osteology

MDPI

Review

A Macroscopic Exploration of the Ideoscape on Exosomes for Bone Regeneration

Maria Teresa Colangelo [1], Marco Meleti [2], Stefano Guizzardi [1] and Carlo Galli [1,*]

[1] Histology and Embryology Laboratory, Department of Medicine and Surgery, University of Parma, Via Volturno 39, 43126 Parma, Italy; mariateresa.colangelo@unipr.it (M.T.C.); stefano.guizzardi@unipr.it (S.G.)
[2] Department of Medicine and Dentistry, Dental School, University of Parma, 43126 Parma, Italy; marco.meleti@unipr.it
* Correspondence: carlo.galli@unipr.it

Abstract: Background: Exosomes, nanoscale extracellular vesicles, play a crucial role in tissue physiology and regeneration. This study uses infometric techniques to explore the structure of exosome-based tissue and bone regeneration research. **Methods:** We applied BERTopic, an advanced topic modeling algorithm, to a comprehensive corpus of the scientific literature on exosomes and tissue regeneration, identifying key themes such as stem cell studies, tissue healing, and regenerative applications, with orthopedics and dentistry emerging as dominant subfields. To further investigate the 'ideoscape', i.e., the conceptual landscape that maps how ideas, methods, and themes are interconnected across the field, we extracted significant concepts from abstracts using GPT 3.5 turbo and created knowledge graphs. **Results:** Our analysis revealed rapid growth in the field of dental stem cell regeneration, which has outpaced other bone regeneration topics by twofold. This analysis highlighted central themes such as periodontal stem cells and their cellular processes—proliferation, migration, and differentiation—along with their clinical applications. Our approach provided a clear visualization of the field's intellectual structure, showing how emerging topics are interconnected. Our findings offer a comprehensive view of the evolving trends in exosome-based bone regeneration, revealing not only the most active research areas but also gaps and opportunities for further investigation. **Conclusions:** This study exemplifies the utility of combining topic modeling with knowledge graph creation to map research trends, offering a flexible and largely automated tool for researchers to explore the vast bodies of literature and guide future research directions.

Keywords: exosomes; bone regeneration; BERTopic; knowledge graph; osteogenesis; tissue engineering; regenerative medicine

Citation: Colangelo, M.T.; Meleti, M.; Guizzardi, S.; Galli, C. A Macroscopic Exploration of the Ideoscape on Exosomes for Bone Regeneration. *Osteology* **2024**, *4*, 159–178. https://doi.org/10.3390/osteology4040013

Academic Editor: Rossella Bedini

Received: 31 July 2024
Revised: 18 September 2024
Accepted: 1 October 2024
Published: 8 October 2024

1. Introduction

Exosomes are extracellular vesicles (EVs) that appear to be released by all cell types that are unable to replicate on their own and are composed of a lipid bilayer containing a wide variety of cargo, depending on the donor cell type and their physiological conditions [1]. Along with ectosomes, they represent a subtype of EVs defined as small extracellular vesicles (sEVs), and their diameter is generally smaller than 200 nm. The term "exosome" is linked to biogenesis and refers to its endosomal origin [2]. It denotes EVs that are specifically distinguished by their formation within the internal compartments of the cell and are released through the multivesicular body (MVB) containing intraluminal vesicles (ILVs), ultimately secreted via exocytosis as exosomes [3].

The interaction of MVBs with other vesicles and intracellular organelles contributes to the diversity of exosome constituents, such as DNA, RNA, lipids, metabolites, and cytosolic and cell-surface proteins [4]. By transferring these bioactive molecules between cells, exosomes are natural endogenous nano-carriers that play a crucial role in near- and long-distance intercellular communication in health and disease [5]. Their intrinsic properties in

regulating complex intracellular pathways have drawn significant interest and enhanced their potential utility in the field of early diagnosis and therapy of tissue diseases [6–9].

Bone tissue engineering (BTE) that uses stem cells has become a potential strategy for tissue defect regeneration, and in this regard, mesenchymal stromal/stem cells (MSCs) have been shown to be efficacious for the treatment of bone defects and diseases [10,11]. Interestingly, accumulating evidence indicates that the regenerative effect of MSCs is primarily to promote the activity of tissue-resident recipient cells through paracrine, such as exosomes [12]. As a key element of cell-free therapy, exosomes have become a promising tool for bone regeneration in recent decades [13,14] because of their promoting osteogenesis and osteogenic differentiation function in vivo and in vitro by overcoming the disadvantages of their cellular counterparts due to their low immunogenicity, small size, and wide range of sources [15]. Furthermore, in order to enrich the exosomal cargo and increase exosome targeting efficiency, engineering exosomes and exosome-integrated biomaterials have been developed and shown to achieve ideal capabilities, such as specific bone targeting, better osteogenic and angiogenic properties and bone healing promotion [16]. In light of these findings, exosomes have been shown to be a valid alternative to traditional cell-based therapies and an attractive strategy for therapeutic purposes in bone repair, regeneration, and age-related defects.

To better understand the potential of EVs, and exosomes in particular, for bone regeneration, it is essential to map the current landscape of research in this rapidly evolving field. In this study, we adopt a computational approach to explore the knowledge structure surrounding exosome research focused on bone regeneration. A useful heuristic concept we use is the "macroscope", a metaphorical tool introduced by de Rosnay in the 1970s [17]. Unlike a microscope, which zooms in on tiny details, or a telescope, which looks at distant objects, a macroscope provides a broad, integrative view of large-scale patterns and structures. Instead of focusing on the results individual reports have contributed to the field, we focus our attention on the themes that compose this research mosaic and how they intersect.

The macroscope is advantageous for exploring the ideoscape of this research field. The concept of "ideoscape" helps us visualize and analyze the flow and distribution of information within a scientific domain. Originally used in anthropology, "ideoscape" combines "idea" with the suffix "-scape", suggesting a landscape of information that can be explored [18]. In anthropology, this concept helps us understand how information moves within and between cultures, shaped by social practices, media, and technology. When applied to scientific research, we use this ideoscape concept to map out how knowledge is produced and disseminated, showing how certain studies or topics become influential and how new ideas spread.

The primary objective of the present report is, therefore, a broad, "low magnification" analysis of the exosome research field of its ideoscape. To this purpose, we employed topic modeling. Topic modeling is an area of computational research that aims at creating algorithms capable of automatically segmenting corpora of documents based on their theme and their topic, using various approaches [19]. The recent introduction of transformer-based embeddings and the BERTopic algorithm [20] have dramatically improved the performance of existing topic-modeling algorithms [21]. Transformer-based embeddings, like those from BERT (Bidirectional Encoder Representations from Transformers), are robust numerical representations of words, phrases, or even whole sentences [22]. BERTopic uses these embeddings to create dynamic topic representations, helping us identify evolving themes and trends in exosome research [23]. By integrating these methods, we systematically analyzed a large collection of scientific articles on exosome research, giving us a comprehensive view of the field.

This macroscopic approach is particularly useful in today's globalized world, where scientific research is shaped by various factors, including funding, collaboration networks, technological advancements, and socio-political contexts [24]. Unlike a traditional narrative or systematic review, which are invaluable for understanding what research has elucidated on a given topic, mapping out knowledge production in a research field helps us understand its development and the forces at play and identify potential research gaps. This report, therefore, aims to gain an understanding of the current research trends and to identify the topics that are being most actively investigated. We also believe that this investigation can

offer valuable insights for readers, particularly those new to the field, by highlighting areas that may be underexplored, thus presenting opportunities for further research.

2. Materials and Methods

Data were analyzed on Google Colab Pro notebooks powered by Python 3.10.12 [25] with computations powered by T4 GPUs [26]. The literature dataset was created with the Biopython library [27] through a query-driven exploration of MEDLINE using the Entrez.esearch function. The query utilized for this exploration was:

> "Exosome*"[MeSH] OR "Exosome*"[Title/Abstract] OR "Extracellular Vesicle*"[Title/Abstract] OR "Microvesicle*"[Title/Abstract]) AND ("Tissue Regeneration"[Title/Abstract] OR "Regenerative Medicine"[MeSH] OR "Regenerative Medicine"[Title/Abstract] OR "Tissue Repair"[Title/Abstract] OR "Regenerative Therapy"[Title/Abstract] OR "Wound Healing"[MeSH]

We iterated the web scraping process through Entrez.efetch covering publication years from 1952 to 2025. This allowed us to retrieve PubMed IDs (PMID), titles, publication years, authors, abstracts, and MeSH keywords for all relevant publications, which were subsequently formatted into a pandas dataframe for analysis [28].

BERTopic was chosen because of its use of advanced transformer-based embeddings that can generate more accurate topics, and its superior performance in generating interpretable topics compared to traditional algorithms like LDA or NMF [29].

We decided to classify articles based on their titles to maximize speed and efficiency, as these are a summary of the article topic [30]: a title is usually designed to communicate the gist of a report in few words, and we thus deemed it useful to characterize the article's topic [31], as supported by our previous experience [32] The dataset was not preprocessed, unlike previous publications [33], besides removing entries without titles. We did not remove stopwords—common grammatical words with no semantic meaning—to let BERTopic use them to create contextual embeddings [34].

Embeddings are dense vectors, i.e., non-zero numerical sequences that semantically encode word meanings into high-dimensional space [35]. An embedding can, therefore, represent the meaning of a word or even of a sentence through a series of numbers. Numerous algorithms have been proposed to create embeddings that better capture the meaning of natural language [36]. Unlike previous algorithms [37], BERT understands context and produces different embeddings for the same word based on its context [22,38], which is crucial for nuanced semantics in a text.

We used BERTopic to extract topics from publication titles in our dataset. BERTopic involves using embedding models, reducing their dimensionality, clustering them, and tagging the clusters with an algorithm known as class-based Term Frequency-Inverse Document Frequency, or cTf-Idf [39,40]. For this, we selected Huggingface's 'all-mpnet-base-v2' model, pre-trained on over one billion token pairs. These 768-dimensional embeddings are very complex to manipulate and, therefore, require dimensionality reduction to make them shorter and easier to handle while maintaining the data's topological structure. We achieved this with UMAP, which constructs a high-dimensional graph and optimizes it in lower-dimensional space [41]. Once embeddings have been reduced to a smaller size, they can be clustered into groups based on their similarity. For clustering, we employed HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise), which is an algorithm designed to find clusters in data by identifying areas of high density and classifying points of lower density as noise [42]. Unlike previous clustering tools, HDBSCAN does not require the number of clusters to be specified beforehand [43], making it highly effective for detecting semantic areas with varying densities. cTf-Idf was then applied to highlight key topic words within clusters. Unlike traditional Tf-Idf, which weighs terms based on their frequency within individual documents, cTf-Idf operates at the cluster level. It identifies terms that are most representative of each cluster by comparing their frequency in the cluster to their frequency across the entire dataset, making it ideal for topic modeling [20].

