط Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). تسمح هذه الرخصة بالاستخدام غير التجاري، وينبغي نسبة العمل إلى صاحبه، مع بيان أي تعديلات عليه. كما تتيح حرية نسخ، وتوزيع، ونقل العمل بأي شكل من الأشكال، أو بأية وسيلة، ومزجه وتحويله والبناء عليه، طالما يُنسب العمل الأصلي إلى المؤلف

دار نشر جامعة قطر
Qatar University Press

مركز ابن خلدون للعلوم الإنسانية والاجتماعية
Ibn Khaldon Center for Humanities and Social Sciences

OPEN ACCESS
Submitted: 6 April 2023
Accepted: 14 July 2023

Research Article

Advocating the Use of Computational Text Network Analysis in the Doha Historical Dictionary of the Arabic Language, An Applied Model on "Kashshāf iṣṭilāḥāt al-Funūn wa-al-ʻUlūm"

Hassan Muthaffar Al-Rizzo

Senior Research Scientists, Computer Simulation, Institute of Computation of Quraan and Islamic Sciences, Iraq
halrizzo@gmail.com

Abstract

In the realm of computational linguistics, our research endeavors to unlock the vast potential of intelligent computational treatments for the enrichment of lexical content in Arabic language studies. Specifically, we embarked on an extensive review of the digital content within the Doha Historical Dictionary of the Arabic Language, aiming to harness a portion of its linguistic blog as valuable raw material. By employing advanced computational intelligence techniques, we localized the text network analysis approach to amplify the lexicon material present in the dictionary.

However ,encountering limitations with the digital material of the Doha dictionary prompted us to explore alternative avenues .In this pursuit ,we turned to" Kashshāf iṣṭilāḥāt al-Funūn wa-al-ʻUlūm," an indispensable terminological dictionary ,to serve as the basis for our computational analysis. Focusing on three carefully selected paragraphs—dedicated to the letter" Ba'a ",Science ,and Knowledge — we applied intelligent computational analysis techniques to generate a multitude of text networks .These networks effectively extracted embedded knowledge concepts while providing visually captivating representations that can offer readers and researchers the opportunity to review the content before reading it and invest their time in the process of screening more texts instead of wasting time on resources that may not meet the needs of the reader or researcher.

Keywords: Doha Historical Dictionary of Arabic Language ;Computational Intelligence ;Text Network Analysis ;Data Mining ;Concept Networks

Cite this article as: Al-Rizzo, Hassan. "Advocating the Use of Computational Text Network Analysis in the Doha Historical Dictionary of the Arabic Language, An Applied Model on "Kashshāf iṣṭilāḥāt al-Funūn wa-al-ʻUlūm". *Tajseer Journal*, Vol. 5, Issue 2 (2023).

https://doi.org/10.29117/tis.2023.0139

© 2023, Al-Rizzo, licensee QU Press. This article is published under the terms of the Creative Commons Attribution-Non Commercial 4.0 International (CC BY-NC 4.0), which permits non-commercial use of the material, appropriate credit, and indication if changes in the material were made. You can copy and redistribute the material in any medium or format as well as remix, transform, and build upon the material, provided the original work is properly cited.

مقدّمة

لا شك أن التأثير المصاحب لتوغّل تقنية المعلومات والاتصالات في حياتنا المعاصرة قد تولّدت عنه تأثيرات كثيرة؛ حيث ظهرت تحديات من نمط جديد في التعامل مع معطيات مجتمع المعرفة المعاصر؛ نظرًا للقدرات الكبيرة لأدوات الحوسبة، وبزوغ مفهوم البيانات الضخمة التي باتت تتكاثر مفرداتها تنتشر في الفضاء السيبراني.

وقد برزت بواكير عملية رقمنة الخطاب العربي، وبدأت مدونات اللغة العربية بمختلف علومها واختصاصاتها تستوطن في الفضاء الرقمي، وأُبتكرت أدوات مختلفة لتسهيل الوصول إليها وتفكيك مفرداتها؛ لبلوغ ما نروم نواله من حصيلتها المعرفية.

ولم تمر مدة طويلة حتى بدأنا نلحظ حضور معاجم اللغة العربية في البيئة الرقمية[1]، فتعددت الجهات التي عكفت على إعدادها لغايات علمية، أو تجارية صرفة. ولم تكن عملية ولوج المعاجم العربية في البيئة الرقمية موفقة إلى حدٍّ كبير، في بداياتها، فلم يتجاوز حضورها إدخال مادة المعجم إلى البيئة المحوسبة مع وجود أدوات بدائية للبحث عن المفردة بعيدًا عن حضور مصرف آلي[2] قادر على تتبع الاشتقاقات، أو مدقق نحوي، وغيرها من الأدوات اللغوية التي نجدها قد توطّنت في المعاجم المحوسبة بلغات أخرى[3].

وقد برز مشروع الدوحة التاريخي للغة العربية ورأى النور عام 2013 بوصفه سجلًا رقميًا مفتوحًا اتّسم ببناء لغوي تراكمي، يتيح للمستخدمين مجالًا رحبًا لرصد التطور الحاصل في معاني ألفاظ لغتنا العربية ومبانيها ودلالاتها عبر العصور[4].

وقد شهد مشروع معجم الدوحة تقدمًا ملحوظًا خلال السنوات الأخيرة، وانتهى فريقه من العمل على المرحلة الثانية من المشروع وفق ما جاء في إعلان المركز العربي للأبحاث ودراسة السياسات بتاريخ 1يونيو من عام 2022م، الذي تضمّن نشر عشرة قرون من تاريخ الكلمة العربية، وبحصيلة معجمية بلغت زهاء مائتي ألف مدخل معجمي[5]، مؤكدًا على تجاوز فريق العمل على تطوير برامج حاسوبية للمعالجة الآلية للغة العربية، وتصميم منصة تتضمن قواعد بيانات عملاقة، وأدوات بحث متقدمة، وخدمات متنوعة؛ لتيسير وصول مادته للمواطن العربي.

إلا أن زياراتنا المتكررة لموقع معجم الدوحة التاريخي للغة العربية، واستثمار جزء يسير من المادة العلمية الخصبة المودعة في موقعه العملاق، قد جعلتنا نقف عند الجهود الكبيرة التي بذلها المتخصصون بالحوسبة في توظيف أدوات الصرف الآلي،

1 - إيمان شاشه، وآمنة مناع، وحسام الدين تاوريريت، "المعالجة الرقمية للمشاريع العربية المعجمية المحوسبة: عرض وقراءة لجهود عربية منجزة"، **مجلة البحوث التربوية والتعليمية**، مج11، عدد خاص (2022)، ص60.

2 - أحمد هاشم السامرائي، "حوسبة المعجم العربي والقضايا المعاصرة"، **مجلة سرّ من رأى**، مج9، ع34، السنة التاسعة (تموز 2013)، ص2.

3 - عبدو خيرة، **معمارية المعجم العربي الإلكتروني**، أطروحة ماجستير، قسم اللغة العربية وآدابها، كلية الآداب والفنون، جامعة عبد الحميد بن باديس، الجزائر، 2015، ص55.

4 - حنان غياط، "صناعة المعجم التاريخي والتقنيات الحاسوبية: معجم الدوحة التاريخي أنموذجًا"، **مجلة جسور المعرفة**، مج 8، ع3 (سبتمبر 2022)، ص808.

5 - "بيان: المركز العربي للأبحاث يعلن عن الانتهاء من المرحلة الثانية في عمل معجم الدوحة التاريخي للغة العربية "،المركز العربي للأبحاث ودراسة السياسات، 2023/1/1، تاريخ الزيارة: 2023/4/22، الرابط :
https://www.dohainstitute.org/ar/News/Pages/acrps-announces-completion-of-phase-two-of-the-doha-historical-dictionary-of-arabic.aspx.

وتطوير أدوات حاسوبية لتحليل محتوى المدوّنة اللغوية الضخمة واستخلاص الحزم الاشتقاقية منها[6]، بالمقابل فإنها قد أثارت لدينا جملة من التساؤلات بخصوص كيفية زج نهج الحوسبة الذكية Computational Intelligence [7] ومعالجاتها المتطورة للارتقاء بأسلوب طرح مادة الكلام العربي، وتوفير آليات محوسبة تدعم المستخدم في إنشاء مستويات متعددة من شبكات المفاهيم، ودلالات الخطاب من هذه المادة الثرية بحيث يستطيع المستخدم (سواء كان قارئًا أو باحثًا) من سبر تطور دلالة اللفاظ عبر العصور، وتتبع تداعيات التغاير الذي يطرأ على الدلالات الاصطلاحية، وتبين شبكة العلاقات المقيمة بين جذور الكلمات ومعانيها.

فأردنا أن نوطّن آليات الحوسبة الذكية في فضاء المدوّنة اللغوية لمعجم الدوحة، من خلال توظيف تقنية تحليل شبكة النصوص Text Network Analysis وتقنيات التنقيب في المعلومات الداعمة لآليات المعالجة التي توظفها في إنتاج النسيج الشبكاتي للنصوص في محاولة لمنح المعجم مساحة معرفية جديدة، وفضاء استخدام يوسّع، إلى حدٍّ كبير، من قدرة المستخدم على تقطير الخلاصة المعرفية من مادة مفرداته الخصبة.

أولًا: خطاطة القراءة ــ عن بعد Distant-Reading

أصدر المفكر فرانكو موريتي (*Franco Moretti*)[8] في عام 2005 كتابه الشهير *Graph, Maps & Trees*، الذي أرسى فيه أسس الخطاطة المفاهيمية الجديدة[9] للتعامل مع النصوص، والتي أطلق عليها خطاطة القراءة ـ عن بعد، بوصفها البديل عن القراءة التقليدية للنصوص المطبوعة، والتي أطلق عليها القراءة ـ عن قرب Close Reading.

دعا فرانكو موريتي في كتابه إلى الإقلاع عن طريقة القراءة التقليدية، التي لم تعد متوافقة مع الحجم الهائل للمحتوى الرقمي الذي ازدحم به الفضاء الرقمي، والتوجّه نحو إعادة تشكيل النص وعرضه على شكل خرائط، وأشكال رسومية تحتوي الكم الكبير من النصوص في خلاصة معرفية مرئية ترشد الباحثين والقراء إلى ضالتهم الفكرية في زحام المحتوى الرقمي العولمي[10].

وإذا كان الإمام الغزالي قد طالع عن قرب، وبشغف مؤلفات شيخه الجليل الإمام الجويني لكي تتضح لديه آراؤه في العقيدة، وعلم الفقه وأصوله[11]، فإن علماء المستقبل سيكونون بحاجة إلى ممارسة آلية القراءة عن بعد؛ لكي تتوفر لديهم فرصة الاطلاع على العدد الهائل من الكتب التي صنفت في علوم العقيدة وأصول الفقه، والتي لم تعد تقتصر على الكتب التقليدية، بل باتت متوفرة في الكتب الإلكترونية، ومواقع الويب، ومواقع التواصل الاجتماعي، إضافة إلى المستودعات

6 - غياط، ص811.

7 - يعد نهج الحوسبة الذكية من المعالجات المحوسبة الحديثة التي تسعى إلى توظيف الآليات التي يمارسها الذهن البشري، في برمجيات تتسم بالذكاء في التعامل مع الإشكاليات والمسائل التي يعجّ بها واقع الانسان المعاصر، وبمقاربة لآليات الذهن التقليدية مثل: التكيّف مع الواقعة كما في نهج الشبكات العصبونية Neural Networks، أو استخدام المنطق البشري بعيدًا عن المنطق الرياضي الصارم كما في نهج المنطق المضبّب Fuzzy Logic واستخدام شبكات تحليل النص Text Network Analysis لتكوين تصوّر عن شبكة المفاهيم التي تستودع في النصوص.

8 - المفكر والمؤرخ والأديب الإيطالي ولد عام 1950 بمدينة سوندريو الإيطالية، وتدرج بالدراسة والتدريس في جامعاتها قبل أن ينتقل الى جامعة كولومبيا في الولايات المتحدة عام 1990، ثم انتقل إلى جامعة ستانفورد، وألقى محاضرات في جامعة بيركيلي. يعدّ المؤسس لمبدأ القراءة عن بعد والدعوة إلى تأسيس خطاطة جديدة للتعامل مع المصادر التاريخية والأدبية في زمن الذكاء المحوسب.

9 - Vita Mozuraite, "Change of the Reading Paradigm in the Age of E-Book", *Libellarium*, VII, 1 (2014), p. 85.

10 - S. Jänicke et al., "Visual Text Analysis in Digital Humanities", *Computer Graphics Forum*, Vol. 36, No. 6 (2016), p. 2.

11 - حاولنا في هذا المثال بيان أن الإمام الغزالي بجلالة علمه وكونه تلميذًا للإمام الجويني كان بحاجة إلى مطالعة كتبه عن قريب للوقوف على ما ورد فيها، بينما عند استخدامنا للتقنيات الذكية لن نكون بحاجة إلى مطالعة الكتاب، وإنما معاينة محتوياته بواسطة الأدوات الذكية التي تمنحنا فرصة تبويب المحتوى، وشبكة المفاهيم بصورة سريعة.

الرقمية العملاقة.

لقد تحوّلت الخطاطة الجديدة إلى اعتماد أدوات الحوسبة الذكية، والمعالجات الإحصائية لتقطير العصارة المعرفية من ملايين الصفحات المنتشرة في الفضاء السيبراني، وعرض المفردات بصورة مرئية لكي نستطيع أن نتبين ونحدد طبيعة الأنماط والتوجهات التي استودعت النصوص، كما أن مخرجاتها ستوفر لنا فرصة تحليل شبكة المفاهيم والاصطلاحات ومستوى الترابط الذي يسود نسيج النص، وهل أن المادة المطروحة ترقى إلى مستوى معنوي يتوافق مع ما نريد الحصول عليه من معرفة. انظر الجدول (1) الذي حاولنا من خلاله بيان أهم الفروق المعرفية بين القراءة ـ عن قرب، والقراءة ـ عن بعد[12].

