



Afarand Scholarly  
Publishing Institute



Deputy of Health,  
Rescue & Treatment

## Toxins as Biological Weapons; a Text-Mining Approach to Biomedical Literature

### ARTICLE INFO

#### Article Type

Systematic Review

#### Authors

Moazamiyanfar R.<sup>1</sup> MSc,  
Rouhani Nejad H.\* PhD

#### How to cite this article

Moazamiyanfar R, Rouhani Nejad H. Toxins as Biological Weapons; a Text-Mining Approach to Biomedical Literature. Journal of Police Medicine. 2018;7(1):45-50.

### ABSTRACT

**Aims** Bioterrorism is an invasive attack that can cause disease or death in humans, using viruses, bacteria or toxic substances. In recent years, due to an increase in the number of online articles in databases, much attention has been paid to the application of text mining and information extraction strategies from biomedical articles. The purpose of this study was to evaluate the importance of toxins as biological weapons by searching in medical texts and medical databases.

**Information & Methods** This text mining and data research study was carried out in 2015-2016. The Carrot 2 software was used to cluster the keyword search results into the network and especially the biomedical databases. When searching for keywords, the search engine was set up on PubMed and the cluster type was based on K-means, and at the end of the results, the results were considered as foam trees.

**Findings** The highest neurotoxins record was related to tetrodotoxin with a total record of 18970, The highest cytotoxins record was related to Pertussis toxin with a total record of 14390, The highest dermally hazardous cytotoxins record was related to Zearalenone with a total record of 2656.

**Conclusion** Priorities of Bioterrorism Tracking and Biological Warfare Tracking are early detection, public health and control of these agents, So the micro and macro policies should focus on these.

**Keywords** Biological Warfare; Text Mining; Neurotoxins

\*Institute of Biotechnology & Bioscience, Malek-Ashtar University of Technology, Tehran, Iran

<sup>1</sup>Department of Nano Technology, Advance Sciences & Technology Faculty, Pharmaceutical Sciences Branch, Islamic Azad University, Tehran, Iran

#### Correspondence

Address: Institute of Biotechnology & Bioscience, Malek-Ashtar University of Technology, Lavizan, Shahid Babaee Highway, Tehran, Iran. Post-al Code: 158751774  
Phone: +98 (21) 26754457  
Fax: +98 (21) 22974604  
rohaninejhad@gmail.com

#### Article History

Received: December 10, 2016  
Accepted: August 9, 2017  
ePublished: January 3, 2018

### CITATION LINKS

[1] The looming threat of bioterrorism [2] Bioterrorism [3] Biological warfare: A historical perspective [4] Bioterrorism and critical care [5] Biological warfare, biodefence and the biological and toxin weapons convention [6] History of biological warfare and bioterrorism [7] Analysis of biological processes and diseases using text mining approaches [8] Biomedical text mining and its applications [9] Mining for gems of information [10] Text mining the biomedical literature for identification of potential virus/bacterium as bio-terrorism weapons [11] Undiscovered public knowledge [12] The history of biologic warfare and bioterrorism [13] A nursing qualitative systematic review required MEDLINE and CINAHL for study identification [14] Database resources of the National Center for Biotechnology Information [15] Comparison of PubMed, Scopus, web of science, and Google scholar: Strengths and weaknesses [16] Text classification using string kernels [17] Effective term based text clustering algorithms [18] Berry and Jacob Kogan, editors. Text mining: Applications and theory [19] A survey on terrorist network mining: Current trends and opportunities [20] Total decontamination cost of the anthrax letter attacks [21] Potential biological weapons threats [22] Bioterrorism-related inhalational anthrax: The first 10 cases reported in the United States [23] A new role for scientists in the Biological Weapons Convention [24] Terror Detection Using Text Mining

## سموم به عنوان سلاح‌های زیستی؛ رویکرد متن‌کاوی ادبیات زیست‌پزشکی

رضا معظیان فر MSc

گروه نانو تکنولوژی، دانشکده علوم و فناوری‌های نوین، واحد علوم دارویی، دانشگاه آزاد اسلامی، تهران، ایران