BERTopic supports topic fine-tuning with additional representation models. We used KeyBERT and Maximal Marginal Relevance (MMR) models. KeyBERT uses the transformers library to extract keywords [44], using a Bag of Words approach [45]. The similarity between document embeddings and keyword embeddings is then compared to select the most similar keywords. MMR selects keywords similarly to KeyBERT, but it is designed to boost their diversity to capture a broader spectrum of meaning [46].

Default topic labels in BERTopic are just a string of its four main keywords, which are often insufficiently informative or intuitive for deeper interpretation. We then used OpenAI's GPT-3.5 Turbo to generate more readable labels for the topics. A Large Language Model (LLM) like GPT-3.5 Turbo is capable of understanding, interpreting, and producing human language in a coherent and contextually appropriate manner [47]. LLMs need a prompt from users, which serves as the starting point for the model to generate relevant text based on the information provided. We set the following prompt:

> *I have a topic that contains the following documents:*
>
> *[DOCUMENTS]*
>
> *The topic is described by the following keywords: [KEYWORDS]*
>
> *Based on the information above, extract a short but highly descriptive topic label of at most five words. Make sure it is in the following format:*
>
> *topic: <topic label>*

Stopwords were removed after topic creation using the sklearn Countvectorizer function [48].

The evaluation of BERTopic's performance and the quality of the GPT-generated labels were conducted using a combination of qualitative and manual review methods. We assessed the coherence of the keywords within each cluster to ensure that they reflected a meaningful and cohesive topic. A sample of articles from each topic group (total n = 200) was manually reviewed to evaluate their relevance to the identified topic. For the GPT-generated labels, we compared the labels to the underlying keywords to verify that they accurately captured the essence of the respective topics. Additionally, we ran multiple iterations of the analysis to ensure the stability and consistency of the topic clusters across different runs.

Data were then visualized using BERTopic's inbuilt functions, and the matplotlib [49] and seaborn libraries [50]. In addition, the Datamapplot library was utilized for effective cluster visualization [51].

The trend in research topics was analyzed by a linear regression model using the scipy Linregress library [52]. The independent variable was the publication year in different time intervals, and the dependent variable was the number of papers within a given topic. The slope indicated how quickly the number of papers on a given topic rose over the time interval.

To investigate the structure (or ideoscape) of individual topics, we resorted to knowledge graphs [53]. A knowledge graph is a structured representation of information, where entities ("things", such as cells, diseases, or concepts) are connected by relationships, enabling data to be organized and linked in a way that captures their meaning. To extract entities that were representative of the scientific reports in the dataset, we decided to use abstracts because they supposedly explain the topic, methods, and results of a manuscript in greater detail. To capture the semantic meaning of abstracts, we employed the GPT-3.5-turbo model using the following prompt:

> *You are an expert in extracting key concepts, keywords, and relationships from scientific text to create a knowledge graph.*
>
> *Please extract the key concepts and their relationships from the following text. Make sure to provide meaningful labels and categories for each concept and normalize similar keywords to the same form. Use a concise and clear format.*

Here's the text:

\n\n{text}\n\n

Return the result in JSON format with 'nodes' and 'relationships' keys.

The results were returned in JSON format containing 'nodes' and 'relationships' keys. Each concept was also categorized according to a predefined taxonomy, including categories such as disease, symptom, treatment, medicine, process, component, and others. We then utilized the SpaCy library for lemmatization and tokenization to normalize keywords and relationship labels extracted from the abstracts [54].

A powerful aspect of knowledge graphs is the possibility to represent them graphically. To do that, we relied on Neo4j, an open-source graph database management system [55]. We established a connection to a Neo4j graph database [55] to store and visualize the relationships between the extracted concepts, with each article title represented as a 'Title' node and relationships created between these nodes and their corresponding concept nodes. 'Year' nodes were also created to indicate the publication year of each article and connected to the 'Title' nodes. The free Neo4j aura online browser was used to execute Cypher queries and to create and manage nodes and relationships in the Neo4j database.

3. Results and Discussion

3.1. Overview of the Dataset

As the purpose of the current investigation was to focus on research on the role of exosomes in tissue regeneration, and more precisely, on bone regeneration, our resulting dataset was quite small (n = 2868); the earliest articles did not date to before 1995, and the dataset mostly comprised recent or very recent research products, as shown in Figure 1.

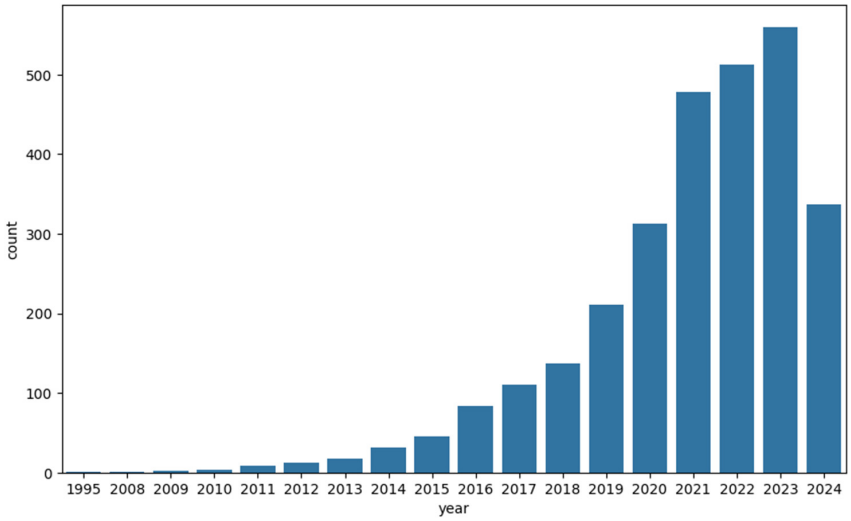


Figure 1. This bar chart represents the distribution of the number of publications in our dataset over the years. The number of publications has been increasing exponentially. Fewer publications appeared in 2024 because the year is still ongoing.

As with most of the biomedical literature, the number of articles has been exponentially increasing in the last few years and, in the case of exosomes in tissue regeneration, within the last decade. The number of published articles in 2024 is understandably lower than in 2023 at the time this manuscript is written.

A recent paper conducted a thorough bibliometric exploration of the literature production about exosomes and bone [56]. This investigation, however, focused exclusively on bone diseases, so we decided to broaden the scope of our investigation and comprehend tissue regeneration to better capture the diversity in the research themes across various areas of application of regenerative medicine.

To get an overview of the knowledge structure of this science field, we focused our subsequent analysis on the titles of the articles in the dataset, and we ran the BERTopic algorithm on them. BERTopic segments a dataset of unstructured documents by clustering them based on their semantics. This is made possible by encoding every document, in this case, every title, using transformer-based models, which generate sentence embeddings, i.e., a numerical representation of the text meaning. BERTopic can thus cluster these embeddings based on their similarity and, extracting representative keywords for each title cluster, can produce a description of its content. To obtain more readable topic labels and get a more 'human' feeling, we decided to use OpenAI's GPT to label them, in addition to BERTopic default meaning representations. GPT is a large language model (LLM), i.e., an artificial intelligence model capable of understanding and generating language with a high degree of similarity to human language [57]. We selected GPT-3.5 Turbo over GPT-4 for the task of labeling topics in BERTopic. While GPT-4 generally outperforms GPT-3.5 Turbo in complex tasks [58], the labeling task in our study was relatively straightforward. GPT's role was primarily to generate a representative phrase from a list of keywords, as it is more intuitive for humans to interpret topics as phrases rather than isolated words. In multiple test runs with both models, we observed that they produced comparable labels, and blinded investigators could not distinguish between labels generated by GPT-3.5 Turbo and GPT-4. Given the similarity in output and the significantly lower cost of using GPT-3.5 Turbo, we opted for this model to optimize resource use without compromising quality.

Though we used titles for the present work, it could be reasonably argued that abstracts are preferable because they provide more detailed information and context on a manuscript's topic. While abstracts undeniably offer greater depth, our decision to focus on titles was based on our previous experience with transformer models [32,59] and on the performance efficiency of the selected algorithm. Abstracts are usually much longer than titles and, therefore, require a longer processing time, even with hardware acceleration.

This level of granularity in topic analysis can be fine-tuned arbitrarily. For our purposes, we set the minimum cluster size to 30, i.e., we excluded clusters with fewer than 30 articles. Articles that could not be classified into any other groups were allocated to a '−1' group.

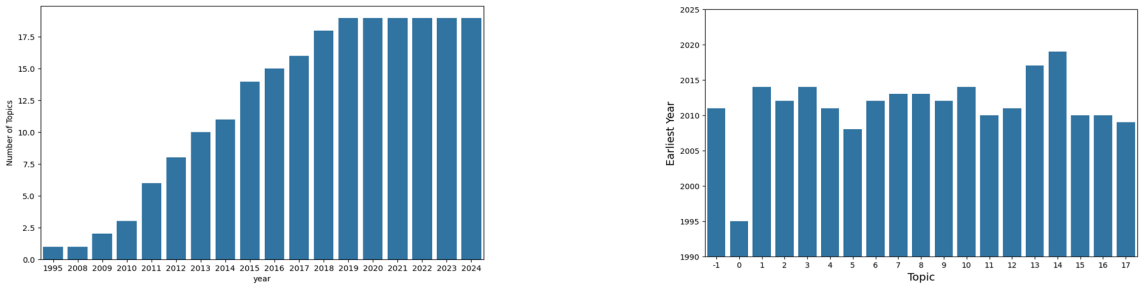
BERTopic analysis, when executed with these parameters, revealed several key research areas within the academic literature on exosomes, and a list of the identified topics can be found below (Table S1). For every topic group, we indicated the number of articles in it, its default name, which BERTopic creates by chaining together the four most relevant keywords that characterize it, the keywords extracted with the KeyBERT and MMR algorithms, and the label obtained from the LLM. Readers are encouraged to examine the table to assess the range of topics identified by BERTopic and observe how these topics can be described with varying nuances depending on the algorithm used to extract keyword descriptors.