الجدول (1): أوجه التباين بين القراءة. عن قرب والقراءة. عن بعد

القراءة . عن قرب	القراءة . عن بعد
يطلق على من يمارس هذا النمط من القراءة اصطلاح قارئ.	يطلق على من يمارس هذا النمط من القراءة اصطلاح مُعايِنْ[13].
تعتمد مبدأ مطالعة محتوى الكتاب التقليدي.	تعتمد مبدأ معاينة مفردات النصوص الرقمية بمختلف أشكالها.
يحرص الباحث على قراءة كل كلمة وردت في النص دون أن يغفل مفردة واحدة لضمان استيعاب فحوى الخطاب.	يحرص الباحث على عدم تضييع جهده في قراءة كلمات، واصطلاحات، وعبارات لا تقع في دائرة اهتمامه يوجّه جل اهتمامه نحو كلمات وعبارات محددة ومشيحًا بوجهه عن بقية النص.
يسعى القارئ إلى تفسير عبارات النص المكتوب، وسبر طبقات المعاني المودعة فيها، وتتبع تحولات الخطاب في العبارات.	يسعى القارئ إلى تشكيل مشهد مجرّد لمادة المحتوى لحجم كبير من النصوص والتحوّل من الفعل القرائي باتجاه العرض الرسومي للمفردات التي تتميز بمحمولات معرفية مكثفة.
يحاول دراسة ما يرد في العبارات من مصطلحات، وتداعيات بالمعاني، وترابط الحوادث، وبنية النص اللغوية والصرفية، وغيرها من التفاصيل التي تمنحه فرصة فهم جميع مستويات الخطاب المطروح في حدود النص.	يتعامل القارئ المعاصر مع محتوى النص بوصفه حقيبة من الكلمات والمصطلحات التي يعيق فهمها مجموعة من الأفعال، والحروف واللواصق التي يعدّها عقبة تحول دون قدرته على تحليل نسيج مادة الخطاب واستخلاص عصارتها المعرفية، لذا يفرغ النص منها ليستبقي المفاهيم في معمارية ذكية تتوافق مع حاجاته.
يتعامل القارئ التقليدي مع محتوى النص بوصفه كتلة متكاملة تضم جميع تفاصيل الخطاب المعرفي، وتداعيات فكر المؤلف بالتعامل مع جميع أشكال الأحداث، مهما كان الحقل الذي تنتمي إليه.	يتعامل القارئ المعاصر مع محتوى النص بآليات تتناسب مع خصائصها، فيوظّف العرض الرسومي لبيان مستويات حضور المصطلحات وتداعياتها داخل حدود النص، بينما يميل إلى استخدام الخرائط لوصف المادة التاريخية، وعناقيد المشجّرات لوصف تقاسيم المفاهيم ونشوء المصطلح، وتوسع دائرة استخدامه، والتوسع في دائرة المعاني الملتحقة به.

ورغم الدعوات المتكاثرة للتحوّل باتجاه القراءة. عن بعد وهجر نهج القراءة. عن قرب[14] إلا أننا نرى أن الأساس في نهل المعرفة من مصادرها واستيعاب تفاصيلها لن يكتب له النجاح إلا بالقراءة. عن قرب، وستبقى القراءة. عن بعد آلة من آليات دعم القارئ في الاطلاع على ما تحتويه البيانات العملاقة والنصوص المتكاثرة من محتوى معرفي، لكي يوجّه اهتمامه بالقراءة. عن قرب لما وفرت له القراءة. عن بعد من موارد معرفية خصبة.

12 - S. Jänicke, G., Franzini, & G. Scheuermann, "On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges", *The Eurographics Association* (2015), p. 2.

13 - عاين يعاين معاينةً وعِيانًا، فهو مُعايِن. تطلق عبارة معاين على من عاين النص فرآه أو شاهده بعينه، وتحقَّق منه بنفسه بنظرة عامَّة أو شاملة عليه دون قراءة مفرداته.

14 - Fabio Ciotti, "Distant Reading in Literary Studies: A Methodology in Quest of Theory", *Testo e Senso*, No. 23 (2021), p. 198.

ثانيًا: التحليل المحوسب للنصوص Text Mining

يُعرّف النص في المدوّنة المعلوماتية المعاصرة بأنه شكل من أشكال البيانات غير المهيكلة التي يتركب من مفرداتها الخطاب المعرفي الإنساني، وتتألف من مجموعة جمل Sentences، وفقرات Paragraphs، تتلاحم لإنتاج محتوى، يشتمل على مجموعة الفقرات والجمل التي تترابط مفرداتها للانخراط بمضامين الفقرة التي تنتسب إليها. بينما تتكون هيكلة الجملة من سلسلة مترابطة من المفردات اللغوية (اسم، فعل، وحرف)، تحكم ارتباطات معانيها مجموعة من القواعد النحوية والصرفية، لضمان إنتاج معنى محدد.

تنتمي عملية التحليل الآلي للنصوص إلى حقل التنقيب عن البيانات Data Mining وينصبّ اهتمامها على معالجة النصوص التي يتألف منها الخطاب الإنساني، وتتألف من مجموعة من المعالجات الذكية[15] التي تهدف إلى استخلاص المادة المعرفية ـ الضمنية Implicit Knowledge التي تستبطنها النصوص بمختلف أشكالها.

وتتضمن هذه المعالجات مجموعة من المهام الذكية[16] التي تمارسها خوارزميات الحاسوب مثل: تنقية المحتوى من الشوائب اللغوية، وتبويب مادة المحتوى Classification على أساس أقسام الكلم، وهيكلتها في عناقيد وكتل لغوية، تحمل معان وموضوعات محددة بواسطة التحليل العنقودي Clustering Analysis، وتحديد هوية الارتباطات القائمة بين مفرداتها، وتوجهات إنتاج المعاني والدلالات Trends[17].

وتستثمر مخرجات عمليات التحليل المحوسب للنصوص في استخلاص الخلاصات المعرفية الموجودة في النصوص المنتشرة في الكتب، والمجلات، ومواقع الويب ومستودعات البيانات الضخمة، لسبر المعاني والدلالات المستوطنة فيها، وصناعة القرارات، والتنبؤ بتوجهات النشاط الإنساني.

ثالثًا: النص: معالجة المفهوم في دائرة الخطاطة المحوسبة[18]

يعد النصّ المورد الأساسي لعمليات التنقيب في النصوص Text Mining وقد استوطن هذا المصطلح في دائرة العلوم الأدبية والإنسانية منذ أزمنة متباعدة، فأصبح يطلق على أي نتاج إنساني في هذه الحقول، تتألف مادته من نسيج من مفردات الكلم، صيغت في عبارات مثقلة بالدلالات والمعاني، وصيغت لنقل مضمون خطاب معرفي إلى متلقٍ، أو جمهور من المتلقين[19].

مادة النص وعناصره

تتعامل معالجات حوسبة النصوص والتنقيب في النصوص Text Mining في تربتها المعرفية مع النص بوصفه كتلة

15 - تتميز المعالجات الذكية Intelligent Processing باستخدام خوارزميات برمجية تتسم بذكاء في معالجة المفردات، وإنتاج مخططات المفاهيم بواسطة آليات الشبكات العصبونية أو المنطق المضبب التي تحاكي الذكاء الإنساني بآليات تتجاوز المعالجات الآلية التقليدية.

16 - يطلق اصطلاح المهام الذكية على مجموعة الخطوات والإجراءات التي تُنفذ في ضوء منطق برمجي تحدد من خلاله طبيعة الخطوات المحوسبة المطلوبة إزاء كل مسألة يطرحها الواقع بمعزل عن تدخل الإنسان المباشر؛ حيث تمارس المعالجات والمهام بصورة آلية محوسبة.

17 - JO, Taeho, *Text Mining: Concepts, Implementation, and Big Data Challenge* (Gewerbestrasse: Springer, 2019), p. 3.

18 - يطلق اصطلاح الخطاطة على اصطلاح Paradigm الذي يذهب البعض إلى ترجمتها الحرفية إلى اللغة العربية براديغما، والخطاطة المحوسبة هي عبارة عن مجموعة المفاهيم والقواعد والإجراءات التي توظف الذكاء المحوسب Computational Intelligence في معالجة النصوص والمسائل الفيزيائية عبر مجموعة من النماذج الرياضية والمنطقية لاستخلاص المعاني والدلالات وإنتاج الحلول.

19 - Ronen Feldman & James Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data* (UK, Cambridge: Cambridge University Press, 2006), pp. 3-4.

لغوية صيغت بواسطة اللغة الطبيعية[20] (NL) Natural Language التي نتواصل بمفرداتها في الخطاب المدوّن ويتألف قوامها من شبكة متلاحمة من الكلمات، والجمل، والفقرات.

بصورة عامة، يتألف لباب أي نص من النصوص، من مجموعة فصول، تحتوي على مجموعة من الفقرات، بينما تضم الفقرات عددًا متباينًا من الجمل. ويمكن تعريف الفصل بأنه عبارة عن مجموعة من الفقرات التي تنتمي إلى تبويب موضوعي مشترك. أما الفقرة فتتألف من مجموعة من الجمل/العبارات التي كُتبت بواسطة اللغة الطبيعية، وتلتحق مضامينها بعنوان الفقرة إلى تنتمي إليها.

وتعدّ الكلمات أصغر وحدة بنيوية في النص تُشكّل من خلالها الجمل بتوظيف القواعد النحوية والصرفية لإنتاج معنى محدد من الحروف التي تتألف منها. ولا يلتفت التحليل المحوسب إلى أقسام الكلم[21]، سواء كانت الكلمة اسمًا أو فعلًا، أو صفة، أو حرفًا، وإنما يتعامل مع مجموعة الحروف التي تشكل كائنًا لغويًا مستقلًا يحمل معنى تنطق به الكلمة، التي ينصب اهتمامه على معالجة مضامينها.

وتستخدم الكلمات لصياغة مجموعة متنوعة من الجمل من خلال توظيف القواعد النحوية والصرفية، وتنفصل كل كلمة عن الكلمة التي تليها بواسطة فراغ، وتنتهي الجملة بنقطة النهاية، أو علامة استفهام، أو علامة تعجب. ويفترض بالجملة أن تحمل مضمونًا يعبّر عنه بالبناء اللغوي لمفرداتها.

بصورة عامة يطلق على البنية اللغوية التي تستقل بذاتها من حيث المعنى المحمول في عناصرها اصطلاح العبارة أو الجملة المستقلة ((Independent Clause، بينما يطلق اصطلاح العبارة غير المستقلة أو العبارة الثانوية (Subordinate Clause) على العبارة التي لا تتضمن معنى مفهومًا قائمًا بمفرداتها وتفتقر إلى عبارة مستقلة؛ لكي تصبح جملة تامة قادرة على الإفصاح عن مفهوم جليّ[22].

وتحتوي الجملة على الأسماء التي تعبّر عن مواضيع أو ذوات، ترتبط فيما بينها بواسطة أفعال للدلالة على مستويات الحدوث، وضمائر، وأدوات جرّ لاستكمال بناء معمارية المعاني المودعة في الجمل.

وتعتمد مجموعة من القواعد في تمييز حضور الجملة داخل النصوص (وفق منطق الحوسبة) بحيث يمكن ممارسة المعالجات المحوسبة على عناصرها بوصفها جملًا /عبارات مستقلة، لعل أهمها[23]:

- تبدأ الجمل في فقرة جديدة، أو بعد علامة تؤشر نهاية جملة سابقة (.).
- تنتهي الجمل بعلامة قد تؤشر نهايتها (.) أو علامة استفهام (؟)، أو علامة تعجّب (!).
- الجملة تحتوي على موضوع (اسم) الذي يذكر لمرة واحدة.
- الجملة تحتوي على فعل أو ضمائر.
- ينبغي أن تحتوي الجملة على معنى مفهوم تستقل بوصفه دون وجود حاجة إلى أمر إضافي لإرساء المعنى.

20 - لا تنتمي النصوص التي نعدها باللغات البرمجية أو تلك التي تنتجها الأدوات الرقمية إلى حقل النصوص، وتعامل بوصفها بيانات تعالج بواسطة تقنيات التنقيب في البيانات Data Mining.

21 - يختلف المنطق المحوسب في المعالجات اللغوية إلى حد كبير مع المنطق التقليدي حيث يتعامل في البداية مع المفردة بوصفها كيانًا قابلًا للحوسبة، بصرف النظر عن كون المفردة مفردة أم مركبة. ويميز بين المفردات على أساس الفراغ الذي يفصلها عن سابقتها أو اللاحقة. ثم تباشر لاحقًا عمليات التقاط الاصطلاحات، وتمييز أجزاء الكلام، والتحليل الصرفي للجمل.

22 - Walden University, "Grammar: Sentence Structure and Types of Sentences" Accessed on 7/10/2022.Available At: https://academicguides.waldenu.edu/writingcenter/grammar/sentencestructure

23 - نقصد بكلامنا هنا حوسبة الجمل المكتوبة باللغة الإنجليزية، بينما سنعالج مسألة الجمل العربية في فصل لاحق.

وفق هذه القواعد وقواعد أخرى قد يحتمها منطق الحوسبة الذي يستخدم في التنقيب بمادة النص تباشر عملية التقاط الجملة من الفقرات، والنصوص، ثم يصار في مراحل لاحقة إلى انتخاب المواضيع، والكشف عن الذخيرة المعرفية المودعة فيها24. ويتباين التعقيد في البنية اللغوية للجمل، فهناك جمل بسيطة تتألف من عبارة مستقلة واحدة، بينما نجد في الوقت نفسه جمل معقّدة تتألف من عبارة مستقلة تلتحق بها عبارة أو مجموعة عبارات غير مستقلة، تتكامل فيما بينها لإنتاج معنى محدّد.

رابعًا: الوصف الشبكاتي المحوسب للنصوص

توجهت الأنظار في السنوات الأخيرة إلى تمثيل نصوص الخطاب المدوّن في المقالات والكتب بواسطة مجموعة متنوعة من المعماريات الشبكاتية التي تعالج مادتها على أساس وصف مفردات النص بالعقد التي يتألف منها نسيج النص، وتوظّف البنى اللغوية والدلالية التي تسود مادة الخطاب المدوّن لإنشاء شبكات العلاقات القائمة بين المفردات، والتي يعبّر عنها بالحافات في مضمار التمثيل الرسومي للشبكات.

وتعد عملية تحليل النص Network Text Analysis (TNA) من المقاربات التي أفرزتها علوم الشبكة، والتي جعلت من مفردات النص مادة خصبة لاستخلاص المادة المعرفية التي تستوطن في مادة خطابها، وعرضها في أشكال مرئية تصف معمارية المفاهيم التي تسود مادة النص، مع وصف العلاقات الرابطة بين مفرداتها[25] والتي تعد مؤشرًاعلى البنية اللغوية والدلالية المقيمة في نسيجه المدوّن.

وتشكّل المفاهيم البنية الأساسية التي ترتكز إليها معمارية العرض الرسومي لمادة النص؛ حيث تُمثّل المفاهيم وفق النسق الشبكاتي بمجموعة من العقد التي تصف الكلمات والاصطلاحات التي تحفل بها النصوص، بينما تمثّل الحافات العلاقات القائمة بين عقدتين من عقد النسق المفاهيمي. ويعبّر النسيج الشبكاتي (الذي يعرض حضور المفاهيم وترابطاتها) عن الأنموذج العقلي Mental Model الذي يستوطن المادة المعرفية للخطاب المدوّن.