حمیده روحانی‌نژاد\* PhD

پژوهشکده علوم و فناوری‌های زیستی، دانشگاه صنعتی مالک‌اشتر، تهران، ایران

### چکیده

**اهداف:** بیوتوریسم حمله حساب‌شده‌ای است که منجر به ایجاد بیماری یا مرگ در انسان، با استفاده از ویروس‌ها، باکتری‌ها یا مواد سمی می‌شود. در سال‌های اخیر، به دلیل افزایش رکورد مقالات آنلاین موجود در پایگاه‌های اطلاعاتی، به کاربرد متن‌کاوی و استراتژی‌های استخراج اطلاعات از مقالات زیست‌پزشکی، توجه بسیاری شده است. هدف از این مطالعه، بررسی اهمیت سموم به عنوان سلاح‌های زیستی با جستجو در متون پزشکی و پایگاه‌های اطلاعاتی پزشکی بود. **اطلاعات و روش‌ها:** این پژوهش متن‌کاوی و بررسی داده‌ای در سال ۱۳۹۵-۱۳۹۶ انجام شد. به منظور خوشه‌بندی نتایج جستجوی کلیدواژه‌های اختصاصی در شبکه و به‌خصوص پایگاه‌های اطلاعاتی پزشکی از نرم‌افزار آنلاین Carrot 2 استفاده شد. در زمان جستجوی کلیدواژه‌ها، موتور جستجوگر PubMed بر نوع خوشه‌بندی بر اساس K-means تنظیم شد و در پایان نتایج به صورت نمودارهای حبابی مورد بررسی قرار گرفت.

**یافته‌ها:** بیشترین رکورد نورو توتوکسین‌ها مربوط به تترودوتوکسین با مجموع رکورد ۱۸۹۷۰، بیشترین رکورد سایتوتوکسین‌ها مربوط به پرتوسیس‌توکسین با مجموع رکورد ۱۴۳۹۰ و بیشترین رکورد سایتوتوکسین‌های خطرناک پوستی مربوط به زیرالون با مجموع رکورد ۲۶۵۶ بود.

**نتیجه‌گیری:** اولویت‌های حوزه پیگیری بیوتوریسم و سلاح‌های زیستی، تشخیص زودرس، سلامت عمومی و کنترل این عوامل است؛ در نتیجه سیاست‌گذاری‌های خرد و کلان می‌بایست معطوف به این زمینه‌ها شود.

**کلیدواژه‌ها:** سلاح‌های زیستی، متن‌کاوی، نورو توتوکسین‌ها

تاریخ دریافت: ۱۳۹۵/۰۹/۲۰

تاریخ پذیرش: ۱۳۹۶/۰۵/۱۸

\*نویسنده مسئول: rohaninejhad@gmail.com

### مقدمه

بیوتوریسم یک حمله حساب‌شده است که منجر به ایجاد بیماری یا مرگ در موجودات زنده از جمله انسان، با استفاده از ویروس‌ها، باکتری‌ها یا مواد سمی می‌شود. این عوامل ممکن است به‌طور طبیعی وجود داشته باشند یا به منظور افزایش قدرت انتشار و بیماری‌زایی توسط انسان دستکاری شوند. سلاح‌های زیستی می‌توانند از طریق هوا، آب یا غذا منتشر شوند. دشواری بودن شناسایی سلاح‌های زیستی و همچنین تأخیر چندین‌روزه در مشخص شدن تأثیر این عوامل در ایجاد بیماری، از مهم‌ترین دلایل استفاده تروریست‌ها از این سلاح‌ها است [1,2].

هرگونه ماده سمی که توسط حیوان، گیاه یا میکروب تولید شود، سم نامیده می‌شود. سنتز اکثر سموم در مقادیر گسترده و به‌واسطه مواد شیمیایی سنتی، امر دشواری است؛ زیرا سموم، ترکیبات پیچیده‌ای دارند، با این حال امکان تولید گسترده آنها توسط مخازن غنی‌شده یا میکروب‌های مهندسی ژنتیک‌شده وجود دارد. همواره ادعای استفاده از سموم در میدان جنگ و ترورهای سیاسی مطرح شده است. محققان فهرستی از مهم‌ترین سموم مورد استفاده در بیوتوریسم را ارائه داده‌اند [3-5].