The largest topic group, interestingly, is topic #0, Stem Cells for Bone Regeneration (n = 443), followed by topic #1, Exosomes for Diabetic Wound Healing (n = 354). Some topics are more general and reflect lexical differences rather than a specific focus, such as topic #2, Extracellular Vesicles in Regenerative Medicine. Other topics are more focused on the application of exosome research in specific fields, such as #5 Extracellular Vesicles in Skin Healing (n = 146), #8 Stem Cell Therapy for Spinal Injury (n = 119), and topic #10 Cardiac Repair with Extracellular Vesicles (n = 91). Unsurprisingly, these topics have not been consistently represented over the last 30 years. Figure 2A shows the number of topics over time in our dataset. The number of topics increased progressively since 1995 and reached the final level of 19 topics only in 2019. This suggests that research in this area has been expanding not only in volume but also in diversity, with new areas of investigation emerging over time.

Figure 2B indicates that the oldest—and largest—topic is #0 Stem Cells for Bone Regeneration. Most other topics emerged during the 2010–2015 period, with the most recent being topic #14, Exosome Therapy for Rotator Cuff (n = 47).

To visualize the structure and knowledge architecture of this research field, we reduced the embeddings to two dimensions using the UMAP dimensionality reduction algorithm, allowing each title to be represented as a dot in a Cartesian semantic space. Dots

positioned close together indicate article titles with similar meanings, while those farther apart represent more distant topics. However, the exact placement of the dots is random. Figure 3 presents a scatter plot of the 2868 titles in the exosome dataset. Grey dots represent unclassified titles scattered across the semantic space, categorized by HDBSCAN as cluster −1. Notably, the left side of the semantic space is populated by articles on wound healing, such as skin and diabetic wound healing, which are naturally grouped together. Bone regeneration research clusters in a different area, located near studies on mesenchymal stem cells, rotator cuff injuries, and, somewhat unexpectedly, corneal wound healing.



(**A**) (**B**)

Figure 2. (**A**) Bar chart representing the number of topics in the dataset per year. Most topics appeared progressively in the 2009–2019 decade. By 2019, all the topics identified by BERTopic in our dataset were present. (**B**) Bar chart representing the earliest publication per topic. Topic #0, which is also the largest topic, contains the oldest paper in the dataset.

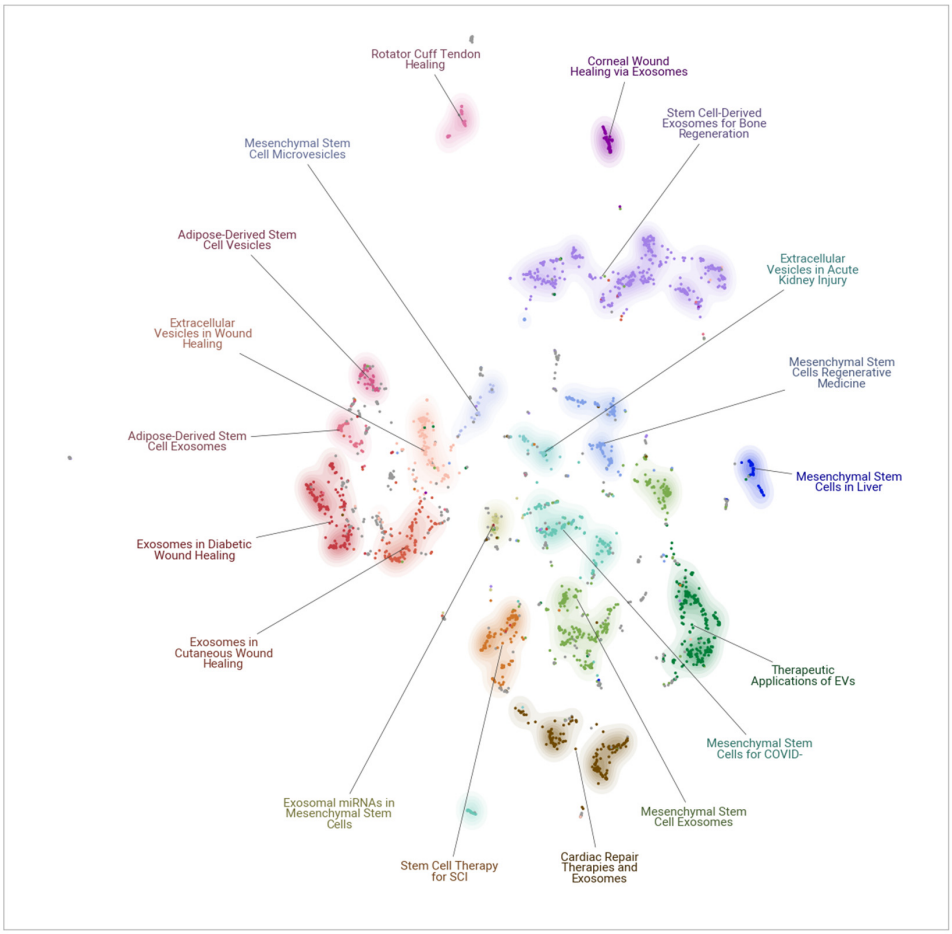


Figure 3. Scatterplot of the semantic distribution of the dataset of titles of scientific articles on exosomes and tissue regeneration. Each article is represented by a dot within the semantic space. The coordinates

of the dots were obtained by reducing the embeddings to 2-dimensional tuples, which were then used as cartesian coordinates. Grey dots belong to group −1, unclassified papers. Articles are not homogeneously distributed, but rather, form clusters that correspond to topics. Clusters of articles are highlighted by a colored halo. Every topic is also marked by a label generated by the LLM.

3.2. Exosomes for Bone Regeneration

We then decided to focus our attention on bone regeneration alone and isolated the papers belonging to topic #0, re-running BERTopic on this subset alone. BERTopic thus identified seven topics (Table S2) within the bone subset. Noticeably, only 14 articles remained unclassified in the −1 cluster, while most of the articles of the dataset were allocated to a specific group.

The largest topic, #0 (n = 126), was labeled quite generically "Mesenchymal Stem Cell Therapy" by GPT. Insights into its content can be provided by its keyword descriptors, which situate it quite firmly in the orthopedic area, with special focus on osteoarthritis:

['cells osteoarthritis', 'cartilage regeneration', 'exosomes derived', 'stem cells', 'stem cell', 'treatment osteoarthritis', 'extracellular vesicles', 'derived exosomes', 'osteoarthritis treatment', 'mesenchymal stem']

Visual inspection of the topic group revealed several papers devoted to this theme, e.g.,

> "*Mesenchymal stem cell-derived exosomes as a promising cell-free therapy for knee osteoarthritis.*" [60]

> "*Exosomes Reshape the Osteoarthritic Defect: Emerging Potential in Regenerative Medicine-A Review.*" [61]

> "*Exosomes Derived from Bone Marrow Mesenchymal Stem Cells Alleviate Rheumatoid Arthritis Symptoms via Shuttling Proteins.*" [62]

> "*Exosomes derived from MSC as drug system in osteoarthritis therapy.*" [63]

Topic #1, a similarly large cluster (n = 103), appeared to mostly contain research on bone fractures and bone healing, and LLM thus labeled it "Bone fracture healing exosomes". A closer inspection of the cluster revealed that it contained at least two main research lines, i.e., both articles investigating the physiological role of exosomes in bone repair, e.g.,

> "*Proteomic analysis of exosomal proteins associated with bone healing speed in a rat tibial fracture model.*" [64]

> "*Cellular and Molecular Connections Between Bone Fracture Healing and Exosomes.*" [65]

and research articles devoted to the biotechnological use of exosomes to improve bone healing, e.g.,

> "*Role of nano-hydrogels coated exosomes in bone tissue repair.*" [66]

> "*Enhancing bone regeneration and immunomodulation via gelatin methacryloyl hydrogel-encapsulated exosomes from osteogenic pre-differentiated mesenchymal stem cells.*" [67]

In contrast to Chen et al. [56], a sizeable number of articles were centered on dental topics. Topic #2 Dental Stem Cell Regeneration (n = 80) and topic #4 'Dental Pulp Stem Cell Exosomes (n = 48), were mostly centered on dental and periodontal applications of exosomes. As their LLM label might not fully account for the difference between these two topic clusters, we inspected the keyword descriptors for topic #2:

['periodontal regeneration', 'vesicles periodontal', 'vesicles dental', 'periodontal tissue', 'pulp regeneration', 'regeneration dental', 'pulp periodontal', 'extracellular vesicles', 'extracellular vesicle', 'dental pulp']

which suggests that this topic contains articles focusing on periodontics applications of exosomes. This was confirmed through a visual inspection of this group, which revealed academic articles in the periodontics area, e.g.,

> "*Clinical Efficacy of Extracellular Vesicle Therapy in Periodontitis: Reduced Inflammation and Enhanced Regeneration.*" [68]

"*3D bioprinted small extracellular vesicles from periodontal cells enhance mesenchymal stromal cell function.*" [69]

"*Effects of periodontal cells-derived extracellular vesicles on mesenchymal stromal cell function.*" [70]

On the contrary, the KeyBERT-derived descriptors for topic #4 were:

['exosomes dental', 'exosomes regenerative', 'exosomes oral', 'exosomes derived', 'cell exosomes', 'exosomes', 'cells exosomes', 'exosomes enhance', 'derived exosomes', 'regenerative endodontic']

The MMR algorithm confirmed that this latter topic focuses on dental pulp rather than periodontal tissues as topic #2 (as suggested by the LLM label):

['exosomes', 'derived', 'stem', 'pulp', 'derived exosomes', 'cell', 'dental', 'exosomes derived', 'cells', 'dental pulp']

which justifies the distinction between these two groups. Inspection of this cluster confirmed the descriptors, highlighting the focus on the dental pulp as the origin or therapeutic target of the exosomes, e.g.,

"*Exosomes as Promising Therapeutic Tools for Regenerative Endodontic Therapy.*" [71]

"*Human dental pulp stem cell-derived exosomes decorated titanium scaffolds for promoting bone regeneration.*" [72]

"*Exosomes derived from odontogenic stem cells: Its role in the dentin-pulp complex.*" [73]

Topic #3, Stem Cell-Derived Vesicles for Bone Regeneration, seems to substantially overlap topic #1, but its descriptors reveal an additional and possibly distinctive focus, i.e., special attention to biomaterial scaffolds, which was absent in topic #0:

['bone regeneration', 'vesicles bone', 'hydroxyapatite scaffold', 'osteoblast derived', 'extracellular vesicles', 'bone healing', 'extracellular vesicle', 'regeneration extracellular', '3d hydroxyapatite', 'tissue regeneration']

To better understand the composition of this group, we manually screened its titles and identified articles that focused on materials, scaffolds, and the use of exosomes for therapy rather than the study of their physiological role, e.g.,

"*Functionalized 3D Hydroxyapatite Scaffold by Fusion Peptides-Mediated Small Extracellular Vesicles of Stem Cells for Bone Tissue Regeneration.*" [74]

"*Transplanted MSCs promote alveolar bone repair via hypoxia-induced extracellular vesicle secretion.*" [75]

"*Bioengineering extracellular vesicles: smart nanomaterials for bone regeneration.*" [76]

"*Biomimetic synthesis and optimization of extracellular vesicles for bone regeneration.*" [77]

The last topic is small (n = 16) but very specific, as it is labeled "Exosome Therapy for Disc Degeneration" and contains very consistently themed articles on this specific niche area, e.g.,

"*Nanoscale Treatment of Intervertebral Disc Degeneration: Mesenchymal Stem Cell Exosome Transplantation.*" [78]

"*MSC-Derived Exosomes Ameliorate Intervertebral Disc Degeneration By Regulating the Keap1/Nrf2 Axis.*" [79]

"*Exosomes: A promising therapeutic strategy for intervertebral disc degeneration.*" [80]

Figure 4 represents the semantic distribution of these topics, which confirms that topic #5 is located at a certain distance from the other bone topics, possibly due to the use of a very specific language that positions it quite far apart from the rest of the orthopedic literature. On the contrary, topic #1, Bone fracture healing exosomes, is visibly juxtaposed with topic #3, Stem Cell-Derived Vesicles for Bone Regeneration, confirming that there is some continuity between the two topic groups.