ويمكن لأي نص من النصوص أن تعالج مفرداته لغرض تمثيله رسوميًا، حيث يعبّر عن الكلمات أو الاصطلاحات أو المفاهيم بواسطة العقد، بينما تمثّل العلاقات على شكل الارتباطات أو الحافات التي تتألف منها مادة النسيج الشبكاتي للنص.

وتسهم عملية التمثيل المرئي للنص بإعادة تشكيل مادته بحيث يمكن توظيف أكثر من أداة رياضية، وخورازمية برمجية لممارسة سلسلة من المعالجات الكمية وتصنيف وتبويب مفرداته، لتحديد طبيعة الأنماط السائدة في نسيجه، والكشف عن المجاميع الواصفة لمحتواه المعرفي، وبيان حجم الدور الذي تمارسه الكلمات والاصطلاحات والمفاهيم في إنتاج المعنى داخل حدود النص.

من جهة أخرى فقد شاع استخدام سُحُبْ الكلمة Word Clouds بوصفه أحد أنماط العرض الرسومي لمحتوى النصوص، فتعرض الكلمات رسوميًا بأحجام تتناسب مع مستويات تكرارها داخل حدود النص[26]. ويعد هذا النمط من الأنماط المبسطة في عرض مادة النصوص لاعتماده على تكرار الكلمات دون الأخذ بعين الاعتبار المعاني والدلالات المستودعة فيها[27].

---

24 - Walden University, "Grammar: Sentence Structure and Types of Sentences"

25 - العلاقة الرابطة (بصورة أولية) بين المفردات هي علاقة تابع ومتبوع تتشكل عنها مادة النسيج الشبكاتي للنص، ثم تأتي عملية تشكيل المفاهيم من مجموعة العلاقات التي تربط المفردة الواحدة مع بقية المفردات في أكثر من جملة وفقرة يمكن بواسطة الذكاء المحوسب الكشف عن شبكة المفاهيم.

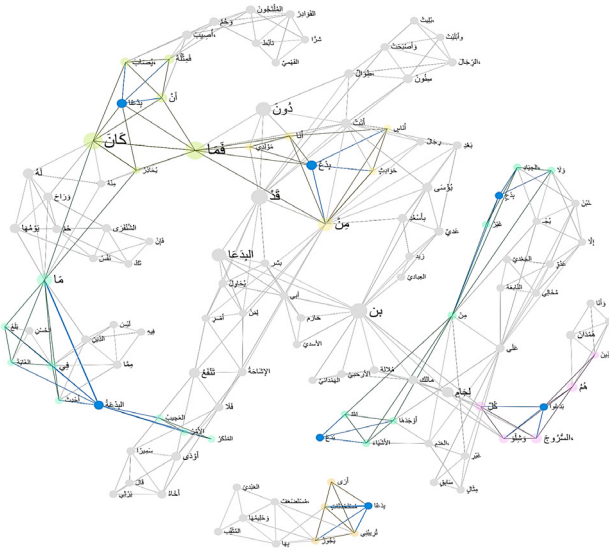
26 - Ji ,Wang ,*Clustered Layout Word Cloud for User Generated Online Reviews*, Thesis (Virginia: Virginia Polytechnic Institute and State University, 2012), p. 16.

27 - Marita Hearst & Daniela Rosner, *Tag Clouds: Data Analysis Tool or Social Signaller?*, *School of Information* (Berkeley: University of California, Berkeley, 2008), p. 3.

وقد برزت خلال السنوات الأخيرة محاولات جادة لتضمين دلالات الألفاظ في العرض الرسومي لسحب الكلمات من خلال توظيف معالجات اللغة الطبيعية، والتي نجحت بإنشاء سُحُبْ رسومية تأخذ بعين الاعتبار البعد الدلالي للكلمات، والمسافة التي تفصل فيما بينها داخل حدود النص إضافة إلى تكراراتها في تشكيل مادة العرض الرسومي[28].

خامسًا: النص المنتخب

في البداية حاولنا التنقير في مادة معجم الدوحة التاريخي للغة العربية، وسعينا إلى انتخاب مفردات لغوية متعددة لتكون مادة لمعالجاتنا المحوسبة، فحاولنا التفتيش في معجم الدوحة عن الجذر الثلاثي لكلمة بَدَعَ واشتقاقاتها وحضورها في تاريخ المدونة اللغوية العربية، بيد أن الموقع كان عصيًّا علينا؛ فلم يوفر لنا سوى استنساخ بضعة عبارات من مادته الثرية[29]، كما أنه لم يكن من السهل الوصول إلى موارده الرقمية. فلم نفلح في معالجاتنا الأولية سوى بالحصول على نتائج متواضعة حاولنا أن نودعها في الشكل (1).



الشكل (1): التمثيل الشبكاتي لجذر "ب. د. ع" في معجم الدوحة التاريخي للغة العربية[30].

يبدو واضحًا من الشكل أعلاه أن الشبكة الدلالية لاشتقاقات (بدع) لا تبدو كافية لكي توفر لنا صورة واضحة المعالم عن المفاهيم والاصطلاحات المستوطنة في مجال هذا المصطلح؛ حيث نلاحظ محدودية عقد شبكة العلاقات الرابطة فيما بينها (6 عقد مفاهيمية) رغم تكاثر جذور الكلمة. ويعزى ذلك إلى صعوبة الوصول إلى مستودع المعاني الخاص بالكلمة في الموقع والذي لا يتيح للمستخدم العادي الوصول إلى قواعد بيانات المعجم.

إن صعوبة الوصول إلى الموارد الرقمية للمعجم التاريخي للغة العربية، وعدم كفاية عمليات المسح الرقمي في مدوّنة معجم الدوحة قد أجبرتنا على أن نتوجه إلى أحد المعاجم اللغوية والاصطلاحية الجامعة، لكي نمارس على مادته عمليات المعالجة المحوسبة، ونستطيع توليد مجموعة متنوعة من التمثيلات الرسومية للمفردات الموجودة فيه، لكي نبرّر دعوتنا

28 - Guneri, Fatma, Hayırcıl, Gulrenk & Selin Iman, "A Word Visualization Observation: The Hidden Meaning of the Words in Work Environments", *European Scientific Journal*, *ESJ*, Vol. 17, No. 36 (2021), p. 22.

29 - موقع معجم الدوحة التاريخي للغة العربية: تاريخ الزيارة: 2022/2/7، الرابط: https://www.dohadictionary.org/dictionary

30 - الشكل من إعداد الباحث.

إلى النمط الجديد الذي ندعو إلى توطينه في المجال الرقمي لمعجم الدوحة للغة العربية.

فوقع اختيارنا على معجم كشاف اصطلاحات الفنون والعلوم لمؤلفه محمد بن علي التهانوي ليكون مادة لعمليات المعالجة المحوسبة، وعمليات التمثيل الرسومي التي سندعو إلى أن تعتمد في مرحلة لاحقة بمعجم الدوحة التاريخي للغة العربية.

سادسًا: البيئة البرمجية والمكتبات اللغوية _ المحوسبة التي استخدمت في معالجة النص

بصورة عامة تتوفر بيئات برمجية متعددة يمكن توظيفها في معالجة وتحليل مادة النصوص، وتعد لغة بايثون الذكية Python من أكثر البرمجيات استخدامًا هذه الأيام في التحليل المحوسب للنصوص بسبب كثرة المكتبات العربية ـ البرمجية التي تعمل في بيئتها المحوسبة، ووفرة أكثر من نهج لحوسبة النص العربي على مستوى أقسام الكلام، والتحليل الصرفي، وتوجهات الخطاب، وغيرها من المتغيرات اللسانية التي يمكن حوسبتها بواسطة هذه المكتبات[31].

لقد اعتمدنا في عملية تحليل نص كتاب كشّاف اصطلاحات الفنون إلى المبدأ الذي اعتمدته نظرية تحليل شبكة النصّ Text Network Analysis في التعامل مع مادة النصوص بوصفها عبارة عن نسيج من العقد المترابطة التي تمثل كل منها كلمة، أو هوية نصيّة مهما كان نوعها. أما حافات الشبكة ومادة خيوطها فتعبّر عن طبيعة العلاقات التي تربط بين هذه العقد.

ولترجمة هذه النظرية إلى منطق برمجي استخدمنا لغة بايثون Python البرمجية مستثمرين المجموعة الخصبة من المكتبات اللغوية الملتحقة بها والتي تسخّر للمستخدمين تحليل مفردات النص، بمستويات سطحية تتناول تفكيك مفرداته، وتصنيفها، وإحصاء تراتبية ومستويات حضورها في نسيج النص؛ ومستويات عميقة تشمل معالجات لغوية، وصرفية، ودلالية، مع تحديد هوية المفاهيم المودعة فيه.

وقد أُنشئت معمارية برمجية ضمن الخدمات التي توفرها أدوات معالجة اللغة الطبيعية NLTK، ضمت مجموعة من الخوارزميات المنطقية والرياضية التي استثمرنا من خلالها الإيعازات البرمجية التي تمارس عملية تحليل ومعالجة اللغة الطبيعية Natural Language Processing (NLP) لضمان استكمال المعالجات بشقّيها السطحي والعميق. ويظهر في الجدول (2) أهم المكتبات البرمجية التي استخدمناها في عملية تحليل نسيج نص الكتاب المنتخب.

الجدول (2): خصائص ووظائف المكتبات البرمجية التي استخدمت في تحليل النصّ المنتخب[32]

مستوى المعالجة	المكتبة البرمجية	الخصائص والوظائف
معالجة سطحية	Panda -NLTK	قراءة النص
	Panda-NLTK	ترميز مفردات النص Tokenization.
	Panda-NLTK	إزالة مستبعدات الفهرسة Removal Words Stop.
	SpaCy	إنشاء سحابة الكلمات Cloud Word.
	Panda-NLTK	إحصاء تكرار حضور الكلمة Frequency وتراتبيتها Sequence.

31 - Osama Obeid et al., "CAMeL Tools: An Open Source Python Toolkit for Arabic Natural Language Processing", *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, pp. 7022–7032 (Marseille, 11-16 May 2020), p. 7024.

32 - GitHub, "Python Package For Arabic Natural Language Processing", Accessed on 7/5/2023. Available At: https://github.com/adhaamehab/arabicnlp.

معالجة عميقة	CAMEL	التحليل الصرفي لمادة النص Analysis Morphological.
	CAMEL	تحليل انتماء المفردة إلى أجزاء الكلام Analysis Speech-of-Part.
	CAMEL	إزالة اللواصق واللواحق من الكلمات Lemmatization.
	InfraNodus	إعداد الوصف الشبكاتي لنسيج النص Visualization Text. مكتبة Networkx.
	Farasa	إنشاء شبكات المفاهيم Mapping Concept..

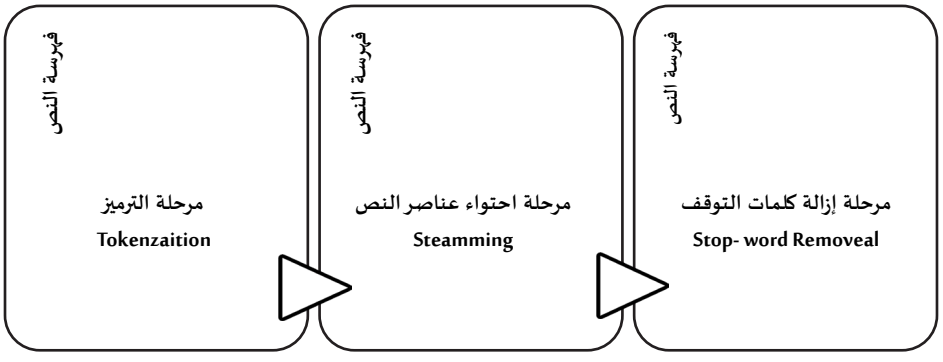
وقد تكامل عمل الخوارزميات البرمجية التي استثمرنا بواسطتها القدرات والخدمات التي توفرها هذه المكتبات البرمجية في إجراء عمليات ترميز مفردات النص، وإزالة مستبعدات الفهرسة، وممارسة تحليل أقسام الكلام، والتحليل الصرفي، وإعداد الوصف الشبكاتي لنسيج النص، وتحليل شبكات المفاهيم المودعة في الفقرات التي انتخبناها من كتاب الكشاف.

سابعًا: المعالجات المحوسبة ــ المسبقة لنص كتاب كشاف اصطلاحات الفنون والعلوم

ترتكز عملية المعالجة المسبقة للنص إلى مبدأ تحويل مادة النص غير المهيكلة إلى معمارية مهيكلة يمكن للحاسوب أن يتعامل معها بوصفها مدخلات للبيئات البرمجية المختلفة. ويطلق على هذه المرحلة تفكيك النص وفهرسة كلماته Text Indexing والتي من خلالها تجرى عملية تفكيك فقرات النص وعباراته وصولًا إلى الكائنات اللغوية بمختلف أنماطها (الاسم/ الفعل/ الحرف/ الصفة/ الظروف/ الضمائر/ ...) والتي تدرج في قائمة يؤشر فيها تسلسل حضور الكائن اللغوي داخل حدود النص.

إن النص الذي نتداوله في حياتنا اليومية يُعدّ (وفق منظور الحوسبة) مادة خام لا يمكن أن نتعامل معها بأدوات الحوسبة؛ لأن النص يمثل كتلة لغوية كبيرة تصعب حوسبتها، ما لم نشرع بتفكيكها إلى أجزاء وكائنات أصغر (كلمات)، يمكن أن تتقبلها البيئات البرمجية كمورد لمعالجتها المحوسبة وفق البرنامج الذي أُعد بواسطة لغة بايثون Python Language.

بصورة عامة تتألف عملية فهرسة النصوص من ثلاث مراحل أساسية، هي مرحلة الترميز Tokenization، ومرحلة احتواء مفردات النص Stemming، وأخيرًا مرحلة إزالة كلمات التوقف Stop-Word Removal. انظر الشكل (2).



الشكل (2): مراحل فهرسة النصوص غير المهيكلة[33].

33 - الشكل من إعداد الباحث.