واکنش سموم در بدن، ترکیبی از اثر فلج‌کنندگی و کشندگی است. اکثر سموم از نظر نظامی می‌توانند در دسته طبقه‌بندی شوند؛ دسته اول، نورو توتوکسین‌هایی که باعث اختلال در سیستم عصبی و

روند انتقال پالس‌های عصبی می‌شوند و دسته دوم سایتوتوکسین‌هایی هستند که سبب اختلال در فرآیند سلولی و در نتیجه مرگ سلولی می‌شوند. علائم این دسته از سموم با مواد التهاب‌آور و تاول‌زا یا علائم مسمومیت غذایی شباهت دارد. برخی سموم نیز ممکن است علائم ترکیبی از دسته‌بندی‌های فوق داشته باشند [3]. تعداد کمی از سموم نیز نظیر مایکوتوکسین‌ها، با آلودگی پوست یا چشم سبب تولید ضایعات پوستی و بیماری‌های سیستمیک می‌شوند.

به‌طور کلی، سموم علائم هشداردهنده خوبی ندارند زیرا غیرفزار و بدون‌بو هستند و در برخی موارد حتی سبب تحریک پوست و چشم هم نمی‌شوند. البته نحوه آلوده‌شدن با سم و دوز آن نیز می‌تواند در شدت بروز علائم و اثرات حاصل از آن و دوره کمون نیز مؤثر باشد. بسیاری از نورو توتوکسین‌ها در تماس با هرگونه ساییدگی یا بریدگی پوستی منجر به درد شدیدی می‌شوند. شدیدترین اثرات، مربوط به زمانی است که سم از راه‌های تنفسی اثر بگذارد. آلودگی توسط برخی سموم، به‌خصوص سایتوتوکسین‌ها ممکن است در عرض چند دقیقه رخ دهد اما علائم آنها پس از گذشت چندین ساعت نیز ممکن است نمایان نشود.

برخی از سموم به عنوان میانجی‌های زیستی عمل می‌کنند و بدن را وادار به انتشار مقادیر زیادی از انواع مواد شیمیایی می‌کنند که به‌طور طبیعی، در حد نرمال در بدن تولید می‌شوند و مقادیر بیشتر آن برای بدن مضر است [4, 5]. اگرچه سموم در دسته سلاح‌های زیستی طبقه‌بندی می‌شوند اما به دلیل این‌که زنده نیستند و نمی‌توانند مانند عوامل بیماری‌زا، همانندسازی نمایند، نوعی ماده شیمیایی محسوب می‌شوند. سموم واگیردار نیستند و یک فرد باید در تماس مستقیم با سم قرار گرفته باشد تا به بیماری دچار شود. سموم ممکن است در حلال مدنظر حل شوند و در قالب محلول‌های رقیقی که خطرات پوستی را در پی خواهند داشت، منتقل شوند [3, 6].

در دهه اخیر، متن‌کاوی در پزشکی و زیست‌شناسی مورد توجه قرار گرفته است. این علم در واقع از استخراج دانش از بانک‌های اطلاعاتی (KDD: Knowledge Discovery In Database) به‌وجود آمده است. متن‌کاوی نوین شامل کشف واقعیت‌های قبلی برگرفته‌شده از متون و تولید فرضیه‌ها است. به عنوان مثال با جست‌وجو پیرامون چکیده‌ها و متون مرتبط با علم پزشکی در ابزار مبتنی بر شبکه MEDLINE، با فرضیه‌های جدید و توضیحاتی در این رابطه روبه‌رو می‌شویم [7, 8]. تمامی سازمان‌ها و نهادهایی نظیر دولت و مؤسسات اجرایی، نهادهای آموزشی و ارگان‌های اقتصادی، نیاز به ذخیره، بازیابی و تحلیل متون مرتبط با فعالیت‌های اجرایی خود دارند. بنابراین، تعداد زیادی از سازمان‌ها که معمولاً در حوزه‌های اجتماعی، سیاسی، صنعتی و پزشکی به فعالیت می‌پردازند، با مدیریت مستندات متفاوت و وظایف تحلیل متن روبه‌رو هستند [9].