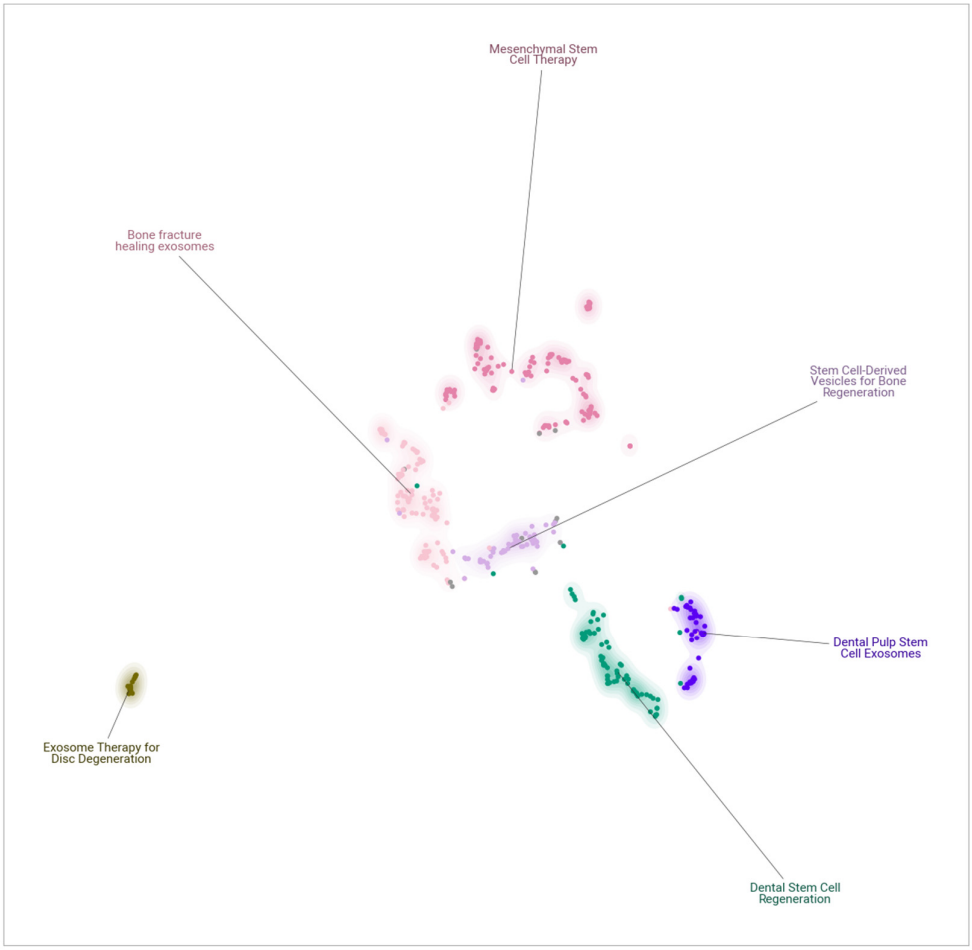


Figure 4. Scatterplot of the semantic distribution of the dataset of titles of scientific articles on exosomes and bone regeneration. Each article is represented by a dot within the semantic space. The coordinates of the dots were obtained by reducing the embeddings to 2-dimensional tuples, which were then used as cartesian coordinates. Grey dots belong to group −1, unclassified papers. Articles are not homogeneously distributed, but rather, form a few well-defined clusters, which correspond to topics and are highlighted by a colored halo for easier reading. Every topic is also marked by a label generated by the LLM.

3.3. Evolution of the Field

To better understand how the bone research field evolved over time, we plotted the growth of the topics as a number of articles in the topic group over time (Figure 5). The plot clearly shows that, just like the rest of the dataset, this particular subset of the literature went through a phase of expansion around the year 2017, with a rapid increase in the number of publications. While most of the topics flourished in the last five years, with a drop in the last part of the chart that is attributable only to the fact that the number of articles published in 2024 is only partial, it seems that Topic #0 of this subset peaked around the year 2020, while Topic #1 Bone fracture healing exosomes has taken over as the fastest growing topic.

To gain further insight into the publication trends in these topics, we performed a linear regression over the number of papers published in the different clusters. Linear regression finds the equation of a line that minimizes the distance between the points on the line and the points in the dataset. The slope of the fitted line indicates the publication trend (Table 1), and it can be considered a crude indicator of how fast the number of publications is changing over time. Though most topics are growing (Table 1), with the only possible exception of topic #3, Stem Cell-Derived Vesicles for Bone Regeneration, which has been maintaining similar levels of publications in the analyzed period, with a

slope of 1, two topics stand out as growth, i.e., topic #2 Dental Stem Cell Regeneration and topic #1 Bone fracture healing exosome, with slopes exceeding 4. As Figure 6 shows, while Topic #0, Mesenchymal Stem Cell Therapy, and Topic #1, Bone fracture healing exosomes, are fast-growing research fields, Topic #2, Dental Stem Cell Regeneration, is the fastest growing topic, with a slope of 4.5 over the 2019–2023 period.

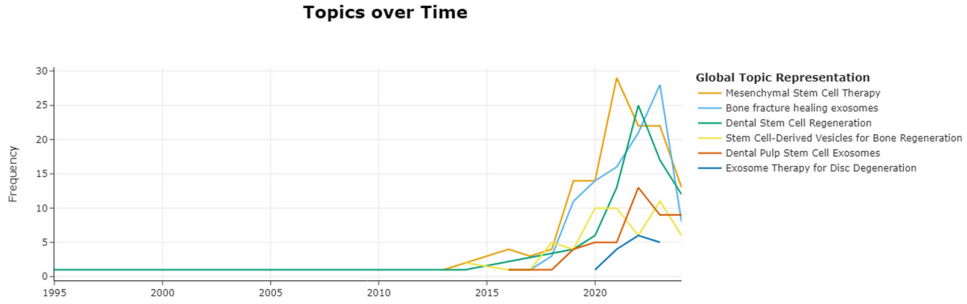


Figure 5. Line chart tracking the frequency of papers published by year by topic in the bone regeneration subset. Each topic is indicated by a solid line of different colors. Most topics start to gain traction only in the 2015–2020 quinquennium.

Table 1. List of bone regeneration topics in the 2019–2023 period sorted by the slope of the linear regression of the number of publications/year. Dental Stem Cell Regeneration and Bone fracture healing exosomes have a much higher slope than the rest of the topics.

Topic	Trend (Slope)
Dental Stem Cell Regeneration	4.5
Bone fracture healing exosome	4.1
Mesenchymal Stem Cell Therapy	2.4
Dental Pulp Stem Cell Exosomes	1.8
Exosome Therapy for Disc Degeneration	1.4
Stem Cell-Derived Vesicles for Bone Regeneration	1.0

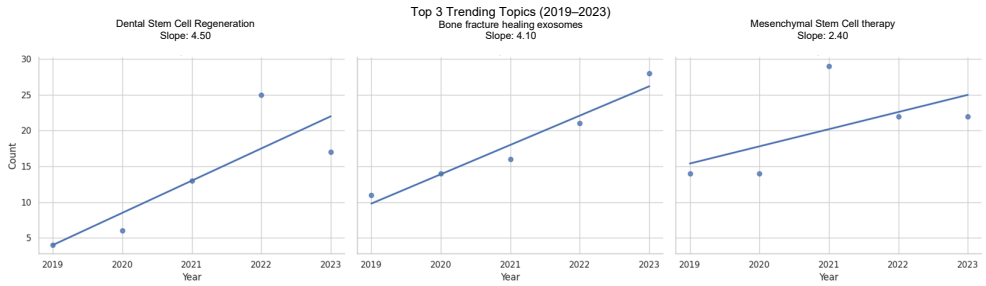


Figure 6. Trend analysis for the three highest-ranking topics in the bone regeneration subset. Linear regression (solid blue line) was used to fit the number of publications per year in the 2019–2023 period (indicated by a blue dot) and to infer the publication trend from its slope. The steeper the slope, the closer the line is to vertical, and the higher the increase in the number of publications.

3.4. Knowledge Structure of Exosome Research: The Ideoscape

To get a better understanding of how ideas and themes are shared in this rapidly growing field, we decided to limit our observation to topic #2, Dental Stem Cell Regeneration, which has been undergoing the steepest growth. We constructed a knowledge graph from the abstracts of articles in this cluster, using GPT 3.5 turbo to extract meaningful keywords and identify relationships between key concepts. Abstracts were selected over titles because they provide a more comprehensive summary of the methodologies, research approaches, and key findings of individual studies, whereas titles often only indicate the general topic or research area—making them more suitable for topic modeling. Unlike the

BERTopic approach, where we utilized keyword-extraction algorithms like KeyBERT, we relied entirely on the LLM (GPT 3.5 turbo) with a tailored prompt to extract core concepts and findings from the abstracts. This enabled us to build a network of ideas that represents a simplified ideoscape of the field, which we then explored and visualized using Neo4j, a platform specifically designed for querying and visualizing knowledge graphs. Neo4j is an interactive tool to visualize graphs, which allows users to search through the graph using a query language called Cypher, similar to what is performed with the SQL language for relational databases [81].