تباشر عملية الترميز من خلال تجزئة محتوى النص إلى كيانات Objects منفصلة، بناء على حضور الفراغات أو علامات الترقيم بين البنيات اللغوية التي تتألف منها مادة النص. فيصار إلى تفكيك الفقرات، والجمل إلى عناصرها اللغوية الأولية مع استبعاد الفراغات، وعلامات الترقيم، ومختلف أنماط الفواصل المستخدمة بالنص. فيتحول النص الذي أُدخل بصيغته التقليدية (كلمات تستوطن في جمل وعبارات، وعبارات تستوطن فقرات، وفقرات تستوطن في كامل بنية النص) إلى قائمة من الكلمات، ثم يصار إلى تنقية الكلمات من الرموز الرقمية، مع تحديد وزن تكرارها في النص ومواقعها في الخطاب الذي يحمله النص، بدلًا من تكرار حضورها في قائمة الكلمات . انظر الجدول (3)؛ حيث أودعنا في الجدول أنموذجًا من مخرجات عملية الترميز مع تكرارات حضور هذه المفردات في نص الكتاب.

الجدول (3): أنموذج من مخرجات ترميز محتوى النص المنتخب

الرمز	التكرار	الرمز	التكرار	الرمز	التكرار	الرمز	التكرار
!	13	آباء	3	مظنة	3	مع	1618
'	9	آباءَكُمْ	2	مظنون	2	معا	148
"	1	آباءَنا	1	مظنونا	2	معابدهم	1
s'	25	آباؤك	1	مظنونة	1	معايير	1
)	11955	آباؤُنا	1	مظنّة	1	معاد	5
(	11944	آبائك	1	مظهر	38	معادا	1

وتوظف معالجة الاحتواء Stemming Process لتقليص حجم النص، وزيادة قدرة التحليل المحوسب للنص في إدارة البيانات، وتحليل مضامينها من خلال توحيد أجزاء الكلم، وإلحاقها بالجذور التي اشتقت منها. وتمارس عملية الاحتواء على كل من عناصر: الأسماء، الأفعال، والصفات. كذلك تُحول أسماء الجموع إلى أسماء مفردة.

وقد اعتمدت آلة الصرف العربي على مبدأ تقسيم ما يلتحق بالكلم العربي إلى ثلاثة أقسام، هي: السوابق واللواحق والمقحمات.

فالسوابق حروف ظاهرة تدلّ على معان، تتصل بأول الكلمة، فتحدث تغييرًا في البنية الصرفية للكلمة، وتنقسم إلى أربعة أقسام[34]. أما اللواحق فهي ضمائر متصلة تلتحق بالمشتقات الاسمية والفعلية، وكذلك تلتحق بالحروف. تتميز بأن لها القدرة على إحداث تغيير في البنية الصرفية للكلمة، سواء بالإعلال، أو بالإبدال، أو بالحذف، أو بالقلب، أو بالإدغام، أو فكّ التضعيف[35].

34 - القسم الأول: سوابق من حيث النوع: سوابق خاصة بالأفعال/ سوابق خاصة بالأسماء/ سوابق تتصل بالحروف والأدوات/ سوابق مشتركة بين نوعين أو أكثر من الكلام.
القسم الثاني: سوابق من حيث الإفراد والتركيب: سوابق مفردة / سوابق مركبة.
القسم الثالث: سوابق من حيث الوظيفة التي تؤديها: سوابق: الاستفهام / الاستقبال / العطف / الجر / التشبيه / القسم / الابتداء / التعليل / والخلوص إلى زمن فعلي محدّد.
القسم الرابع: سوابق من حيث الزمن: سوابق تختص بالفعل الماضي / سوابق تختص بفعل الأمر / سوابق تختص بالفعل المضارع.
للمزيد، انظر: محمد نصر سيد، **السوابق واللواحق في اللغة العربية، حوليات آداب عين شمس**، عدد خاص (2017)، ص384-385.

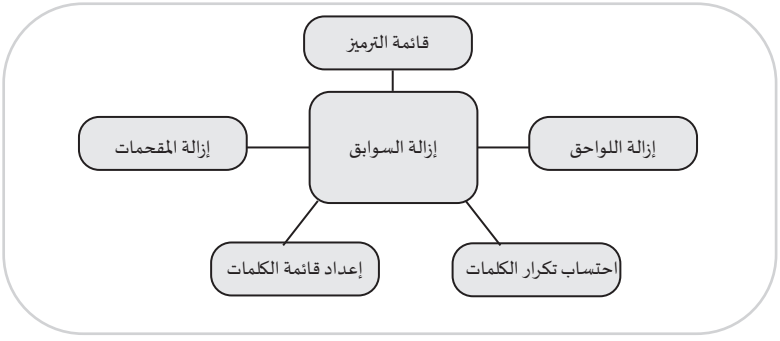
35 - تنقسم اللواحق الى الأقسام التالية:
القسم الأول: لواحق من حيث النوع وتشمل: لواحق تدلّ على المتكلم / لواحق تدلّ على المخاطب / لواحق تدلّ على الغائب.
القسم الثاني: لواحق من حيث العدد وتشمل: لواحق تدلّ على المفرد / لواحق تدلّ على المثنى / لواحق تدلّ على الجمع.
القسم الثالث: لواحق من حيث الزمن، تشمل: لواحق تدلّ على الزمن الماضي / لواحق تدلّ على الزمن المضارع / لواحق تدلّ على زمنين أو أكثر.
القسم الرابع: لواحق من حيث الإفراد والتركيب، وتشمل: لواحق مفردة / لواحق مركبة.
القسم الخامس: لواحق من حيث اللزوم والتعدّي، وتشمل: لواحق لا تدلّ على التعدّي / لواحق مفردة / لواحق مركبة.
للمزيد، انظر المرجع السابق، ص408-412.

أما الدواخل[36] "المقحمات"، فهي من العناصر التي تتوسط الجذر كالألف في اسم الفاعل، أو الواو في اسم المفعول، ولاصقة فاعل..انظر الجدول (4) حيث أودعنا أنموذجًا من مخرجات المحلّل الصرفي الآلي مع جذور الكلمات، مع السوابق واللواحق التي تبدو واضحة من مراجعة الكلمات وجذورها في الأنموذج المعروض بالجدول.

الجدول (4): أنموذج لمخرجات المعالجة المحوسبة العميقة (إزالة السوابق واللواحق ثم التحليل الصرفي لكلمات النص المنتخب)

تسلسل الكلمة	الكلمة	الجذر	تسلسل الكلمة	الكلمة	الجذر
283	تَحْصِيلِ	تَحْصِيل	293	تَبَرَّكا	تَبَرَّك
284	العُلُومِ	عِلْم	294	وَتَطَوُّعًا	تَطَوُّع
285	العَرَبِيَّةِ	عَرَبِيّ	295	فَلَمّا	لَمّا
286	حاجَةٍ	حاجَة	296	فَرَّغْتُ	فَرَّغ
287	إِلَيهِم	إِلى	297	مِن	مِن
288	إِلّا	إِلّا	298	تَحْصِيلِ	تَحْصِيل
289	مِن	مِن	299	العُلُومِ	عِلْم
290	حَيْثُ	حَيْثُ	300	العَرَبِيَّةِ	عَرَبِيّ
291	السِنْدِ	سِنْد	301	وَالشَرْعِيَّةِ	شَرْعِيَّة
292	عَنهُم	عَن	302	مِن	مِن

وعلى هذا الأساس يمارس محلل الصرفي الآلي دوره في تحليل قائمة الكلمات، فيسعى إلى تخليصها من السوابق والمقحمات، واللواصق ليبقي في القائمة المحوسبة على جذور الكلمات، وتكرار حضورها في النص. انظر الشكل (3).



الشكل (3): مراحل عملية احتواء عناصر نص كتاب كشاف اصطلاحات الفنون والعلوم[37].

ثم يصار إلى إزالة الكلمات التي لا تحمل معاني مستقلة ببنيتها اللغوية من قائمة قواعد بيانات الرموز أو الكلمات المشتقة Stemmed Words. ويطلق في بعض الأحيان على هذه الكلمات مستبعدات الفهرسة.

ويطلق اصطلاح كلمات الوقف على البنيات النحوية التي لا صلة لها بالمحتوى المعرفي والمعاني المودعة في النص، وإنما تسهم بربط المفردات لبناء الجملة وفق القواعد الصرفية والنحوية.

وتقع في قائمة كلمات مستبعدات الفهرسة كل من حروف الجرّ، وحروف العطف والاقتران، والضمائر المنفصلة، وظروف الزمان والمكان، وغيرها كثير...، انظر الجدول (5) الذي أودعنا فيها نماذج من مستبعدات الفهرسة.

36 - عيسى العزري، "دلالة اللواصق في اللغة العربية"، **مجلة اللغة العربية**، الجزائر، مج21، ع 43 (2019)، ص88.

37 - الشكل أعد من قبل الباحث.

الجدول (5): أمثلة لكلمات التوقف التي يطلق عليها مستبعدات الفهرسة، وقد استبعدت من مدونة معجم كشاف اصطلاحات الفنون والعلوم

،	أجل	أغسطس	أنا	أى
ء	أجمع	أفريل	أنبأ	أي
ءَ	أحد	أفعل به	أنت	أيا
آ	اخبر	أفٍّ	أنتم	أيار
آب	أخذ	أقبل	أنتما	أيضا
آذار	أخو	أكتوبر	أنتن	أيلول
آض	أخٌ	ألْ	أنتِ	أين
آل	أربع	ألا	أنشأ	أيّ
آمين	أربعاء	ألف	أنه	أيّان
آناء	أربعة	ألفى	أنَّ	أُفٍّ
آنفا	أربعمئة	أم	أنّى	ؤ
آه	أربعمائة	أما	أهلا	إحدى
آهًا	أرى	أمام	أو	إذ
آهٍ	أسكن	امامك	أوت	إذا
آهِ	أصبح	أمامكَ	أوشك	إذًا
أ	أصلا	أمد	أول	إذما
أبدا	أضحى	أمس	أولئك	إذن
أبريل	اطعم	أمسى	أولاء	إزاء
أبو	أعطى	امّا	أؤلالك	إلى
أبٌ	أعلم	إن	أوّهْ	إلي

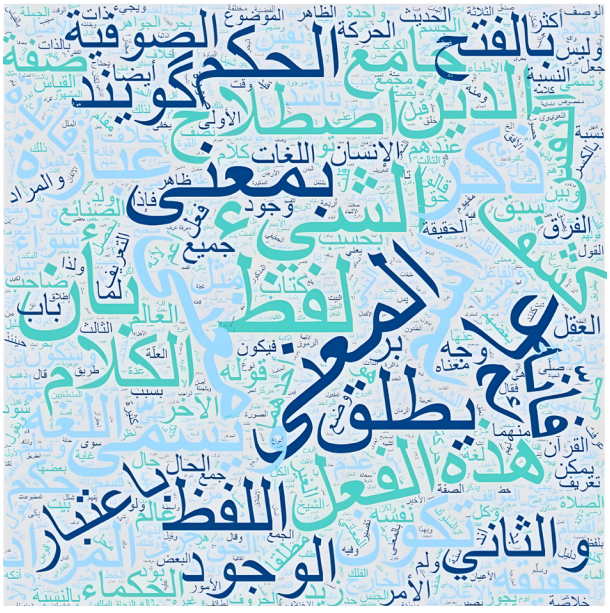
ولا تقتصر كلمات الوقف على هذه الفئات بحسب، وإنما تتوسع دائرتها أو تتقلص بحسب طبيعة النص الذي نروم معالجته حاسوبيًا لاستخلاص المادة المعرفية المودعة في محتواه. فيصار في كثير من الأحيان إلى المعالجة الصرفية للكلمات لإزالة اللواصق بالكلمات، أو اختيار جذر الكلمة أو إحدى بنياتها الصرفية والتي تتوافق مع المادة المعرفية التي ننقّب عنها في مادة النص. فتلتحق، على هذا الأساس، بكلمات التوقف الكلمات الشائعة بين النصوص والتي لا تحمل دلالة تؤثر على مضمون النص[38].

من أجل هذا فإن أي عملية حوسبة للنص تقتضي إعداد ملف يضم جميع كلمات التوقف على شكل قاعدة بيانات يمكن أن تمارس بواسطتها عملية التقاط هذه الكلمات من النص المحوسب لكي نقلل من حجم مفردات النص، ونمنح لأنموذجنا المحوسب فرصة الكشف عن المعاني والدلالات المودعة في الأسماء، والأفعال، والصفات، بعيدًا عن التأثيرات التي قد تصاحب حضور كلمات التوقف فتؤثر سلبًا على سريان المعنى وترابط الدلالات بين الكلمات الأساسية.

38 - ترتبط هوية هذه الكلمات بالحقل الذي ينتمي إليه النص، فهناك الكثير من الأفعال والصفات، والضمائر التي لا تمتلك أهمية في حقل من الحقول المعرفية بينما تمتلك أهمية في حقول أخرى. وتعتمد عملية انتخاب هذه المفردات على النهج الذي يتبناه الباحث في عملية التنقيب بمادة البيانات.

ثامنًا: المعالجات المحوسبة لمدونة كتاب كشاف اصطلاحات الفنون والعلوم

بعد استكمال عملية استيراد النص الرقمي لمدوّنة كتاب كشاف اصطلاحات الفنون والعلوم[39]، حاولنا أن نشكّل تمثلًا مرئيًا لمحتوى الكتاب بعد إجراء عملية المعالجة المحوسبة الأولية، فبدأنا بتوظيف تقنية سحابة الكلمة Word Cloud والتي توظف تقنيات الحوسبة الذكية لتمثيل محتوى النص رسوميًا على أساس تكرار حضور الكلمات في مادته، فحصلنا على الشكل (4).



الشكل (4): سحابة الكلمة في مدونة كشاف اصطلاحات الفنون والعلوم بعد إجراء العمليات الأولية لمادة النص[40].

ورغم حضور كلمات واصطلاحات مهمة في واجهة السحابة (مثل: المعنى/ اللفظ/ الوجود/ الحكم) إلا أن هذا المستوى من المعالجة المحوسبة الرقمية والتمثيل المرئي لا يعد كافيًا لوصف محتوى المدونة وشبكات الدلالة المقيمة بين الاصطلاحات المودعة فيه.

1. **المعالجات المتقدمة لمادة مدونة كشاف اصطلاحات الفنون والعلوم**

بعد أن أتممنا عملية تنقية النص الأولية من مستبعدات الفهرسة، وتمثيل مادة المدونة بواسطة سحابة الكلمة، أصبحت المدونة اللغوية لكشاف اصطلاحات الفنون والعلوم جاهزة لإجراء المزيد من عمليات المعالجة المحوسبة لمادتها؛ لإنتاج المزيد من أنماط التمثيل الرسومي ومحاولة الظفر بشبكات المفاهيم التي استودعت فيه.