مسئله کاوش صریح دانش از مقالات زیست‌پزشکی، اولین بار توسط دکتر سوآنسون در سال ۱۹۸۶ مطرح شد. یکی دیگر از کاربردهای متن‌کاوی در پزشکی، کاوش متون پزشکی و استخراج اطلاعات مرتبط با بیماری‌ها است [10, 11]. این سیستم‌ها در نهایت می‌توانند تحلیلی از نتایج به‌دست‌آمده که به چگونگی تصمیم‌گیری در نحوه اقدامات پیشگیرانه منجر می‌شوند، ارائه دهند. شبکه هوشمند تندرستی عمومی سراسری (GPHIN: Global Public Health Intelligence Network)، به منظور اعلان به موقع هشدارها در رابطه با بیماری‌های مسری، تهدیدات شیمیایی یا

تقسیم اشیاء به‌گونه‌ای است که بیشترین شباهت موضوعی در یک گروه و بیشترین تفاوت با اشیاء گروه‌های دیگر وجود داشته باشد. در این روش، داده‌ها به‌نحوی که در هر زیرگروه ویژگی‌های مشابه بر مبنای فاصله تعریف شده دارند، دسته‌بندی می‌شوند. در طبقه‌بندی، هر داده به یک طبقه از کلاس پیشین تخصیص داده می‌شود ولی در خوشه‌بندی هیچ اطلاعی از کلاس‌های موجود درون داده‌ها وجود ندارد و به عبارتی خود خوشه‌ها نیز از داده‌ها استخراج می‌شوند<sup>[16]</sup>.

الگوریتم‌های K-means، حداکثرسازی امید ریاضی (EM: Expectation Maximization) و خوشه‌بندی سلسله‌مراتبی (HAC: Hierarchical Clustering) از متداول‌ترین روش‌های خوشه‌بندی است<sup>[17]</sup>. در الگوریتم K-means که یک روش پایه برای بسیاری از روش‌های خوشه‌بندی دیگر محسوب می‌شود، ابتدا k عضو که تعداد خوشه‌هاست، به‌صورت تصادفی از میان n عضو به عنوان مراکز خوشه‌ها انتخاب می‌شوند. سپس n-k عضو باقی‌مانده به نزدیک‌ترین خوشه اختصاص می‌یابند. بعد از تخصیص همه اعضا، مراکز خوشه محاسبه می‌شوند و با توجه به مراکز جدید، اعضا به خوشه‌ها تخصیص می‌یابند و این کار تا زمانی که مراکز خوشه‌ها ثابت بماند، ادامه می‌یابد. از جمله مشکلات این روش، وابستگی به‌یابی آن به انتخاب اولیه مراکز و همچنین مشخص‌سازی تعداد خوشه‌ها و صفرشدن خوشه‌ها است<sup>[17, 18]</sup>.

در این پژوهش به منظور خوشه‌بندی نتایج جستجوی کلیدواژه‌های اختصاصی در شبکه و به‌خصوص پایگاه‌های اطلاعاتی پزشکی از نرم‌افزار آنلاین Carrot 2 استفاده شد (<http://search.carrot2.org/stable/search>). این نرم‌افزار از پایگاه‌های اطلاعاتی مختلف مانند PubMed، Wiki، Jobs و PUT یا از جستجوی کلی در شبکه اینترنت استفاده می‌کند و سبب بهبود بصری‌سازی نتایج جستجو در شبکه توسط موتورهای جستجوگر می‌شود؛ علاوه بر این، به کاربر قابلیت انتخاب پایگاه داده موردنظر و الگوریتم‌های خوشه‌بندی را می‌دهد.

به منظور هماهنگی با پایگاه‌های اطلاعاتی پزشکی و مقالات نمایه‌شده در این پایگاه‌ها، در زمان جستجوی کلیدواژه‌ها، موتور جستجوگر بر PubMed و نوع خوشه‌بندی بر اساس K-means تنظیم شد و در پایان نتایج به‌صورت نمودارهای حبابی (Foam Tree) مورد بررسی قرار گرفت<sup>[19]</sup>.

در این پژوهش، بررسی روند مقالات منتشرشده و ارزیابی ارتباط زمانی و اهمیت دوره‌ای بیوتورویسم و سلاح‌های زیستی، با استفاده از مقالات نمایه‌شده در پایگاه اطلاعاتی PubMed انجام

رادیواکتیویته، مقالات روزنامه‌های سراسری و دیگر رسانه‌ها را نظارت می‌کند.