A knowledge graph is a structured representation of knowledge that captures entities, their attributes, and the relationships between them, facilitating both human readability and the adoption of computational procedures to extract insights, or knowledge, from the data through queries [82]. It comprises nodes, representing entities or concepts, and edges, depicting the relationships between these nodes. Both nodes and edges can have multi-layered attributes that provide additional context. Such a graph is, therefore, a complex representation of concepts and how these concepts are related to each other [83].

GPT was capable of not only extracting keywords and assign them a label but also cluster keywords by categories, such as anatomy, cell models, procedures, or therapies, providing a further layer of insight and structure into these key concepts. We only reported selected screenshots of Neo4j graphs or of meaningful subgraphs, although these are only a partial representation of the whole graph and lack their distinguishing interactivity.

Figure 7 shows a graph of the nodes categorized as anatomical sites (orange circles) that were investigated by at least two research articles (large green circles), i.e., that were connected to at least two article nodes by a "Contains" edge.

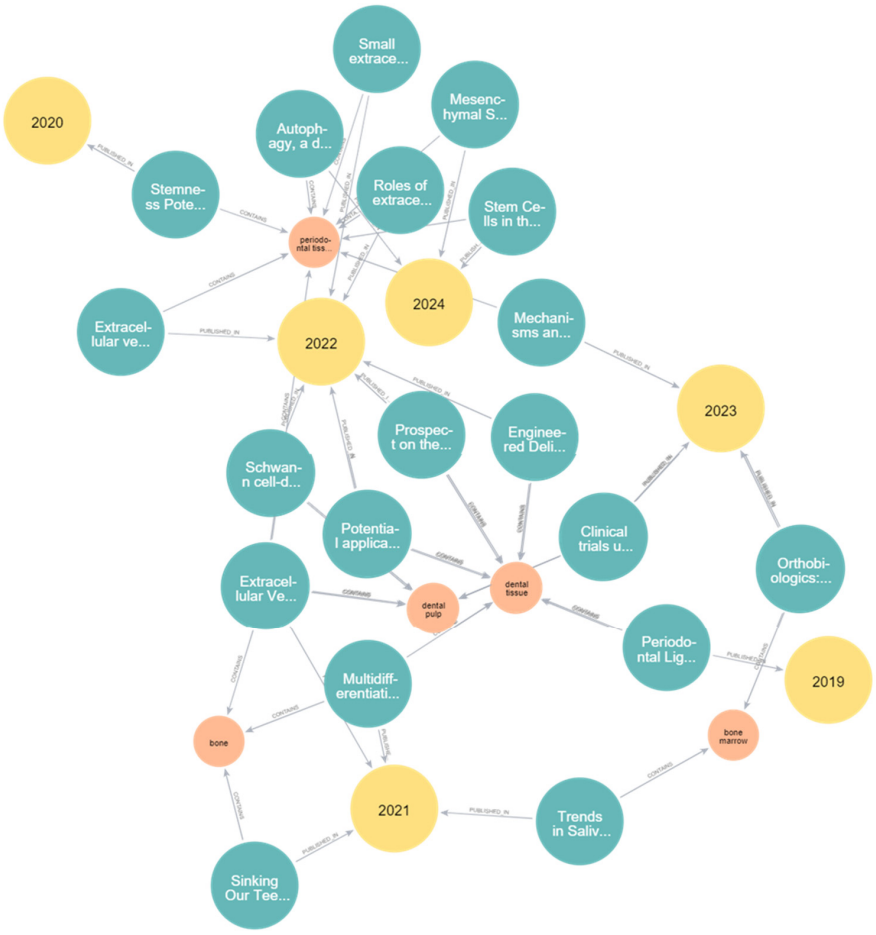


Figure 7. Neo4j knowledge graph of anatomy nodes (orange). Each node contains a concept pertaining to anatomy (e.g., a locus) and is connected to at least two article nodes (in green). Article

nodes represent the articles in the dataset, and a part of their title is written on the node. Yellow nodes indicate publication years. Nodes are connected by edges, which represent a relation linking two nodes. A label on the edge indicates the kind of existing relation.

Anatomy keywords that were connected through only one edge to an article node are not represented in this plot because our purpose was to see how ideas could be shared between papers. So, only a few Anatomy nodes are displayed out of a total of 84 concept nodes for this category. Large Yellow circles are Year nodes, i.e., nodes that represent the publication year of an article and are connected to articles published in that year through a "Published_in" edge.

The graph in Figure 7 shows how periodontal tissue has catalyzed most of the recent attention, with several articles gravitating around the "periodontal tissue" keyword (in agreement with the central interest for Periodontal disease as shown in Figure S1), especially in the 2022–2024 period. This trend aligns with a recent systematic review highlighting the diverse applications of exosomes in periodontal research [84]. Noticeably, in the case of the paper "Periodontal Ligament Stem Cells: Current Knowledge and Future Perspectives" [85], the LLM felt a closer connection to the Dental tissue descriptor node (Figure 7), likely due to the article's focus on the multi-lineage potential of stem cells from the periodontium. Interestingly, the paper was published in 2019, potentially before the periodontal tissue had imposed itself. Overall, this graph underscores the growing interest in dental and periodontal applications, suggesting to investigators that this is a thriving area where research efforts are being concentrated.

Given the popularity of stem cells as a keyword in this dataset, frequently appearing in the descriptors extracted by BERTopic, we chose to analyze the most common stem cell models and topics, investigating the "Cells" nodes. Again, we only focused on those nodes that were connected to at least two article nodes (Figure 8, brown circles) to highlight idea interactions or at least focal points of ideas. While the dataset included 105 different 'Cells' nodes, only a few cell models attracted the majority of publications in the subsets (Figure 8).

Mesenchymal stem cells were—possibly unsurprisingly so—the focus of a vast portion of the literature of this group. This prominence can be attributed to the flexibility of mesenchymal stem cells as a model in tissue regeneration, supported by a vast body of evidence [86]. Additionally, since deep periodontal tissues are mesenchymal in nature, most regenerative approaches naturally center around the use of this cell model [87].

Periodontal ligament stem cells were another common concept in this subset of the literature, and many research reports focused on them, albeit they have long represented quite an elusive cell population [88]. In contrast, dental stem cells and pulp stem cells garnered attention from a smaller niche of the literature. A handful of articles centered around the broader 'Stem cell' term, often simultaneously connected to more specific cell types (Figure 8).

Collectively, these findings serve as a valuable reminder of the most prominent models utilized in the literature (and thus the most robust, which new investigators might prefer as safer) while also highlighting potential cell models that remain relatively underexplored in the field and where room may be available to set up novel investigations. We then focused on the "Process" nodes that were investigated in this literature set, which is an interesting category per se, given its polysemous nature (Figure 9). A total of 129 different Process nodes were identified by GPT, although those that are shared by at least two article nodes are much fewer. Interestingly, we observed two distinct poles of keywords: one focused primarily on cellular processes, such as cell migration, proliferation, and differentiation, and another centered on clinical applications, biogenesis, and tissue repair. A strong 'Tissue regeneration' concept node served as a central theme, linking much of the dataset and providing a key descriptor for this group of articles (Figure 9). Overall, the LLM thus recognized two semantic areas for these processes: natural processes that are affected by exosomes or affect exosomes (i.e., the cell mechanisms) on one end and thematic areas of application of exosomes for therapy purposes.

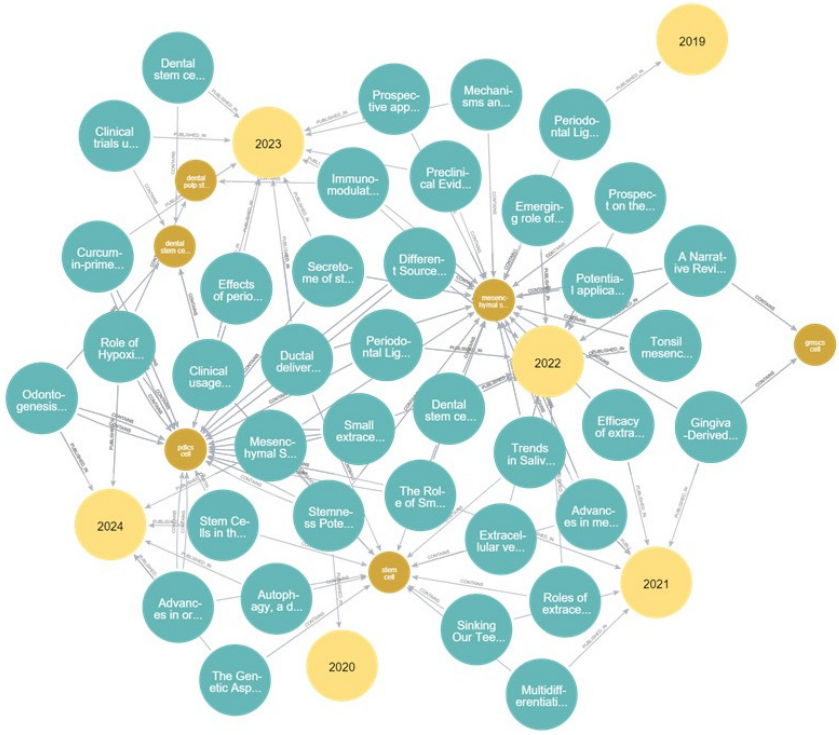


Figure 8. Neo4j knowledge graph of cells nodes (khaki). Each node contains a concept pertaining to a cell model and is connected to at least two article nodes (in green). Cell nodes connecting to only one article were not visualized. Green article nodes represent the articles in the dataset, and a part of their title is indicated on the node. Yellow nodes indicate publication years. Nodes are connected by edges, which represent a relation linking two nodes. A label on the edge indicates the kind of existing relation.

Figure 9. Neo4j knowledge graph of process nodes (light green). Each node contains a concept pertaining to a process (e.g., cell proliferation, migration, etc.) and is connected to at least two article nodes (in darker.

green). Article nodes represent the articles in the dataset, and a part of their title is written on the node. Yellow nodes indicate the publication year. Nodes are connected by edges, which represent a relation linking two nodes. A label on the edge indicates the kind of existing relation

GPT was unable to assign a specific category to a set of nodes, and these nodes were hence generically labeled "Unknown"; this large group ($n = 788$) encompasses many generic keywords from the literature dataset (Figure 10), e.g., 'Chemokine', or 'In vitro research'. However, most of these concept nodes are very specific, e.g., 'beagle modem organism', and are associated with only one article node. The concept nodes that are shared across a number of articles are much more limited in number (Figure 10, light blue). It is easy to observe how "Tissue Regeneration" and "Extracellular Vesicles" are the main attractors for the more recent papers, as highlighted by the high number of connections that they have with articles in the dataset, which is expected because of the nature of the dataset itself.