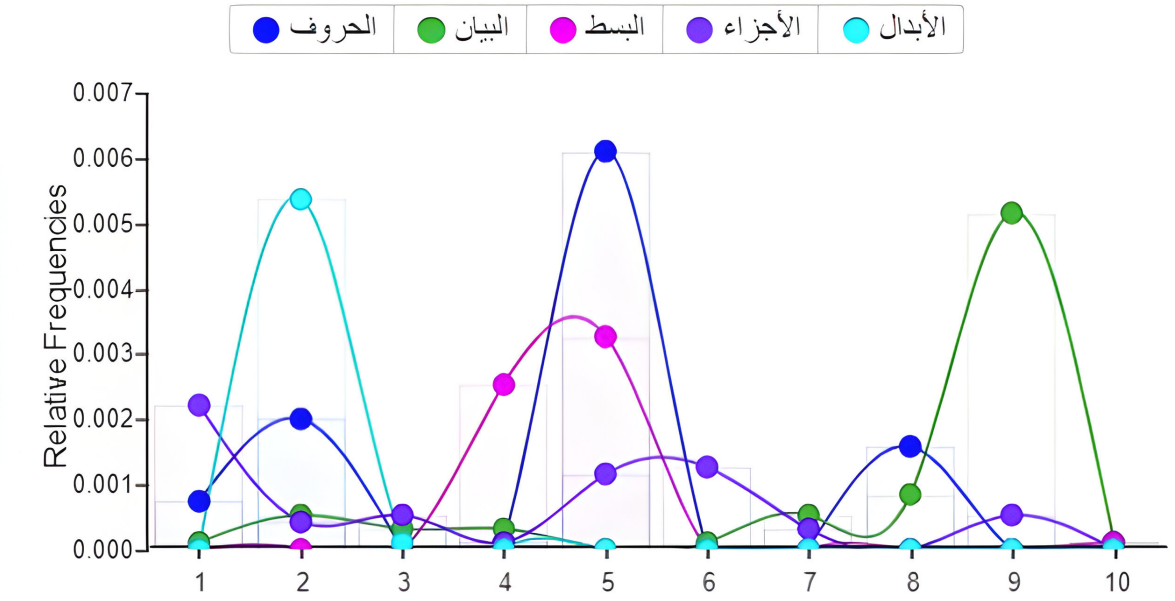
وقد اعتمدت المعالجة إلى مبدأ انتخاب فقرة حرف الباء، وكذلك فقرتي اصطلاح العلم، واصطلاح المعرفة لتكون مادة لعمليات المعالجة المحوسبة المتقدمة للمدونة.

بداية أعددنا أنماطًا رسومية لبيان أكثر الاصطلاحات حضورًا في الفقرات الثلاث بعد تقسيم مادة النص إلى أكثر من مقطع. وقد استمدت هذه الأنماط الرسومية مادتها من المعالجة المحوسبة التي تولّت مسألة إحصاء أعداد المفردات في كل قطاع من قطاعات الفقرات الثلاث.

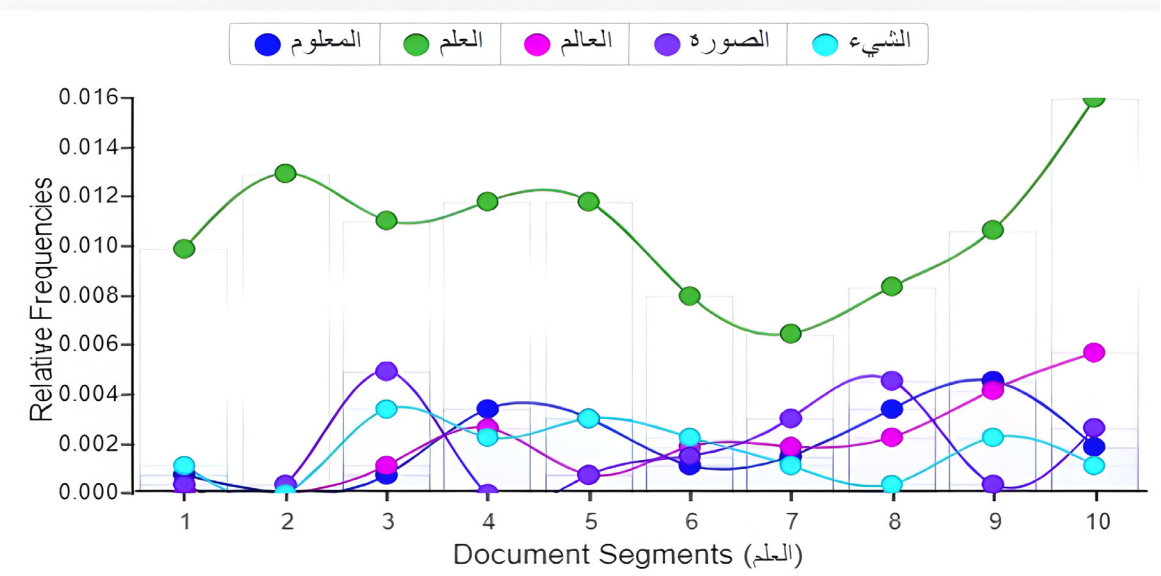
39 - محمد بن علي القاضي التهانوي، **كشاف اصطلاحات الفنون والعلوم**، تحقيق علي دحروج (بيروت: مكتبة لبنان ناشرون،1996)، النسخة الرقمية.

40 - الشكل من إعداد الباحث.

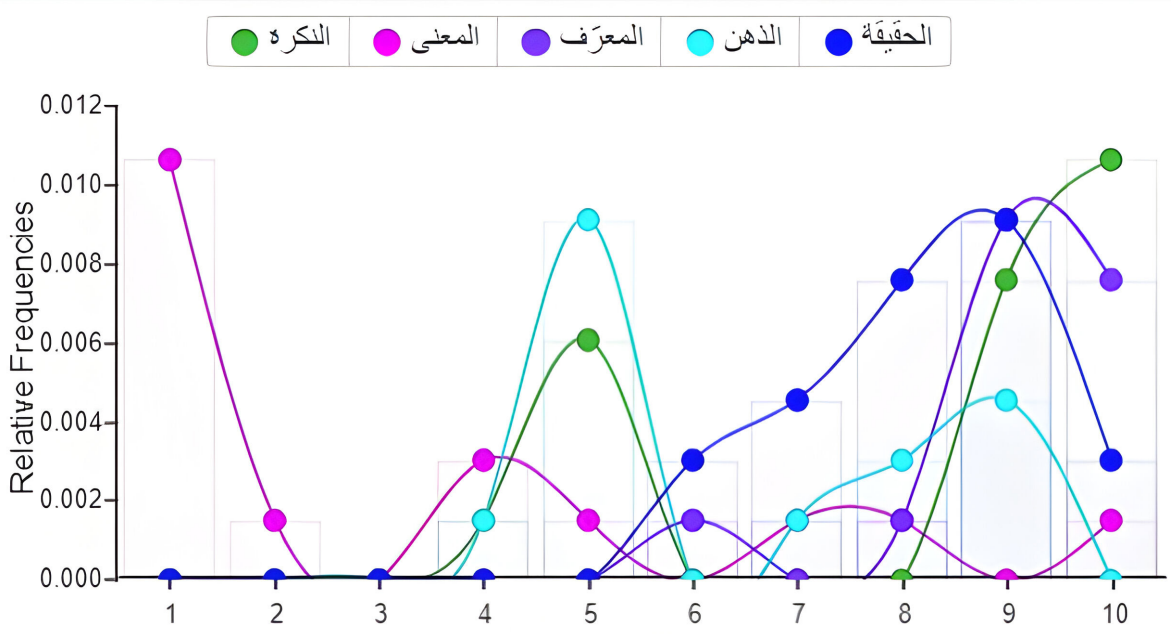
ويظهر في الشكل (5) أكثر الاصطلاحات حضورًا في مقاطع فقرة حرف الباء، أما الشكل (6) فقد تولى اظهار الاصطلاحات الأهم في فقرة العلم، في حين أن الشكل (7) قد أظهر أكثر الاصطلاحات حضورًا في فقرة المعرفة من كتاب كشاف اصطلاحات الفنون والعلوم.



الشكل (5): أكثر الاصطلاحات حضورًا في فقرة الباء.

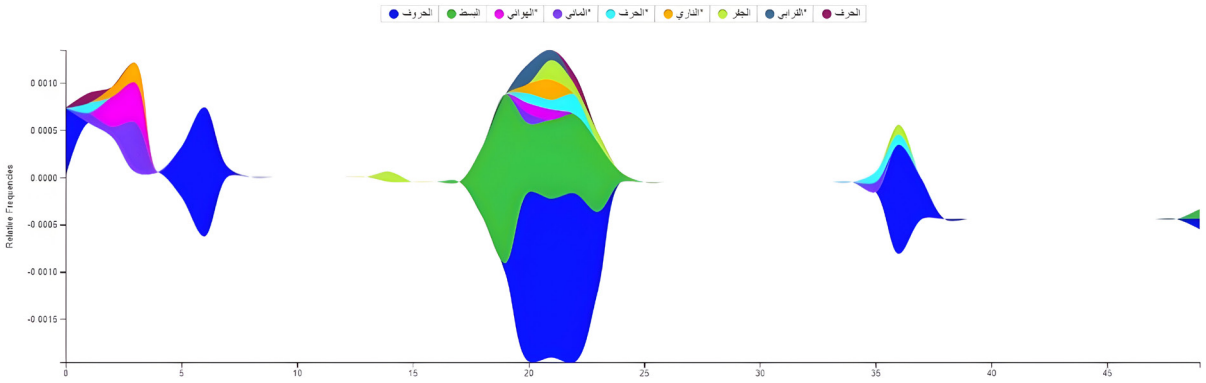


الشكل (6): أكثر الاصطلاحات حضورًا في فقرة العلم.



الشكل (7): أكثر الاصطلاحات حضورًا في فقرة المعرفة[41].

ولكي تكتمل دائرة فهمنا لمحتوى الفقرات الثلاث حاولنا أن نتتبع تدفق الإصطلاحات على طول مقاطع الفقرات الثلاث التي انتخبت لتمثيل محتواه[42]. انظر الأشكال (8، 9، 10).



الشكل (8): تدفق اصطلاحات علوم منتخبة على عموم نص فقرة حرف الباء[43].

ويبدو واضحًا من الشكل (8) أن موضوع الحروف قد احتل مكانة بارزة في خطابه، وقد ناقشه مع مسائل متعددة شملت أنواع الحروف الأربعة (المائية، الهوائية، النارية، والترابية) بإسهاب (يبدو واضحًا من حجم الشكل)، كما أنه ناقش موضوع الجفر باهتمام أقل في المقطع[35].

41 - الأشكال أعدّها الباحث من البيانات التي حصل عليها من التحليل اللساني للمحتوى.

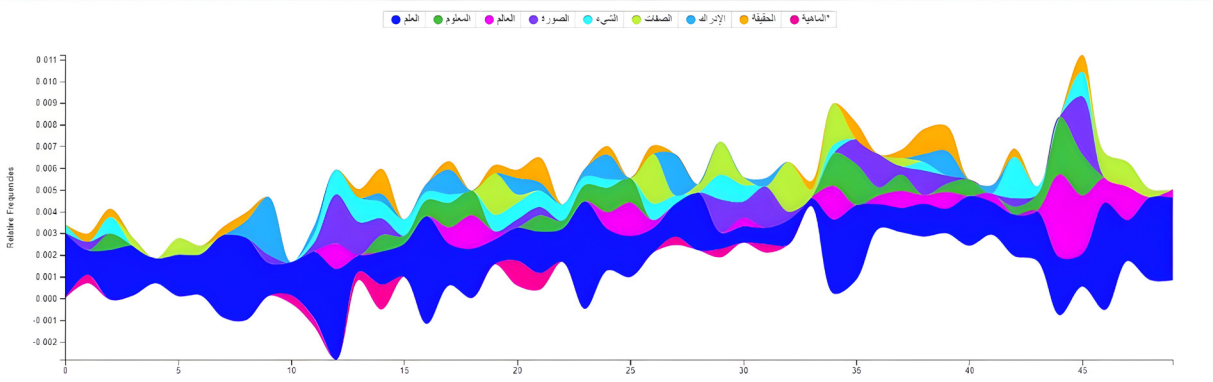
42 - قسمنا كل فقرة إلى مجموعة من المقاطع التي حددها البرنامج المحوسب، ثم تتبعنا تدفق أهم الاصطلاحات (الأكثر ورودًا) على طول مادة النص في الأشكال المذكورة ليتبين لنا حضور كل منها في مادة خطابه المعرفي.

43 - يمثل المحور السيني للشكل الرسومي X-Axis موقع المفردات في كامل محتوى النص، أما المحور الصادي Y-Axis فيمثل عدد تكرارات الكلمة ومستوى حضورها. وتوفر لنا هذه الأشكال الرسومية مستوى تجمّع الكلمات (بحسب الموضوع المنتخب) والتي تعدّ مؤشرًا على مستوى اهتمام المصنّف بموضوع دون آخر.

إلا أنه رغم اهتمام المؤلف باصطلاحات محددة ذات صلة بتوجهاته العقدية والثقافية إلا أنه قد ترك فجوات واضحة عند معالجة أهم هذه الاصطلاحات. فقد غاب علم الحروف عن اهتماماته في المقاطع (6-18)، (25-34).

ويبدو واضحًا من الشكل (9) طبيعة تدفق مسائل العلوم وتفريعاتها في فقرة العلم والتي امتدت على طول الفقرة، وشملت مسائل تخص: العلم، والمعلوم، والعالم، والأشياء، والصفات، والإدراك والماهية، وبمستويات متباينة من المعالجة على طول مقاطع النص.

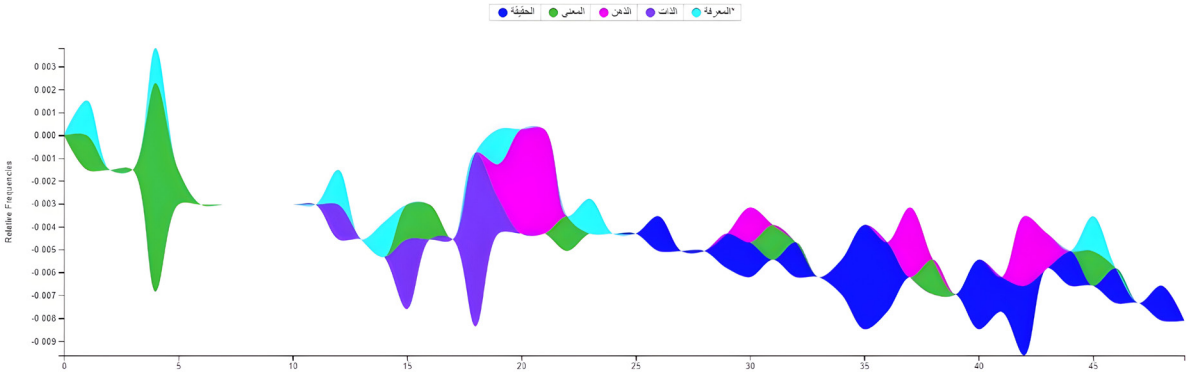
إلا أنه من الواضح استمرار حضور اصطلاح العلم ومفرداته على جميع مقاطع الفقرة دون أن يترك فجوة في مجال معالجاته للموضوع.