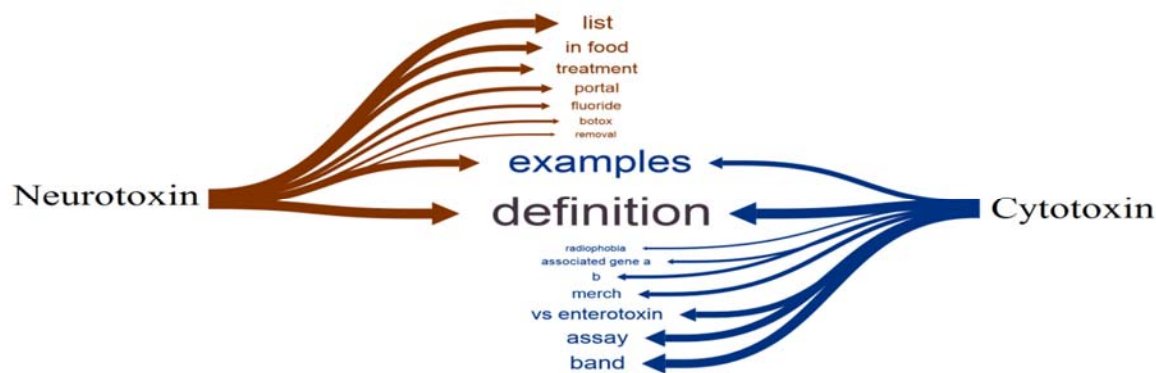
یکی از منابع رایجی که دسترسی به پایگاه داده کتابخانه ملی پزشکی آمریکا در زمینه‌های پرستاری، دندانپزشکی، دامپزشکی و غیره را فراهم می‌کند، PubMed است که از سال ۱۸۷۹ میلادی وظیفه شاخص‌گذاری و فهرست‌بندی مطالب زیست‌پزشکی را بر عهده دارد. این پایگاه به افراد متخصص در حوزه این علوم کمک می‌کند تا به راحتی بتوانند اطلاعات مورد نیاز حوزه خود را در اختیار داشته باشند. پایگاه داده MEDLINE دارای استنادات مجلات و چکیده‌های بیش از ۵۶۰۰ مجله در حیطه زیست‌پزشکی است که از سال ۱۹۹۶ دسترسی آزاد به آن از طریق اینترنت و درگاه PubMed امکان‌پذیر شده است و در آمریکا و سراسر دنیا منتشر می‌شوند. اکثر رکوردهای PubMed، استنادات یا ارجاعات MEDLINE است و سایر رکوردها شامل آنهایی هستند که در مراحل مختلف پردازش قرار دارند (نظیر رکوردهای مستقیم از انتشارات مجلات)، ولی برای استناد در MEDLINE هدف‌گذاری شده‌اند<sup>[12-15]</sup>.

به منظور سیاست‌گذاری‌های خرد و کلان پیش از صرف هرگونه وقت و هزینه می‌بایست از استراتژی متن‌کاوی برای دستیابی به اطلاعات دسته‌بندی‌شده و همچنین تمایز اطلاعات مهم و کاربردی از اطلاعات کم‌اهمیت استفاده شود. از این‌رو این پژوهش در حوزه‌های امنیتی و بهداشت عمومی حائز اهمیت است. عوامل زیادی به عنوان بیوتورویسم شناخته شده‌اند که توانایی اولویت‌بندی کنترل این عوامل در حمله‌های زیستی دارای اهمیت بسیاری است. در این راستا حوزه متن‌کاوی و استراتژی‌های متن‌کاوی می‌توانند مورد استفاده واقع‌شده و در تعیین اولویت‌ها کمک‌کننده باشند.

هدف از این مطالعه، بررسی اهمیت سموم به عنوان سلاح‌های زیستی با جستجو در متون پزشکی و پایگاه‌های اطلاعاتی پزشکی بود.

## اطلاعات و روش‌ها

این پژوهش متن‌کاوی و بررسی داده‌ای در سال ۹۶-۱۳۹۵ انجام شد. با بزرگ‌ترشدن بانک‌های داده‌ای، تلاش محققان برای یافتن روش‌های خوشه‌بندی کارا و مؤثر متمرکز شده است تا از این راه بتوانند زمینه تصمیم‌گیری سریع و منطبق با واقعیت را فراهم آورند. خوشه‌بندی از مهم‌ترین الگوریتم‌های داده‌کاوی است و کاربرد بسیاری در کشف دانش دارد. هدف اصلی خوشه‌بندی،