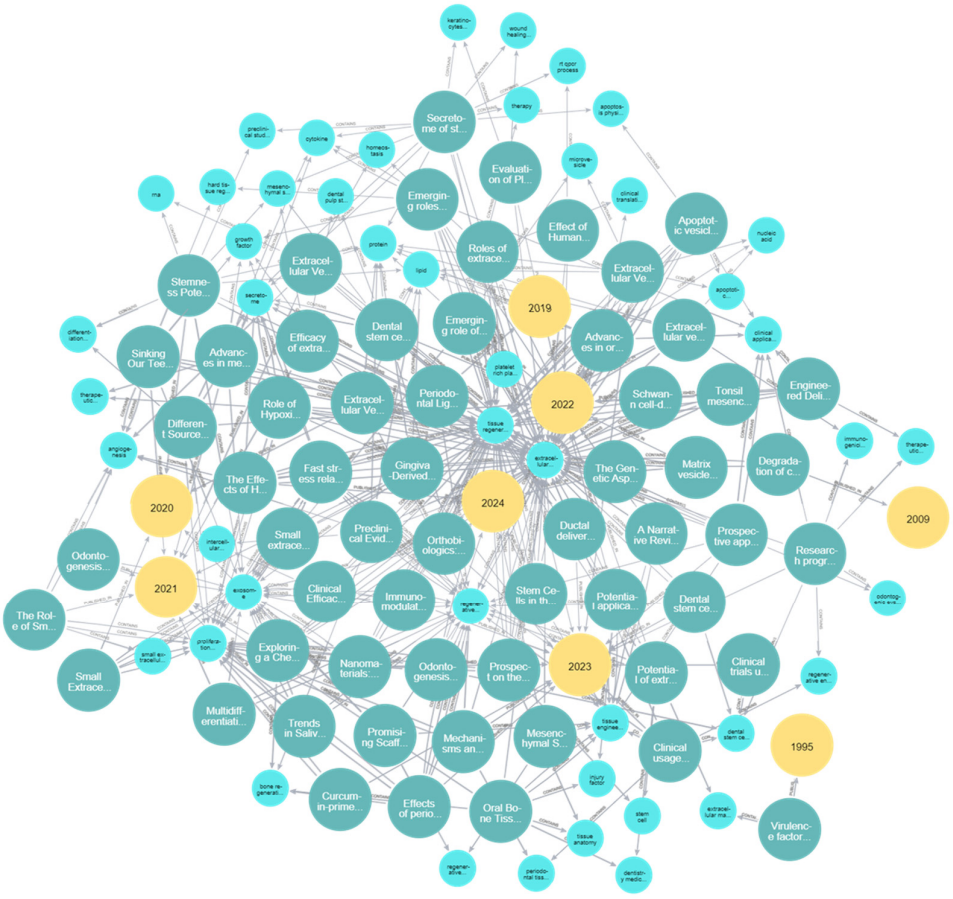


Figure 10. Neo4j knowledge graph of generic "Unknown" category nodes (turquoise). This category contains all the entities that GPT was unable to assign to other categories. Each node is connected to at least two article nodes (in green). Article nodes represent the articles in the dataset, and a part of their title is written on the node. Yellow nodes indicate the publication year. Nodes are connected by edges, which represent a relation linking two nodes. A label on the edge indicates the kind of existing relation.

As expected, "Exosomes" and "Secretome" (albeit to a lesser extent) are also popular buzzwords, although they tend to cluster together with other concepts that are more typical of pre-clinical settings, such as "Proliferation", "Differentiation", "Angiogenesis" or "Intercellular Communication" (Figure 10).

4. Conclusions

The field of exosomes and extracellular vesicles for tissue regeneration is a fast-expanding and multi-faceted area of research, which has grown exponentially in the last few years. Rather than focusing on the details of their biology and applications, our study employed a macroscope approach to try to understand the field in its complexity at an overview level of its topics and trends. This approach, while not exhaustive, offers a complementary perspective to traditional reviews and serves as a valuable tool for assessing how the field is developing over time. Our investigation has shown that EVs, their biology, and their potential applications in tissue regeneration have captivated researchers from several science fields, with a net prevalence of stem cell studies and tissue healing both in the skin, in the musculoskeletal system, and in various organs, such as kidneys or the eye (Figure 3). Specifically, within bone regeneration research, orthopedics, and dentistry emerged as key thematic areas with distinguishable subfields. Dentistry, particularly periodontics, and bone fractures exhibited the most rapid growth, outpacing other bone regeneration topics by at least twofold (Table 1). We then attempted to investigate the main ideas that associated subsets of papers in the field and conveniently focused on the topic that has been growing the most, i.e., topic #2, Dental Stem cell regeneration. To do that, we extracted the main concepts from the abstracts of the articles in this subset and the relations associating them to create a knowledge graph, which represents the 'ideoscape' of the field, i.e., the shared ideas around which most papers in the dataset gravitate. Understanding the knowledge structure of the field is critical, as it allows researchers to identify key trends, emerging areas of focus, and gaps in the literature. For instance, by zooming in on the rapidly growing topic of dental stem cell regeneration, we extracted and analyzed the main concepts and associations from the abstracts in this subset.

Our findings indicate that, over the past three years, the literature has increasingly centered around periodontal stem cells and their associated cellular processes (e.g., proliferation, migration, and differentiation) as well as their clinical applications (Figures 8 and 9). Although our results are necessarily partial and based on specific queries, the insights we provide highlight how the knowledge structure of exosome research is shaping, especially within the context of bone regeneration. The concept of an ideoscape offers a unique perspective compared to the traditional literature reviews or meta-analyses because it maps out the relationships and connections between key concepts across a corpus of the literature. Traditional reviews basically summarize individual findings from studies, while meta-analyses combine quantitative data to get statistical insights. In contrast, an ideoscape approach visualizes the entire intellectual structure of a research field, showing how ideas, methods, and themes are interlinked. For instance, in this study, we were able to highlight the interplay between cellular processes like proliferation, migration, and differentiation with their clinical applications in bone regeneration. By examining these connections, researchers can uncover gaps in the literature, identify emerging areas of focus, or even spot potential collaborations across traditionally separate subfields. This method provides a more holistic and interactive understanding of research trends, offering new avenues for exploration beyond the scope of traditional reviews.

The methodology we applied—combining topic modeling and knowledge graph creation—can be adapted for targeted queries, offering researchers a flexible and largely automated tool to explore and extract meaningful insights from the vast bodies of literature. This not only enhances the ability to keep up with a fast-moving field but also supports more strategic decision-making in future research endeavors.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/osteology4040013/s1, Supplementary Table S1: The table contains the list of topics identified by BERTopic, in order of size, in our initial exosome dataset. The topic column contains a numerical index for the topic; the Count column indicates the number of articles in the topic; the Name column is BERTopic's default output, i.e., the topic, which is indicated by the index followed by the four most common keywords, conventionally joined by an underscore sign.

The KeyBERT and MMR columns contain the keywords extracted through alternative algorithms (see Section 2): the MMR algorithm, in particular, is designed to maximize diversity in the set of keywords. In the LLM columns, the topic label is elaborated by GPT based on the extracted keywords.; Supplementary Table S2: The table contains the list of topics identified by BERTopic, in order of size, in the Bone regeneration subset of articles (Topic #0 of the initial dataset). The topic column contains a numerical index for the topic; the Count column indicates the number of articles in the topic; the Name column is BERTopic's default output. The KeyBERT and MMR columns contain the keywords extracted through alternative algorithms (see Section 2). In the LLM columns, the topic label is elaborated by GPT based on the extracted keywords.; Figure S1: Neo4j Disease nodes.