الشكل (9): تدفق اصطلاحات علوم منتخبة على عموم نص فقرة العلم[44].

أما الشكل (10) فيظهر فيه تدفق موضوعات خطاب التهانوي في فقرة المعرفة، والتي تضمنت مسائل تخص العلاقات الرابطة بين المعرفة والحقيقة، والمعنى، والذهن، والذات.

وقد تباينت مفردات خطابه عند مناقشة هذه المسائل بين حضور لمصطلح بكثافة، أو غيابه على طول مقاطع الفقرة، والتي تشكل تحديًا أمام الباحثين لتبرير هذه الأمور من خلال قراءة النص عن قرب.



الشكل (10): تدفق اصطلاحات علوم منتخبة على عموم نص فقرة المعرفة[45].

44 - الشكل أعدّه الباحث من البيانات التي حصل عليها من عمليات حوسبة مفردات الفقرة.

45 - الشكل أعده الباحث من حوسبة مفردات فقرة المعرفة.

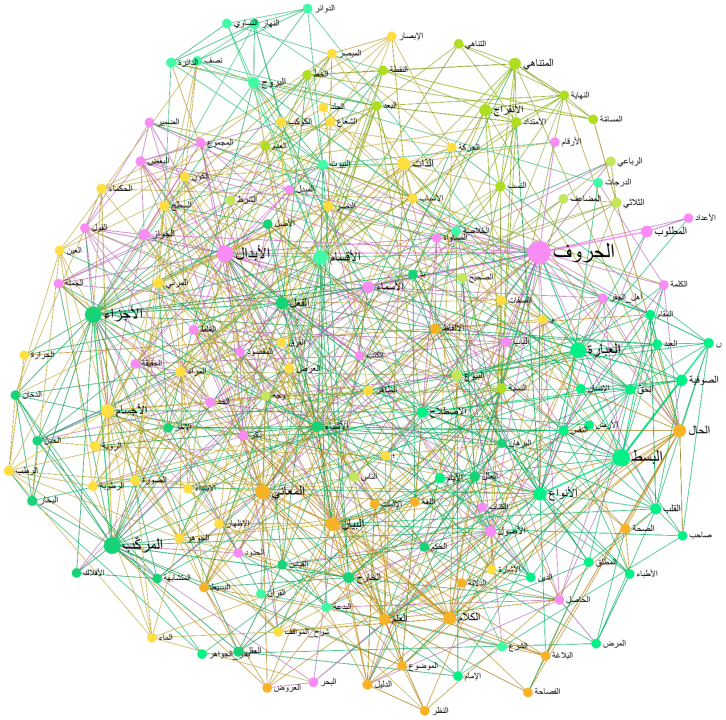
2. التمثيل الشبكاتي للمفاهيم في مدونة كشاف اصطلاحات الفنون والعلوم

يستخدم التمثيل الشبكاتي لنسيج النص Text Network Visualization لتعميق فهمنا بالبنية الدلالية والمفاهيمية لمادة النص، مع بيان طبيعة الروابط المقيمة بين المصطلحات، وتحديد أهم المفاهيم المودعة فيه، وأنماط حضور الكلمات والمصطلحات وترابطاتها وتكراراتها فيه.

ويرتكز نهج التمثيل المرئي لشبكة النص إلى مبدأ التعامل مع مادة النص المهيكل بوصفه شبكة يتألف لبابها من مجموعة من العقد Nodes التي يمكن تمثيلها بواسطة الكلمات والمصطلحات الموجودة بالنص. أما العلاقات والروابط الدلالية والمفاهيمية بين الكلمات والمصطلحات فتمثّل بواسطة الحافات Edges التي تشكّل الروابط التي تتشكل منها مادة نسيج النص.

وقد تناولنا مادة نصوص الفقرات الثلاثة المنتخبة من كشاف اصطلاحات الفنون والعلوم بمزيد من عمليات المعالجة المحوسبة لتمثيلها على شكل نسيج شبكاتي يبرز لنا معمارية شبكة النصوص الثلاثة.

بداية قمنا بتمثيل شبكة نص فقرة حرف الباء لتحديد مستويات حضور الاصطلاحات والمفاهيم فيها، مع أنماط العلاقات التي تربط فيما بينها لتشكيل مادة الخطاب المطروح في هذه الفقرة[46]. انظر الشكل (11).



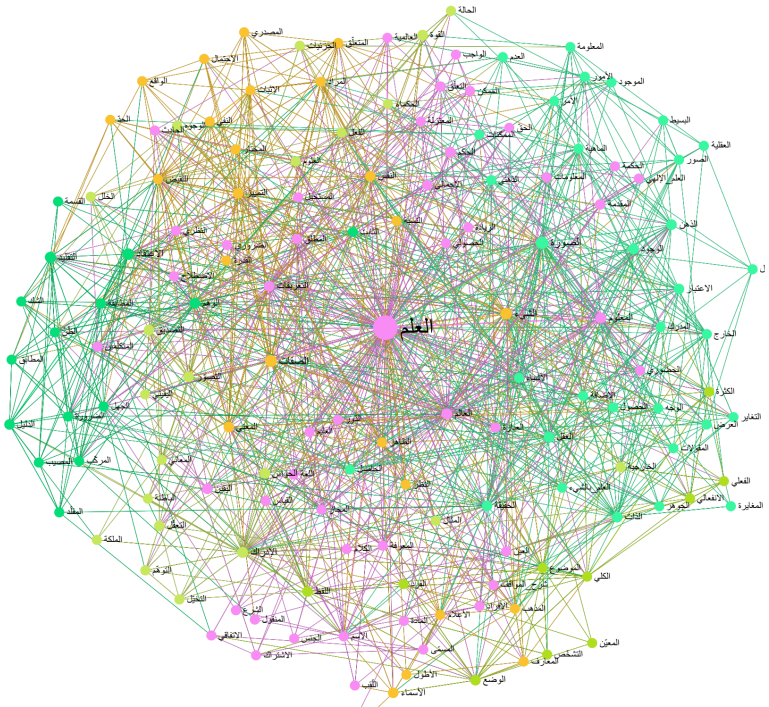
الشكل (11): النسيج الشبكاتي لمحتوى فقرة حرف الباء[47].

46 - يبدو واضحًا وجود أكثر من نمط للروابط المقيمة بين المفردات المنتشرة في كل فقرة من الفقرات التي انتخبناها من الكشاف. وقد حاولنا اختيار ألوان مختلفة لوصف طبيعة العلاقات الرابطة والتي تؤشر إلى مستوى المعنوية في الارتباط Significance؛ حيث يمكن على أساس شبكة العلاقات البادية في الشكل تحديد طبيعة العلاقات وتسلسلها بين مختلف المفردات التي استخدمها المصنّف في كتابه.

47 - النسيج الشبكاتي من إعداد الباحث.

ويبدو واضحًا لنا من معاينة نسيج مادة هذه الفقرة أن التهانوي (مؤلف الكتاب) قد جعل من فضاء هذه الفقرة بابًا للحديث عن علم الحرف، فظهر حضور اصطلاح الحروف واضحًا في عقد النسيج، وتشكلت من عقد النسيج ثلاثة عناقيد من العقد المترابطة (اللون الوردي، والأصفر، والأخضر) والتي تؤشر إلى مستويات تجاور مفرداتها داخل نسيج نص هذه الفقرة[48].

أما مادة فقرة العلم فقد حاولنا تمثيل نسيجها الشبكاتي، فأودعناها في الشكل (12).



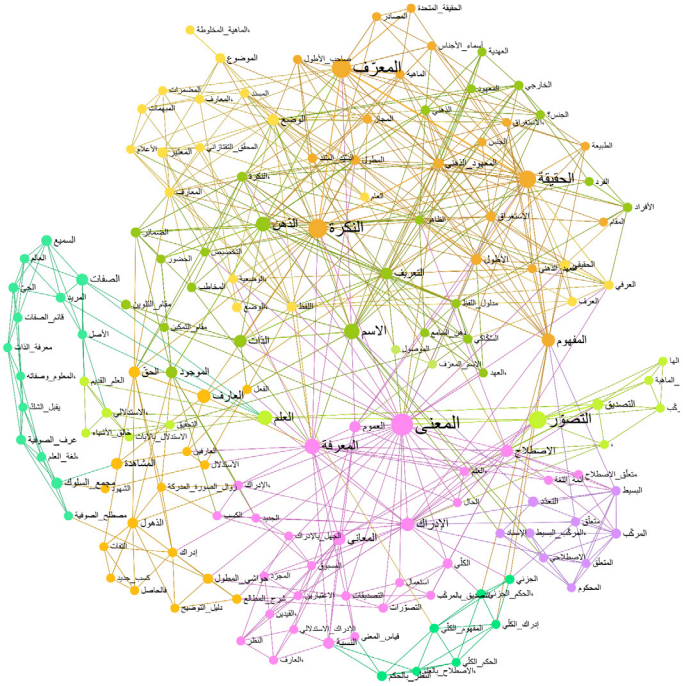
الشكل (12): النسيج الشبكاتي لمحتوى فقرة العلم[49].

وقد تألف نسيج هذه الفقرة من ثلاثة أنواع من العقد بألوانها المشابهة لنسيج فقرة حرف الباء. ويبدو واضحًا اهتمام المؤلف باصطلاح العلم الذي احتلّ مكانة مركزية في النسيج وتفوق بمستوى تكراراته فتفوق على بقية الاصطلاحات بحيث لا نكاد نعثر على مصطلح امتلك حضورًا مقاربًا لحضوره، وأن بقية الاصطلاحات تصب في بوتقته المفاهيمية.

وقد قمنا بالعمل ذاته مع نص فقرة المعرفة بالكشاف، وأودعنا مفرداتها في النسيج الشبكاتي الرقمي الذي يبدو واضحًا في الشكل (13).

48 - تشير أحجام عقد النسيج الشبكاتي إلى مستوى تكرارها داخل كل فقرة من هذه الفقرات، بينما يؤشر لونها إلى انتمائها الى عنقود المعاني والارتباطات التي تلتحق بها كل مفردة من مفردات الفقرات الثلاث.

49 - النسيج الشبكاتي من إعداد الباحث.



الشكل (13): النسيج الشبكاتي لمحتوى فقرة المعرفة[50].

وعلى عكس اصطلاح العلم، نرى أن اصطلاح المعرفة قد تراجع في حضوره قبالة اصطلاحات أخرى اهتم بها المؤلف أثناء إيراده هذه الفقرة، فبرز اصطلاح المعنى بوصفه أكثر الاصطلاحات حضورًا، ثم جاءت الاصطلاحات: الحقيقة، المعرّف، النكرة، والتصور في مستويات حضورها قبل اصطلاح المعرفة الذي لم يتوسع في مناقشته كثيرًا.

إن معاينة التمثيل الرسومي للفقرات الثلاثة المنتخبة سيمنح القارئ والباحث فرصة تحديد أهم المفاهيم المطروحة فيها، ومستوى الاهتمام بالاصطلاحات دون الحاجة إلى مطالعتها مما يوفر له وقتًا ثمينًا، كما سيمنحه فرصة لتحديد الفجوات المفاهيمية في عمليات الطرح ورؤية مبصرة لإجراء بحوث قادمة عند تتبع مسارات الارتباطات، ومستويات حضور المصطلح أو غيابه في الفضاء النسيجي لكل فقرة[51].

3. تشكيل شبكة المفاهيم في الفقرات المنتخبة

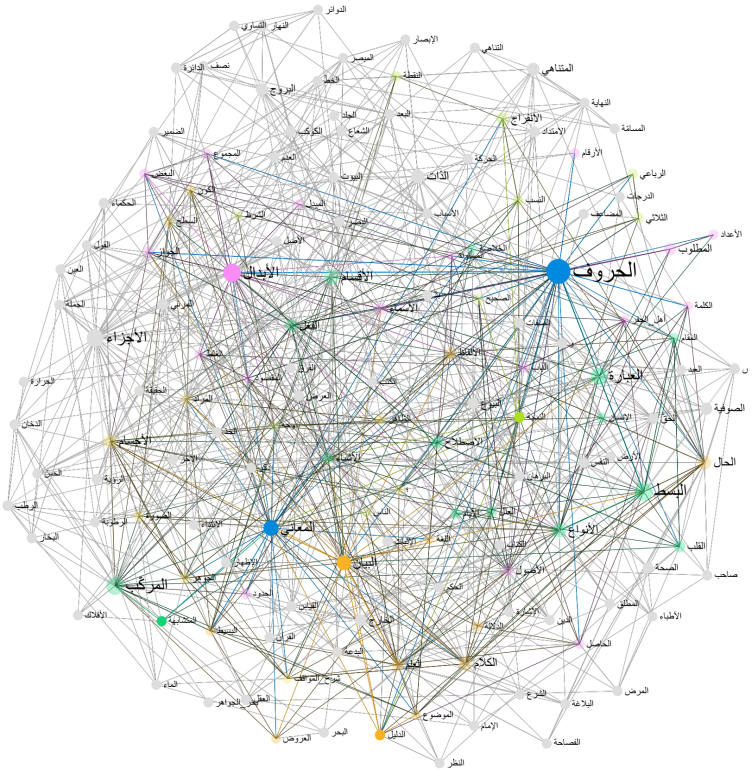
تأتي المرحلة اللاحقة في استثمار الحوسبة الذكية لسبر محتوى الفقرات من خلال التنقير عن شبكة العلاقة الرابطة بين المفردات التي استوطنت في كل فقرة من هذه الفقرات، لكي نوفر للقارئ أو الباحث فرصة تتبع الارتباطات المفاهيمية لاصطلاحات محددة.

50 - النسيج الشبكاتي من إعداد الباحث.