شکل ۱) دسته‌بندی مقایسه‌ای نتایج و ارتباط میان نوروتوکسین‌ها و سایتوتوکسین‌ها

شد [18]. ابتدا مقایسه‌ای میان نوروپروتوکسین‌ها و سایتوتوکسین‌ها با هدف دسته‌بندی مقایسه‌ای و بصری‌سازی نتایج جستجو، انجام شد (شکل ۱). سپس به ترتیب، کلیدواژه‌های بیوتروریسم (Biological Bioterrorism)، عامل جنگ بیولوژیک (Biological Warfare Agent) و سم جنگ بیولوژیک (Biological Warfare Toxin) در نرم‌افزار خوشه‌بندی آنلاین Carrot 2 و در پایگاه اطلاعاتی بررسی شد.

در ادامه به منظور بررسی اهمیت گروه‌های مختلف سموم اعم از نوروپروتوکسین، سایتوتوکسین و سایتوتوکسین خطرناک پوستی، به بررسی و ارزیابی تعداد رکورد ثبت‌شده از هر یک از سموم در مقالات نمایه‌شده در پایگاه اطلاعاتی PubMed پرداخته شد. فهرست تهیه‌شده از توکسین‌ها شامل مهم‌ترین توکسین‌های مورد استفاده به عنوان سلاح‌های زیستی بود [3]. در انتها نیز کلیدواژه‌های متن‌کاوی (Text Mining)، متن‌کاوی پزشکی (Text Mining) و (Medical) ردیابی بیوتروریسم (Bioterrorism Tracking) و (Biological Warfare Tracking) جنگ بیولوژیک (Biological Warfare Tracking) بررسی شدند.

### یافته‌ها

در ابتدا با جستجوی کلیدواژه بیوتروریسم در ابزار آنلاین خوشه‌بندی، بیشترین رکوردهای ثبت‌شده در رابطه با عناوین نشان داده شدند. پس از آن با جستجوی همین کلیدواژه در پایگاه اطلاعاتی و بررسی روند رکوردهای ثبت‌شده در مقالات پزشکی، بیشترین رکورد به میزان ۸۸۹ و در سال ۲۰۰۲ بود. در ادامه با جستجوی کلیدواژه عامل جنگ بیولوژیک در ابزار آنلاین خوشه‌بندی مشخص شد که بیشترین رکوردهای ثبت‌شده در رابطه با عناوین مرتبط با سموم به عنوان سلاح‌های زیستی و ابزار بیوتروریسم بود. پس از آن با جستجوی همین کلیدواژه در پایگاه اطلاعاتی و بررسی روند رکوردهای ثبت‌شده در مقالات پزشکی، بیشترین رکورد به میزان ۱۰۳ و در سال ۲۰۰۲ بود. با بررسی‌های بیشتر در رابطه با سموم به عنوان سلاح‌های زیستی و جستجوی کلیدواژه سم جنگ بیولوژیک، بیشترین رکورد به میزان ۶۲ و در سال ۲۰۰۲ بود.

جدول ۱) بررسی رکوردهای ثبت‌شده از مهم‌ترین سموم مورد استفاده به عنوان سلاح‌های زیستی در پایگاه PubMed از سال ۲۰۱۶-۱۹۴۱

سموم	مجموع رکورد
نوروپروتوکسین‌ها	
Aconitine	۱۷۳۹
Anatoxin A	۵۴۳
Anatoxin-A(S)	۲۲۸
Batrachotoxin	۵۷۵
Botulinum toxin	۱۷۵۱۴
Brevetoxin	۴۲۳
Bungarotoxin	۴۷۲۰
Ciguatoxin	۵۷۰
Cobrotoxin	۴۹۲
Conotoxin	۴۴۹۵
Diamphotoxin	۳
Gonyautoxin	۱۹۸
α-Latrotoxin	۳۹۰
Neosaxitoxin	۱۲۲
Palytoxin	۴۲۵
Saxitoxin	۱۴۸۷
Taipoxin	۸۰
Tetanus toxin	۴۲۴۶
Tetrodotoxin	۱۸۹۷۰
Tityustoxin	۱۶۴
Veratridine	۲۵۴۷
سایتوتوکسین‌ها	
Abrin	۳۷۹
Aflatoxins	۱۱۱۴
Cardiotoxins	۱۲۸۵
Cholera toxin	۱۳۴۹۶
Clostridium perfringens toxins	۲۵۷۱
Maitotoxin	۳۰۷
Microcystin LR	۱۴۴۹
Modeccin	۸۵
Pertussis toxin	۱۴۳۹۰
Ricin	۳۶۷۶
Shiga toxin	۵۷۲۳
Staphylococcal enterotoxin B	۲۲۸۹
Verotoxins	۴۷۱۰
Viscumin	۱۶۴
Volkensin	۶۱
سایتوتوکسین‌های خطرناک پوستی	
T2 Mycotoxin	۱۵۸۳
Deoxyrnivalenol	۲۶۲۱
Diacetoxyscirpenol	۴۱۸
Fusarenon X	۲۰۶
Neosolaniol	۱۵۴
Nivalenol Diacetate	۱
Roridin A	۶۵
Satratoxin H	۳۹
Verrucaric acid	۱۰۴
Zearalenone	۲۶۵۶