Author Contributions: Conceptualization, C.G. and S.G.; methodology, C.G.; formal analysis, C.G. and M.T.C.; data curation, S.G. and M.M.; writing—original draft preparation, M.T.C.; writing—review and editing, M.M.; supervision, S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are available on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lange, M.; Babczyk, P.; Tobiasch, E. Exosomes: A New Hope for Angiogenesis-Mediated Bone Regeneration. *Int. J. Mol. Sci.* **2024**, *25*, 5204. [CrossRef]
2. Théry, C.; Amigorena, S.; Raposo, G.; Clayton, A. Isolation and Characterization of Exosomes from Cell Culture Supernatants and Biological Fluids. *Curr. Protoc. Cell Biol.* **2006**, *30*, 3–22. [CrossRef] [PubMed]
3. Welsh, J.A.; Goberdhan, D.C.I.; O'Driscoll, L.; Buzas, E.I.; Blenkiron, C.; Bussolati, B.; Cai, H.; Di Vizio, D.; Driedonks, T.A.P.; Erdbrügger, U. Minimal Information for Studies of Extracellular Vesicles (MISEV2023): From Basic to Advanced Approaches. *J. Extracell. Vesicles* **2024**, *13*, e12404. [CrossRef] [PubMed]
4. Kalluri, R.; LeBleu, V.S. The Biology, Function, and Biomedical Applications of Exosomes. *Science* **2020**, *367*, eaau6977. [CrossRef] [PubMed]
5. Mathieu, M.; Martin-Jaular, L.; Lavieu, G.; Théry, C. Specificities of Secretion and Uptake of Exosomes and Other Extracellular Vesicles for Cell-to-Cell Communication. *Nat. Cell Biol.* **2019**, *21*, 9–17. [CrossRef] [PubMed]
6. Harrell, C.R.; Jovicic, N.; Djonov, V.; Volarevic, V. Therapeutic Use of Mesenchymal Stem Cell-Derived Exosomes: From Basic Science to Clinics. *Pharmaceutics* **2020**, *12*, 474. [CrossRef]
7. Kučuk, N.; Primožič, M.; Knez, Ž.; Leitgeb, M. Exosomes Engineering and Their Roles as Therapy Delivery Tools, Therapeutic Targets, and Biomarkers. *Int. J. Mol. Sci.* **2021**, *22*, 9543. [CrossRef]
8. Mukherjee, A.; Bisht, B.; Dutta, S.; Paul, M.K. Current Advances in the Use of Exosomes, Liposomes, and Bioengineered Hybrid Nanovesicles in Cancer Detection and Therapy. *Acta Pharmacol. Sin.* **2022**, *43*, 2759–2776. [CrossRef]
9. Arya, S.B.; Collie, S.P.; Parent, C.A. The Ins-and-Outs of Exosome Biogenesis, Secretion, and Internalization. *Trends Cell Biol.* **2024**, *34*, 90–108. [CrossRef]
10. Xue, N.; Ding, X.; Huang, R.; Jiang, R.; Huang, H.; Pan, X.; Min, W.; Chen, J.; Duan, J.-A.; Liu, P. Bone Tissue Engineering in the Treatment of Bone Defects. *Pharmaceuticals* **2022**, *15*, 879. [CrossRef]
11. Egido-Moreno, S.; Valls-Roca-Umbert, J.; Céspedes-Sánchez, J.M.; López-López, J.; Velasco-Ortega, E. Clinical Efficacy of Mesenchymal Stem Cells in Bone Regeneration in Oral Implantology. Systematic Review and Meta-Analysis. *Int. J. Environ. Res. Public Health* **2021**, *18*, 894. [CrossRef] [PubMed]
12. Wang, X.; Thomsen, P. Mesenchymal Stem Cell–Derived Small Extracellular Vesicles and Bone Regeneration. *Basic. Clin. Pharmacol. Toxicol.* **2021**, *128*, 18–36. [CrossRef] [PubMed]
13. Torrecillas-Baena, B.; Pulido-Escribano, V.; Dorado, G.; Gálvez-Moreno, M.Á.; Camacho-Cardenosa, M.; Casado-Díaz, A. Clinical Potential of Mesenchymal Stem Cell-Derived Exosomes in Bone Regeneration. *J. Clin. Med.* **2023**, *12*, 4385. [CrossRef] [PubMed]
14. Tan, S.H.S.; Wong, J.R.Y.; Sim, S.J.Y.; Tjio, C.K.E.; Wong, K.L.; Chew, J.R.J.; Hui, J.H.P.; Toh, W.S. Mesenchymal Stem Cell Exosomes in Bone Regenerative Strategies—A Systematic Review of Preclinical Studies. *Mater. Today Bio* **2020**, *7*, 100067. [CrossRef]
15. Irfan, D.; Ahmad, I.; Patra, I.; Margiana, R.; Rasulova, M.T.; Sivaraman, R.; Kandeel, M.; Mohammad, H.J.; Al-Qaim, Z.H.; Jawad, M.A.; et al. Stem Cell-Derived Exosomes in Bone Healing: Focusing on Their Role in Angiogenesis. *Cytotherapy* **2023**, *25*, 353–361. [CrossRef]
16. Meng, F.; Xue, X.; Yin, Z.; Gao, F.; Wang, X.; Geng, Z. Research Progress of Exosomes in Bone Diseases: Mechanism, Diagnosis and Therapy. *Front. Bioeng. Biotechnol.* **2022**, *10*, 866627. [CrossRef]
17. De Rosnay, J. *Le Macroscope. Vers Une Vision Globale*; Média Diffusion: Paris, France, 2014; ISBN 202118644X.
18. Appadurai, A. *Modernity at Large: Cultural Dimensions of Globalization*; U of Minnesota Press: Minneapolis, MN, USA, 1996; Volume 1, ISBN 145290006X.

19. Kherwa, P.; Bansal, P. Topic Modeling: A Comprehensive Review. *ICST Trans. Scalable Inf. Syst.* **2018**, 159623. [CrossRef]
20. Grootendorst, M. BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure. *arXiv* **2022**, arXiv:2203.05794.
21. Churchill, R.; Singh, L. The Evolution of Topic Modeling. *ACM Comput. Surv.* **2022**, *54*, 1–35. [CrossRef]
22. Liu, Q.; Kusner, M.J.; Blunsom, P. A Survey on Contextual Embeddings. *arXiv* **2020**, arXiv:2003.07278.
23. Wang, Z.; Chen, J.; Chen, J.; Chen, H. Identifying Interdisciplinary Topics and Their Evolution Based on BERTopic. *Scientometrics* **2023**, 1–26. [CrossRef]
24. Börner, K.; Record, E. Macroscopes for Making Sense of Science. In Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact, New Orleans, LA, USA, 9–13 July 2017; pp. 1–2.
25. Bassi, S. A Primer on Python for Life Science Researchers. *PLoS Comput. Biol.* **2007**, *3*, e199. [CrossRef] [PubMed]
26. Jia, Z.; Maggioni, M.; Smith, J.; Scarpazza, D.P. Dissecting the NVidia Turing T4 GPU via Microbenchmarking. *arXiv* **2019**, arXiv:1903.07486.
27. Cock, P.J.A.; Antao, T.; Chang, J.T.; Chapman, B.A.; Cox, C.J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B. Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics. *Bioinformatics* **2009**, *25*, 1422–1423. [CrossRef]
28. Mckinney, W. Data Structures for Statistical Computing in Python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; pp. 51–56.
29. Gan, L.; Yang, T.; Huang, Y.; Yang, B.; Luo, Y.Y.; Richard, L.W.C.; Guo, D. Experimental Comparison of Three Topic Modeling Methods with LDA, Top2Vec and BERTopic. In Proceedings of the International Symposium on Artificial Intelligence and Robotics, Atlanta, GA, USA, 8–10 July 2010; Springer: Singapore, 2023; pp. 376–391.
30. Cook, D.A.; Beckman, T.J.; Bordage, G. A Systematic Review of Titles and Abstracts of Experimental Studies in Medical Education: Many Informative Elements Missing. *Med. Educ.* **2007**, *41*, 1074–1081. [CrossRef]
31. Hartley, J. Planning That Title: Practices and Preferences for Titles with Colons in Academic Articles. *Libr. Inf. Sci. Res.* **2007**, *29*, 553–568. [CrossRef]
32. Galli, C.; Donos, N.; Calciolari, E. Performance of 4 Pre-Trained Sentence Transformer Models in the Semantic Query of a Systematic Review Dataset on Peri-Implantitis. *Information* **2024**, *15*, 68. [CrossRef]
33. Guizzardi, S.; Colangelo, M.T.; Mirandola, P.; Galli, C. Modeling New Trends in Bone Regeneration, Using the BERTopic Approach. *Regen. Med.* **2023**, *18*, 719–734. [CrossRef]
34. Saif, H.; Fernandez, M.; He, Y.; Alani, H. On Stopwords, Filtering and Data Sparsity for Sentiment Analysis of Twitter. In Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, Iceland, 26–31 May 2014.
35. Gutiérrez, L.; Keith, B. A Systematic Literature Review on Word Embeddings. In *Trends and Applications in Software Engineering, Proceedings of the 7th International Conference on Software Process Improvement (CIMPS 2018), Guadalajara, Mexico, 17–19 October 2018*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 132–141.
36. Jurafsky, D.; Martin, J.H.; Vector Semantics and Embeddings. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models, 3rd ed. *Online manuscript released 20 August 2024.* Available online: https://web.stanford.edu/~jurafsky/slp3 (accessed on 30 September 2024).
37. Wang, S.; Zhou, W.; Jiang, C. A Survey of Word Embeddings Based on Deep Learning. *Computing* **2020**, *102*, 717–740. [CrossRef]
38. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. *Adv. Neural Inf. Process Syst.* **2017**, *30*, 6000–6010.
39. Qaiser, S.; Ali, R. Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents. *Int. J. Comput. Appl.* **2018**, *181*, 25–29. [CrossRef]
40. Xu, D.D.; Wu, S.B. An Improved TFIDF Algorithm in Text Classification. *Appl. Mech. Mater.* **2014**, *651*, 2258–2261. [CrossRef]
41. McInnes, L.; Healy, J.; Melville, J. Umap: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv* **2018**, arXiv:1802.03426.
42. McInnes, L.; Healy, J.; Astels, S. Hdbscan: Hierarchical Density Based Clustering. *J. Open Source Softw.* **2017**, *2*, 205. [CrossRef]
43. Haider, M.M.; Hossin, M.A.; Mahi, H.R.; Arif, H. Automatic Text Summarization Using Gensim Word2vec and K-Means Clustering Algorithm. In Proceedings of the 2020 IEEE Region 10 Symposium (TENSYMP), Dhaka, Bangladesh, 5–7 June 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 283–286.
44. Issa, B.; Jasser, M.B.; Chua, H.N.; Hamzah, M. A Comparative Study on Embedding Models for Keyword Extraction Using KeyBERT Method. In Proceedings of the 2023 IEEE 13th International Conference on System Engineering and Technology (ICSET), Shah Alam, Malaysia, 2 October 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 40–45.
45. Zhang, Y.; Jin, R.; Zhou, Z.-H. Understanding Bag-of-Words Model: A Statistical Framework. *Int. J. Mach. Learn. Cybern.* **2010**, *1*, 43–52. [CrossRef]
46. Bennani-Smires, K.; Musat, C.; Hossmann, A.; Baeriswyl, M.; Jaggi, M. Simple Unsupervised Keyphrase Extraction Using Sentence Embeddings. *arXiv* **2018**, arXiv:1801.04470.
47. Thirunavukarasu, A.J.; Ting, D.S.J.; Elangovan, K.; Gutierrez, L.; Tan, T.F.; Ting, D.S.W. Large Language Models in Medicine. *Nat. Med.* **2023**, *29*, 1930–1940. [CrossRef]
48. Akre, P.; Malu, R.; Jha, A.; Tekade, Y.; Bisen, W. Sentiment Analysis Using Opinion Mining on Customer Review. *Int. J. Eng. Manag. Res.* **2023**, *13*, 41–44.

49. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]
50. Waskom, M. Seaborn: Statistical Data Visualization. *J. Open Source Softw.* **2021**, *6*, 3021. [CrossRef]
51. McInnes, L. DataMapPlot. Available online: https://github.com/TutteInstitute/datamapplot (accessed on 10 March 2024).
52. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef] [PubMed]
53. Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; Yu, P.S. A Survey on Knowledge Graphs: Representation, Acquisition, and Applications. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 494–514. [CrossRef] [PubMed]
54. Honnibal, M.; Montani, I. SpaCy 2: Natural Language Understanding with Bloom Embeddings, Convolutional Neural Networks and Incremental Parsing. *To Appear* **2017**, *7*, 411–420.
55. Miller, J.J. Graph Database Applications and Concepts with Neo4j. In Proceedings of the Southern Association for Information Systems Conference, Atlanta, GA, USA, 15–17 August 2013; Volume 2324, pp. 141–147.
56. Chen, X.; Yang, N.; Li, B.; Gao, X.; Wang, Y.; Wang, Q.; Liu, X.; Zhang, Z.; Zhang, R. Visualization Analysis of Small Extracellular Vesicles in the Application of Bone-Related Diseases. *Cells* **2024**, *13*, 904. [CrossRef]
57. Sarker, I.H. LLM Potentiality and Awareness: A Position Paper from the Perspective of Trustworthy and Responsible AI Modeling. *Discov. Artif. Intell.* **2024**, *4*, 40. [CrossRef]
58. Rizzo, M.G.; Cai, N.; Constantinescu, D. The Performance of ChatGPT on Orthopaedic In-Service Training Exams: A Comparative Study of the GPT-3.5 Turbo and GPT-4 Models in Orthopaedic Education. *J. Orthop.* **2024**, *50*, 70–75. [CrossRef]
59. Galli, C.; Cusano, C.; Meleti, M.; Donos, N. Topic Modeling for Faster Literature Screening Using Transformer-Based Embeddings. *Preprints* **2024**, 2024072198. [CrossRef]
60. Luo, D.; Zhu, H.; Li, S.; Wang, Z.; Xiao, J. Mesenchymal Stem Cell-Derived Exosomes as a Promising Cell-Free Therapy for Knee Osteoarthritis. *Front. Bioeng. Biotechnol.* **2024**, *12*, 1309946. [CrossRef]
61. Sankaranarayanan, J.; Lee, S.C.; Kim, H.K.; Kang, J.Y.; Kuppa, S.S.; Seon, J.K. Exosomes Reshape the Osteoarthritic Defect: Emerging Potential in Regenerative Medicine—A Review. *Int. J. Stem Cells* **2024**. [CrossRef]
62. Wang, L.; Li, F.; Wang, L.; Wu, B.; Du, M.; Xing, H.; Pan, S. Exosomes Derived from Bone Marrow Mesenchymal Stem Cells Alleviate Rheumatoid Arthritis Symptoms via Shuttling Proteins. *J. Proteome Res.* **2024**, *23*, 1298–1312. [CrossRef]
63. Wen, S.; Huang, X.; Ma, J.; Zhao, G.; Ma, T.; Chen, K.; Huang, G.; Chen, J.; Shi, J.; Wang, S. Exosomes Derived from MSC as Drug System in Osteoarthritis Therapy. *Front. Bioeng. Biotechnol.* **2024**, *12*, 1331218. [CrossRef] [PubMed]
64. Hong, H.; Lin, C.; Fang, M.; Liu, J.; Hsu, H.-C.; Chang, C.-J.; Wang, H. Proteomic Analysis of Exosomal Proteins Associated with Bone Healing Speed in a Rat Tibial Fracture Model. *Biomed. Chromatogr.* **2024**, *38*, e5846. [CrossRef] [PubMed]
65. Lv, S.; Wang, G.; Dai, L.; Wang, T.; Wang, F. Cellular and Molecular Connections Between Bone Fracture Healing and Exosomes. *Physiol. Res.* **2023**, *72*, 565–574. [CrossRef] [PubMed]
66. Pan, Y.; Li, Y.; Dong, W.; Jiang, B.; Yu, Y.; Chen, Y. Role of Nano-Hydrogels Coated Exosomes in Bone Tissue Repair. *Front. Bioeng. Biotechnol.* **2023**, *11*, 1167012. [CrossRef]
67. Li, X.; Si, Y.; Liang, J.; Li, M.; Wang, Z.; Qin, Y.; Sun, L. Enhancing Bone Regeneration and Immunomodulation via Gelatin Methacryloyl Hydrogel-Encapsulated Exosomes from Osteogenic Pre-Differentiated Mesenchymal Stem Cells. *J. Colloid Interface Sci.* **2024**, *672*, 179–199. [CrossRef]
68. Puletic, M.; Velikic, G.; Maric, D.M.; Supic, G.; Maric, D.L.; Radovic, N.; Avramov, S.; Vojvodic, D. Clinical Efficacy of Extracellular Vesicle Therapy in Periodontitis: Reduced Inflammation and Enhanced Regeneration. *Int. J. Mol. Sci.* **2024**, *25*, 5753. [CrossRef] [PubMed]
69. Han, P.; Raveendran, N.; Liu, C.; Basu, S.; Jiao, K.; Johnson, N.; Moran, C.S.; Ivanovski, S. 3D Bioprinted Small Extracellular Vesicles from Periodontal Cells Enhance Mesenchymal Stromal Cell Function. *Biomater. Adv.* **2024**, *158*, 213770. [CrossRef]
70. Han, P.; Johnson, N.; Abdal-Hay, A.; Moran, C.S.; Salomon, C.; Ivanovski, S. Effects of Periodontal Cells-Derived Extracellular Vesicles on Mesenchymal Stromal Cell Function. *J. Periodontal Res.* **2023**, *58*, 1188–1200. [CrossRef]
71. Kong, Q.; Wang, Y.; Jiang, N.; Wang, Y.; Wang, R.; Hu, X.; Mao, J.; Shi, X. Exosomes as Promising Therapeutic Tools for Regenerative Endodontic Therapy. *Biomolecules* **2024**, *14*, 330. [CrossRef]
72. Zhang, S.; Wang, S.; Chen, J.; Cui, Y.; Lu, X.; Xiong, S.; Yue, C.; Yang, B. Human Dental Pulp Stem Cell-Derived Exosomes Decorated Titanium Scaffolds for Promoting Bone Regeneration. *Colloids Surf. B Biointerfaces* **2024**, *235*, 113775. [CrossRef]
73. Zou, J.; Xia, H.; Jiang, Q.; Su, Z.; Wen, S.; Liang, Z.; Ouyang, Y.; Liu, J.; Zhang, Z.; Chen, D.; et al. Exosomes Derived from Odontogenic Stem Cells: Its Role in the Dentin-Pulp Complex. *Regen. Ther.* **2023**, *24*, 135–146. [CrossRef]
74. Ma, S.; Ma, B.; Yang, Y.; Mu, Y.; Wei, P.; Yu, X.; Zhao, B.; Zou, Z.; Liu, Z.; Wang, M.; et al. Functionalized 3D Hydroxyapatite Scaffold by Fusion Peptides-Mediated Small Extracellular Vesicles of Stem Cells for Bone Tissue Regeneration. *ACS Appl. Mater. Interfaces* **2024**, *16*, 3064–3081. [CrossRef]
75. Liu, Y.; Zhang, Z.; Ma, C.; Song, J.; Hu, J.; Liu, Y. Transplanted MSCs Promote Alveolar Bone Repair via Hypoxia-Induced Extracellular Vesicle Secretion. *Oral Dis.* **2024**, 1–11. [CrossRef] [PubMed]
76. Man, K.; Eisenstein, N.M.; Hoey, D.A.; Cox, S.C. Bioengineering Extracellular Vesicles: Smart Nanomaterials for Bone Regeneration. *J. Nanobiotechnol.* **2023**, *21*, 137. [CrossRef] [PubMed]
77. Song, X.; Xu, L.; Zhang, W. Biomimetic Synthesis and Optimization of Extracellular Vesicles for Bone Regeneration. *J. Control. Release* **2023**, *355*, 18–41. [CrossRef] [PubMed]

78. Hu, Y.-C.; Zhang, X.-B.; Lin, M.-Q.; Zhou, H.-Y.; Cong, M.-X.; Chen, X.-Y.; Zhang, R.-H.; Yu, D.-C.; Gao, X.-D.; Guo, T.-W. Nanoscale Treatment of Intervertebral Disc Degeneration: Mesenchymal Stem Cell Exosome Transplantation. *Curr. Stem Cell Res. Ther.* **2023**, *18*, 163–173. [CrossRef]
79. Xu, G.; Lu, X.; Liu, S.; Zhang, Y.; Xu, S.; Ma, X.; Xia, X.; Lu, F.; Zou, F.; Wang, H.; et al. MSC-Derived Exosomes Ameliorate Intervertebral Disc Degeneration By Regulating the Keap1/Nrf2 Axis. *Stem Cell Rev. Rep.* **2023**, *19*, 2465–2480. [CrossRef]
80. Wang, C.; Guo, S.; Gu, Q.; Wang, X.; Long, L.; Xiao, C.; Xie, M.; Shen, H.; Li, S. Exosomes: A Promising Therapeutic Strategy for Intervertebral Disc Degeneration. *Exp. Gerontol.* **2022**, *163*, 111806. [CrossRef]
81. Francis, N.; Green, A.; Guagliardo, P.; Libkin, L.; Lindaaker, T.; Marsault, V.; Plantikow, S.; Rydberg, M.; Selmer, P.; Taylor, A. Cypher: An Evolving Query Language for Property Graphs. In Proceedings of the 2018 International Conference on Management of Data, Kanazawa, Japan, 28 June–2 July 2018; pp. 1433–1445.
82. Hogan, A.; Blomqvist, E.; Cochez, M.; D'amato, C.; De Melo, G.; Gutierrez, C.; Kirrane, S.; Gayo, J.E.L.; Navigli, R.; Neumaier, S.; et al. Knowledge Graphs. *ACM Comput. Surv.* **2022**, *54*, 1–37. [CrossRef]
83. Peng, C.; Xia, F.; Naseriparsa, M.; Osborne, F. Knowledge Graphs: Opportunities and Challenges. *Artif. Intell. Rev.* **2023**, *56*, 13071–13102. [CrossRef]
84. Wang, T.; Zhou, Y.; Zhang, W.; Xue, Y.; Xiao, Z.; Zhou, Y.; Peng, X. Exosomes and Exosome Composite Scaffolds in Periodontal Tissue Engineering. *Front. Bioeng. Biotechnol.* **2024**, *11*, 1287714. [CrossRef] [PubMed]
85. Trubiani, O.; Pizzicannella, J.; Caputi, S.; Marchisio, M.; Mazzon, E.; Paganelli, R.; Paganelli, A.; Diomede, F. Periodontal Ligament Stem Cells: Current Knowledge and Future Perspectives. *Stem Cells Dev.* **2019**, *28*, 995–1003. [CrossRef] [PubMed]
86. Ju, R.; Gao, X.; Zhang, C.; Tang, W.; Tian, W.; He, M. Exogenous MSCs Based Tissue Regeneration: A Review of Immuno-Protection Strategies from Biomaterial Scaffolds. *J. Mater. Chem. B* **2024**, *36*, 8868–8882. [CrossRef] [PubMed]
87. Zeng, W.-Y.; Ning, Y.; Huang, X. Advanced Technologies in Periodontal Tissue Regeneration Based on Stem Cells: Current Status and Future Perspectives. *J. Dent. Sci.* **2021**, *16*, 501–507. [CrossRef]
88. Isamorad, F.; Ouhestani, F.; Aghandeh, P.; Motamedian, S.R. Application of Periodontal Ligament Stem Cells in Periodontal Regeneration: A Systematic Review. *J. Appl. Biotechnol. Rep.* **2023**, *10*, 1055–1068.