51 - إن هناك متغيرين أساسيين يمكن استثمارهما في تحديد هوية المفاهيم المودعة في النص. الأول حجم العقدة الشبكاتية التي تمتلكها المفردة، والثاني لون العقد العلاقات الرابطة فيما بينها. فالنظرة الأولى إلى الشكل (13) تظهر لنا أن هناك أربع مفردات اهتم بها المؤلف في هذه الفقرة هي: المعنى / المعرّف/ النكرة (بالمرتبة الأولى) ثم الحقيقة / التصور/ والمعرفة (في المرتبة الثانية، بناء على الحجم الذي تمتلكه عقد هذه الكلمات في نسيج النص. كذلك فإن العقد والعلاقات التي تربط فيما بينها تؤشر إلى انتماء هذه المفردات إلى عنقود من المعاني التي ناقشها المؤلف في هذه الفقرة دون غيرها. وهنا يأتي دور المتخصص في العلوم المعجمية واللغة لتفسير هذه العلاقات، كما أنها بالوقت ذاته ستجعله على بينّة حول مستوى اهتمام المؤلف بكل مفردة موضوعية في أي فقرة منتخبة للتحليل المحوسب من فقرات الكتاب.

أ. شبكة المفاهيم في فقرة حرف الباء:

لتكملة الشوط الذي قطعناه مع دراسة تدفق مادة خطاب التهانوي في فقرة حرف الباء، حاولنا رسم شبكة العلاقات الرابطة بين اصطلاح الحروف الذي يحتلّ مكانة بارزة في علم الحرف والجفر، وبقية المواضيع التي أولاها المؤلف اهتمامًا في هذه الفقرة. انظر الشكل (14).



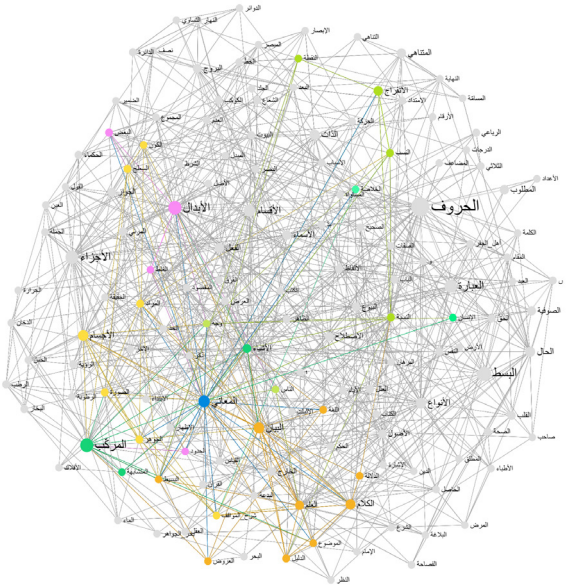
الشكل (14): المفاهيم ذات الصلة باصطلاح الحروف والارتباطات المقيمة معها[52].

لقد تشابكت خطوط المسائل المرتبطة بالحروف عند التهانوي في مادة النسيج الشبكاتي أعلاه، وتلاحمت مع مسألة المعاني بعد أن توغّل حضورها مع مسائل تخص خصائص الحروف، وحضورها في مسألة بسط الحروف لاستنطاقها بالعلوم المختلفة، كذلك علاقتها الحميمة مع الأعداد والأرقام التي تستوطن معها في فضاء علوم الحرف والجفر.

وتشكل هذه الشبكة بتفريعاتها المعقدة تحديًا كبيرًا أمام الباحثين لتبرير أنماط ومستويات الارتباطات المقيمة بين الحروف وبين بقية الاصطلاحات وتفتح بابًا واسعًا للتفسير.

من جهة أخرى عندما حاولنا أن نسبر نسيج العلاقات بين اصطلاح المعاني وما يرتبط به من مفاهيم متنوعة حصلنا على الشكل (15). فوجدنا أن المعالجة المفاهيمية للمعاني داخل حدود هذه الفقرة قد ارتبطت بمسائل شتى منها: اللغة، والأسماء، والمراد، والجوهر، وغيرها من التفريعات المتعددة التي توحي بتداعيات المفهوم لدى التهانوي وانعكاساتها على مادة الخطاب المطروح في فقرة الباء.

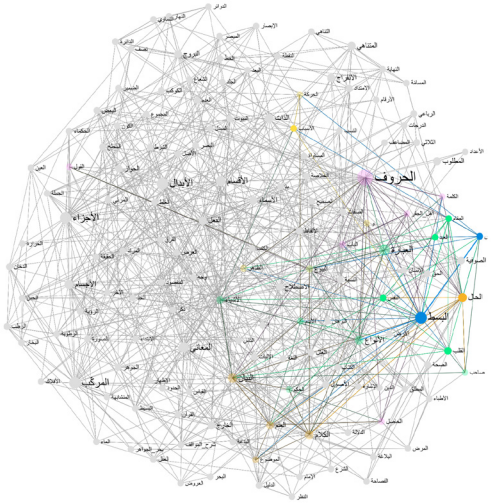
52 - النسيج الشبكاتي من إعداد الباحث.



الشكل (15): المفاهيم ذات الصلة باصطلاح المعاني والارتباطات المقيمة معها[53].

وأخيرًا، حاولنا الكشف عن التوجهات الصوفية في خطاب التهانوي واستثماره لها في فقرة حرف الباء، فوجدنا أنه قد وجد من اصطلاح البسط (من الاصطلاحات التي تبدأ بحرف الباء) بابًا لمناقشة كثير من المسائل التي تناقش في فضاء المعرفة الصوفية فأورد اصطلاح القبض المقابل له، وبدأ بتشكيل نسيج من العلاقات ذات الصلة بهذا الموضوع، فظهرت لنا في مادة النسيج ترابطات مع مصطلحات ذات صلة مثل: الحق، والصوفية، والعبد، والمقام.

ويمكن أن نتوسع في هذه المسائل متى عمدنا إلى انتخاب أي فقرة من هذه الفقرات؛ حيث ستبرز أمامنا تفاصيل لشبكات مفاهيمية ثانوية تربط بين القبض أو البسط أو كليهما مع الاختيارات الجديدة، مما سيمنحنا فرصة للتوغل بعمق أكبر في ماهية نسيج النص.



الشكل (16): المفاهيم ذات الصلة باصطلاحي القبض والبسط والارتباطات المقيمة معها[54].

53 - النسيج الشبكاتي من إعداد الباحث.

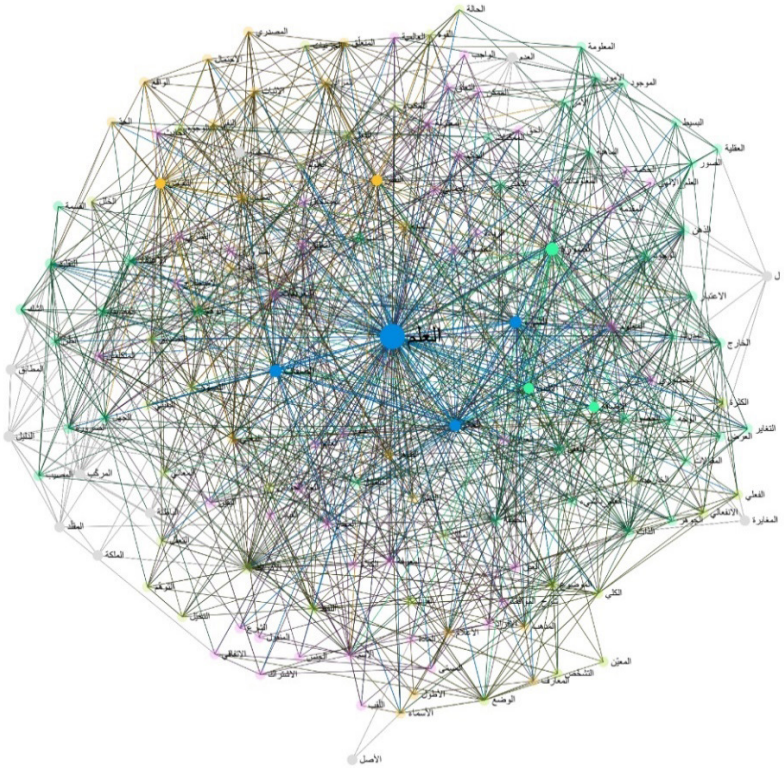
54 - النسيج الشبكاتي من إعداد الباحث.

ب. شبكة المفاهيم في فقرة العلم:

وجهنا اهتمامنا إلى المفاهيم المقيمة في فقرة العلم من كتاب كشاف اصطلاحات العلوم، فوجدنا أن اصطلاح العلم قد شكّل لباب هذه الفقرة، فحاولنا أن نسبر العلاقات التي يشكلها مفهوم العلم مع اصطلاحات أخرى مثل: العالم، والشيء، والصفات، فتولّد لدينا النسيج الشبكاتي الظاهر في الشكل (17).

فوجدنا أن هناك ارتباطات حميمة بين العلم والمعرفة، والتصوّر، واللفظ، والكثرة، والتغاير، والاحتمال، والجزئيات، وغيرها. بالمقابل فإن حضور العالم مع مفردة العلم تتولد عنه امتدادات جديدة في النسيج الشبكاتي مثل: المعلوم والحكم، والعين، والاصطلاح.

أما إضافة الشيء والصفات إلى مادة النسيج الخاصة بمفردة العلم فتتولّد عنه معان جديدة من خلال الارتباط بمفردات جديدة مثل: النفس، والتمييز، والنقيض. وهكذا تتكاثر الترابطات وتتداعى المعاني كلما حاولنا التوّغل أكثر في نسيج مادة الخطاب المطروح في فقرة العلم.



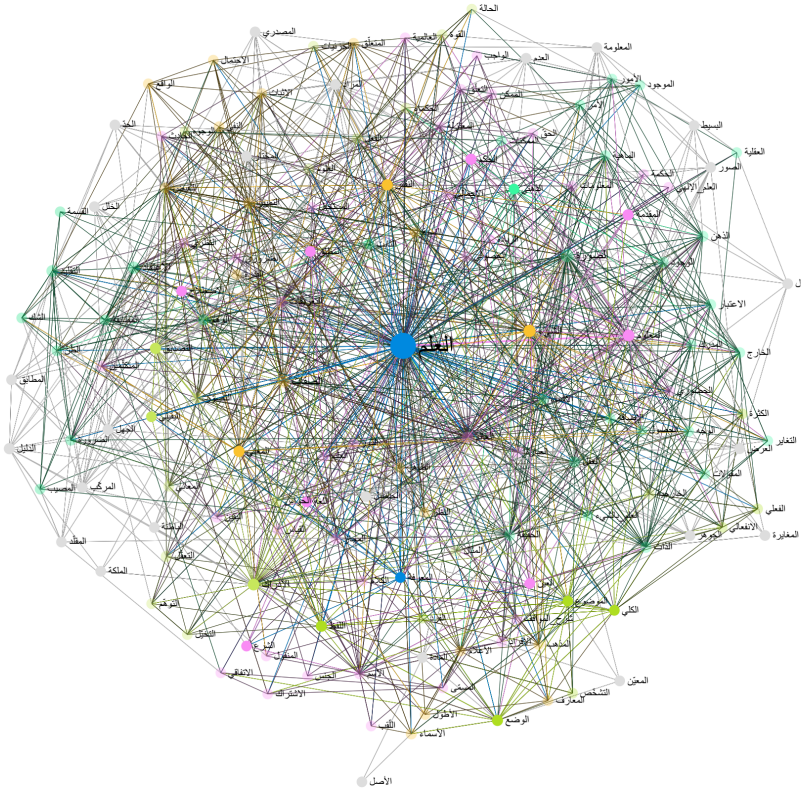
الشكل (17): العلم، والعالم، والشيء، والصفات[55].

ونختم حضورنا في مادة نسيج فقرة العلم بتحليل شبكة العلاقات الرابطة بين العلم والمعرفة، والتي حاولنا إيداعها في الشكل (18).

ونلاحظ وجود تعقيد كبير في نسيج شبكة المفاهيم الرابطة بين العلم والمعرفة لدى التهانوي في فقرة العلم، فثمة

55 - النسيج الشبكاتي من إعداد الباحث.

ارتباطات حاضرة مع: الكلي، والمجاز، واللغة، واليقيني، والتصور، والصفات، والظن، والمتكلمين، والاعتقاد، والظاهر.... وهي محاور رئيسة تفتح أمامنا المزيد من التساؤلات التي ستدفع بنا لا محالة إلى التنقير عن المزيد من خيوط النسيج المعرفي المقيم في هذه الفقرة.



الشكل (18): شبكة المفاهيم التي تنتجها العلاقة بين العلم والمعرفة[56].

ج. شبكة المفاهيم في فقرة المعرفة:

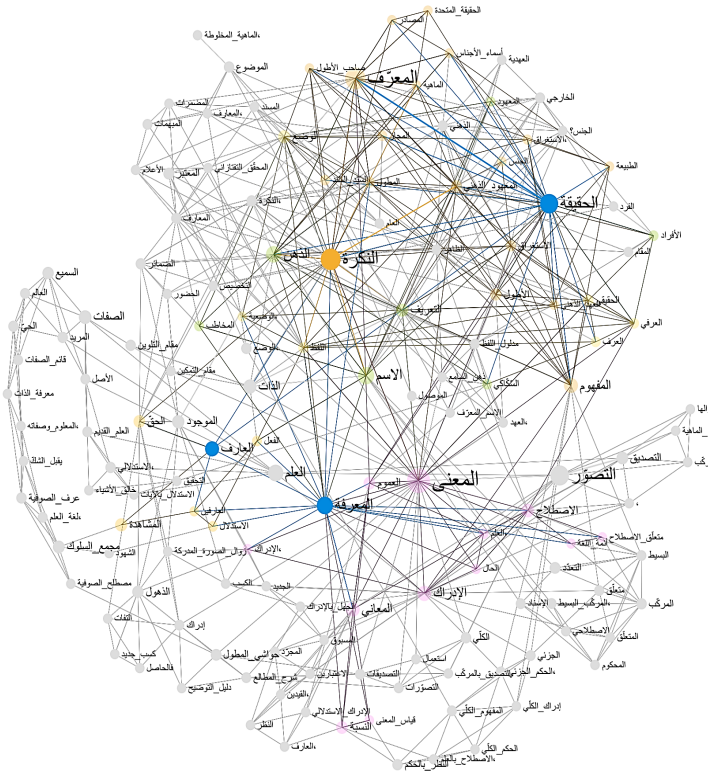
استخدمنا النهج ذاته الذي اعتمدناه في الفقرتين السابقتين في هذه الفقرة، وحاولنا انتخاب بعض الاصطلاحات وتوجيه دفّة حوسبة شبكة النص للإفصاح عن طبيعة الارتباطات القائمة فيما بينها. فحاولنا التنقير في طبيعة العلاقة القائمة بين: المعرفة. العارف. الحقيقة فحصلنا على النسيج الشبكاتي لهذه العلاقة كما في الشكل (19).

فظهر لنا أن المعرفة ترتبط في معالجة التهانوي لمفرداتها بالعارفين، واللغة، والمشاهدة، وزوال الإدراك. أما بالنسبة للعارف فإن ارتباطاته تتجه نحو الحق، والعارفين، والحقيقة، ومسائل أخرى.

أما بالنسبة لارتباطات الحقيقة فتتوجه صوب: المعرّف، والاستغراق، والمفهوم، وذهن السامع.

وينشأ عن تلاحم العلاقة بين الاصطلاحات الثلاثة والتفريعات الأساسية المرتبطة بها مجموعة ثرية من المفاهيم التي يمكن استنباطاها من نسيج النص، أو تداعيات المعنى الظاهرة في مادته.

56 - النسيج الشبكاتي من إعداد الباحث.



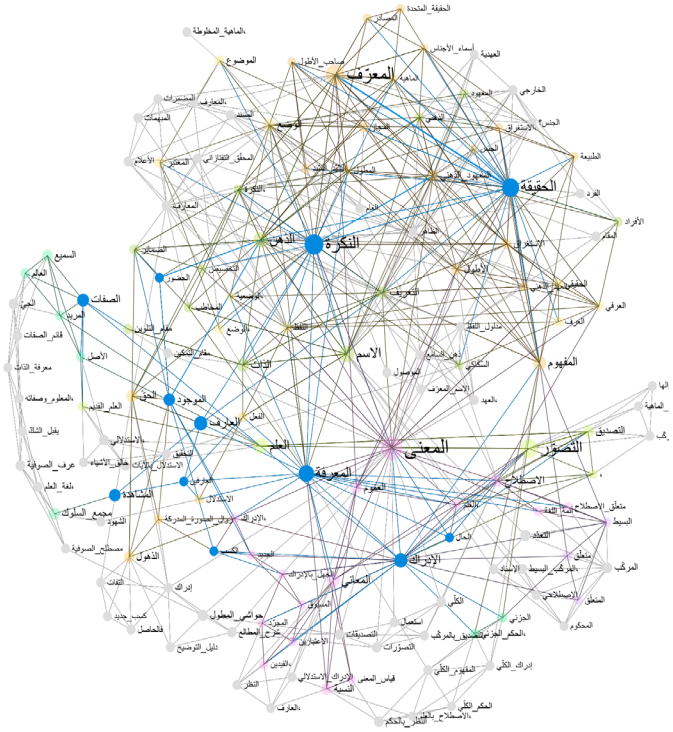
الشكل (19): العارف، المعرفة، الحقيقة[57].

أما الشكل (20) فحاولنا أن نختتم فيه رحلتنا مع شبكة نصوص معجم كشاف اصطلاحات الفنون والعلوم، فحاولنا الكشف عن النزعة الصوفية لدى التهانوي والتي أفصح عنها في معالجة فقرة المعرفة. فانتخبنا اصطلاحات ذات صلة بأهل التصوف تخص أفرادهم، وتفصح عن توجهاتهم في تعريف المعرفة وترابطاتها مع بقية الاصطلاحات التي تستوطن في الرقعة الجغرافية لاصطلاح المعرفة.

لقد تكوّن أمامنا نسيج فريد بترابطاته بات يربط المعرفة مع العارف، والحقيقة، والعارفين، والأفراد، والإدراك، والكسب، والمقام، والمشاهدة، والحضور، والحال. وجميع هذه الاصطلاحات تشكل نواة المعرفة الصوفية.

لقد اقتصرنا في تشكيل الشبكة الظاهرة على أهم الاصطلاحات، بينما توجد طبقات متعددة من شبكات العلاقات التي يمكن أن نظهرها متى ما وجهنا دفة اهتمامنا إلى مساحات معرفية أخرى.

57 - النسيج الشبكاتي من إعداد الباحث.



الشكل (20): الشبكة المفاهيمية للمعرفة الصوفية وأفرادها ومواردها[58].

خاتمة

أثمرت جهودنا في عملية إعادة هيكلة النص المنتخب للفقرات الثلاث من كتاب كشاف اصطلاحات الفنون والعلوم في توفير فرصة مهمة لبيان الإمكانيات التي ستوفرها المعالجات المحوسبة ـ الذكية في التنقير عن المعاني وشبكة الدلالات المقيمة بين الاصطلاحات المنبثة في المعاجم.

وتبين أن عملية إنشاء نسيج النصوص تعد مثمرة لأنها توفر لنا فرصة للتنقيب في مادة المحتوى، وتحديد أنماط تداعيات المعاني وإنتاج المفاهيم والتي تدعم القارئ والباحث في تحديد الفجوات المعرفية، أو العلاقات الحميمية التي قد لا تتوفر أمامنا الفرصة للكشف عنها، وإيلاءها الاهتمام الكافي لسبر المفاهيم والدلالات.

ونعتقد أن مثل هذا النمط من المعالجات سيزيد من ثراء الحصيلة اللغوية الضخمة التي قد أودعت في معجم الدوحة التاريخي للّغة العربية حيث سيوفر لمستخدميه، وبمختلف فئاتهم تمثيلًا مرئيًا يكشف عن طبيعة الارتباطات القائمة بين مفردات المعجم، من جهة، والدلالات والمفاهيم المتولدة عن الاقترانات المفاهيمية. كما أن الصبغة التاريخية التي يتسم بها هذا المعجم يمكن أن تستثمر في توليد شبكات تاريخية تربط بين مختلف العصور بحيث يتسنى للمستخدم تتبع التطوّر التاريخي لولادة المصطلح، وتراكم الدلالات نتيجة لتوطنه في مدوّنة الخطاب العربي.

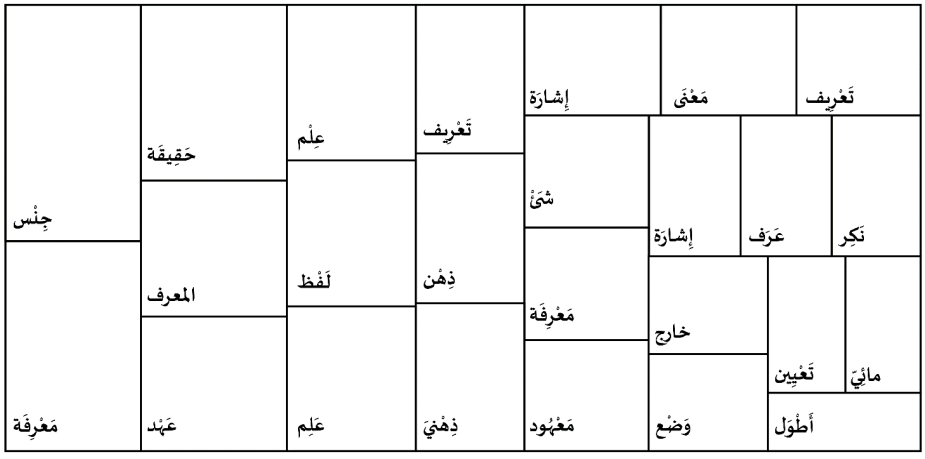
وعلى هذا الأساس فإن تبني هذا النهج في إعادة تشكيل عملية عرض المصطلحات في معجم الدوحة التاريخي سيمنح

58 - النسيج الشبكاتي من إعداد الباحث.

للعاملين عليه فرصة الانفتاح على شبكات المعاني والمفاهيم ووصف تطورها عبر مختلف العصور بدلًا من اعتماد النهج الاستاتيكي في إيراد المفردات مع تاريخ ظهورها، كما سيسهم في توليد المزيد من الشبكات الدلالية الرابطة بين مختلف المصطلحات التي تزدحم بها مادة المعجم الثرية بمفرداتها.

أما بالنسبة إلى المستخدم، فستتوفر له فرصة لتوليد مستويات متعددة من شبكات المفاهيم/ مع القدرة على تتبع مسارات ولادة المعاني والمفاهيم عبر العصور من خلال واجهة رسومية تفاعلية تجعله قادرًا على السفر بين مفردات المعجم، عبر العصور. انظر الشكلين (21 ، 22)، على سبيل المثال، اللذين يظهران أهم المفردات اللغوية والجذور اللغوية التي استخدمت في فقرة المعرفة، على التوالي؛ حيث يمكن من خلال هذين الشكلين تحديد مستوى اهتمام المؤلف بكل جذر من الجذور، والتي يمكن أن نستطيع من خلالها تحديد توجهات المؤلف في معالجة شبكة الاصطلاحات المقيمة في حقل المعرفة، دون أن نضطر لمطالعة الفقرة كاملة، فنوجّه دفّة اهتماماتنا باتجاه ما نروم نواله من هذه الفقرات، أو نُعرِض عنها؛ لأنها تلبي اهتماماتنا.

الشكل (21): مستوى حضور المفردات الموضوعية[59] في فقرة المعرفة[60].



الشكل (22): مستوى حضور جذور المفردات الموضوعية في فقرة المعرفة[61].

59 - الشكل الرسومي من إعداد الباحث.

60 - مساحة الشكل الذي تقيم فيه المفردة أو الجذر يعبر عن مستويات تكرارها وحضورها في فقرة النص.

61 - الشكل الرسومي من إعداد الباحث.

المراجع

أولًا: العربية

التهانوي، محمد بن علي القاضي. **كشاف اصطلاحات الفنون والعلوم**، تحقيق علي دحروج. بيروت: مكتبة لبنان ناشرون،1996. النسخة الرقمية.

خيرة، عبدو. **معمارية المعجم العربي الإلكتروني**، أطروحة ماجستير، قسم اللغة العربية وآدابه، الجزائر: كلية الآداب والفنون، جامعة عبد الحميد بن باديس،2015.

السامرائي، أحمد هاشم. "حوسبة المعجم العربي والقضايا المعاصرة". **مجلة سرّ من رأى**، مج9، ع34، السنة التاسعة (تموز 2013)

سيد، محمد نصر. "السوابق واللواحق في اللغة العربية". **حوليات آداب عين شمس**، عدد خاص (2017).

شاشه، إيمان. مناع، آمنة. تاوريريت، حسام الدين. "المعالجة الرقمية للمشاريع العربية المعجمية المحوسبة: عرض وقراءة لجهود عربية منجزة"، **مجلة البحوث التربوية والتعليمية**، مج11، عدد خاص (2022): 57- 76.

العزري، عيسى."دلالة اللواصق في اللغة العربية"، **مجلة اللغة العربية**، الجزائر، مج21، ع 43 (2019).

غياط، حنان. "صناعة المعجم التاريخي والتقنيات الحاسوبية: معجم الدوحة التاريخي أنموذجًا"، **مجلة جسور المعرفة**. مج8، ع3 (سبتمبر 2022): 803-815.

ثانيًا: الأجنبية

References:

Al-ʻAzrī, ʻĪsá. "The characteristic of derivation in Arab language" (in Arabic). *Majallat al-lughah al-ʻArabīyah*, al-Jazā'ir, Vol. 21, No. 43 (2019).

Al-Sāmarrā'ī, Aḥmad Hāshim. "Ḥawsabat al-Muʻjam al-ʻArabī wa-al-qaḍāyā al-muʻāṣirah" (in Arabic). *Majallat Sirra man ra'á*, Vol. 9, No. 34, al-Sanah al-tāsiʻah (July 2013).

Al-Tahānawī, Muḥammad ibn ʻAlī al-Qāḍī. *Kashshāf iṣṭilāḥāt al-Funūn wa-al-ʻUlūm* (in Arabic), taḥqīq ʻAlī Daḥrūj. Bayrūt: Maktabat Lubnān Nāshirūn, 1996. Al-nuskhah al-raqmīyah.

Ciotti, Fabio. "Distant Reading In Literary Studies: A Methodology In Quest Of Theory", *Testo e Senso*, No. 23 (2021): 195-213.

Feldman, Ronen & Sanger, James. *The Text Mining Handbook: Advanced Approaches in AnalyzingUnstructured Data*. UK, Cambridge: Cambridge University Press, 2006.

Ghayyāṭ, Ḥanān." Historical Dictionary industry and computational techniques Doha Historical Dictionary as a model" (in Arabic). *Djoussour El-maaréfa*. Vol. 8, No. 3 (September 2022): 803-815.

GitHub, "Python Package For Arabic Natural Language Processing", Accessed on 7/5/2023. Available At: https://github.com/adhaamehab/arabicnlp.

Guneri, Fatma, Hayırcıl, Gulrenk & Selin, Iman. "A Word Visualization Observation: The Hidden Meaning of the Words in Work Environments", *European Scientific Journal*, ESJ, Vol. 17, No. 36 (2021).

Hearst, Marita & Rosner, Daniela. *Tag Clouds: Data Analysis Tool or Social Signaller?*, School of Information. Berkeley: University of California, Berkeley, 2008.

Jänicke S., Franzini G., & Scheuermann G. "On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges", *The Eurographics Association* (2015).

Jänicke, S. et al. "Visual Text Analysis in Digital Humanities", *Computer Graphics Forum*, Vol. 36, No. 6 (2016).

Khayrah, ʻAbdū. *miʻmārīyah al-Muʻjam al-ʻArabī al-iliktrūnī* (in Arabic), uṭrūḥat mājistīr, Qism al-lughah al-ʻArabīyah wa-ādābihā. al-Jazā'ir: Kullīyat al-Ādāb wa-al-Funūn, Jāmiʻat ʻAbd al-Ḥamīd ibn Bādīs, 2015.

Mozuraite, Vita. "Change of The Reading Paradigm In The Age Of E-Book", *Libellarium*, VII, 1 (2014).

Obeid, Osama et al. "CAMeL Tools: An Open Source Python Toolkit for Arabic Natural Language Processing", *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, Marseille, 11-16 May 2020.

Sayyid, Muḥammad Naṣr. "Al-Sawābiq wa-al-lawāḥiq fī al-lughah al-ʻArabīyah" (in Arabic) Ḥawlīyāt ādāb ʻAyn Shams, special issue (2017).

Shāshh, Īmān, wa Mannāʻ, Āminah, wa Tāwryryt, Ḥusām al-Dīn. "Digital processing of computerized Arabic lexical projects Presentation and reading of accomplished Arab efforts" (in Arabic). *Majallat al-Buḥūth al-Tarbawīyah wa-al-taʻlīmīyah.* Vol. 11, No. 1 (special issue) (2022): 57-76.

Taeho, JO. *Text Mining: Concepts, Implementation, and Big Data Challenge*. Gewerbestrasse: Springer, 2019.

Walden University, "Grammar: Sentence Structure and Types of Sentences", Accessed on 7/10/2022. Available at: https://academicguides.waldenu.edu/writingcenter/grammar/sentencestructure

Wang, Ji. *Clustered Layout Word Cloud for User Generated Online Reviews*. Thesis. Virginia: Virginia Polytechnic Institute and State University, 2012.