شکل ۲) خوشه‌بندی نتایج جستجوی کلیدواژه‌های متن‌کاوی (بالا) و متن‌کاوی پزشکی (پایین)

### بحث

هدف از این مطالعه، بررسی اهمیت سموم به عنوان سلاح‌های زیستی با جستجو در متون پزشکی و پایگاه‌های اطلاعاتی پزشکی بود. بیشترین رویدادها و حوادث در رابطه با بیوتروریسم و سلاح‌های زیستی در سال‌های ۲۰۰۱ تا ۲۰۰۲ رخ داده است و این موضوع سبب شده تا محققین بسیاری در زمینه‌های مختلف به بررسی این موضوع بپردازند. حوادث ۱۱ سپتامبر سال ۲۰۰۱ در ایالات متحده آمریکا نشان داده است که کشتار جمعی غیرنظامیان، مسأله دور از ذهنی نیست و سلاح‌های کشتار جمعی و مسائل مربوط به تکثیر تهدیدکننده‌های زیستی، شیمیایی و هسته‌ای امور قابل توجهی هستند. در این میان، سلاح‌های زیستی که همانند سلاح‌های هسته‌ای مرگ‌آور هستند و آسان به‌دست می‌آیند، از اهمیت بیشتری برخوردارند. با تحولاتی که در ۲ دهه گذشته رخ داده است، از نظر استراتژیست‌های نظامی و امنیتی، سطح تهدید به‌کارگیری عمدی این عوامل، از حالت بالقوه به بالفعل تبدیل شده است. چنین حملاتی می‌توانند در مدت کوتاهی باعث مرگ میلیون‌ها انسان شوند. سلاح‌های زیستی به دلیل توانایی بالقوه انتشار عوامل بیماری‌زا (همچون آبله، تب زرد، آنتراکس و بیماری‌های عفونی دیگر) و همچنین با توجه به سرعت بالای انتقال بیماری و رسیدن به مرحله بحرانی قبل از شناسایی عامل، از مهم‌ترین عوامل مورد استفاده در حملات بیوتروریسم و کشتار جمعی هستند [20-22].

برای مقابله با چنین حملات مرگباری، روش‌هایی مانند واکسیناسیون افراد در معرض خطر، پیشنهاد شده که البته این مسیر دارای معایبی مانند درخطر انداختن سلامت گروه خاصی از جامعه (بیماران دارای نقص ایمنی، افراد مسن و نوزادان) است. بهترین روش مبارزه با بیوتروریسم، شناسایی و مقابله با آن با استفاده از تکنولوژی‌هایی مانند بیوسنسورها، متن‌کاوی و تکنولوژی مدیریت داده‌ها است [23]. متن‌کاوی نوین شامل تحلیل داده‌های اکتشافی، یعنی کشف واقعیت‌های قبلی برگرفته شده از متون و تولید فرضیه‌ها است. این روش در پزشکی و زیست‌شناسی با موفقیت به‌کار گرفته شده است [24].

از آنجایی که مطالعات متن‌کاوی تماماً به بررسی مقالات صورت‌گرفته در موضوع مشخصی معطوف است و به‌وسیله نرم‌افزار آنلاین انجام می‌پذیرد و نیاز به سخت‌افزار و ابزارآلات آزمایشگاهی ندارد، در نتیجه، محدودیت ویژه‌ای در روند اجرای آن وجود نداشت. البته در برخی موارد، کشور یا پایگاه مبدأ برای کاربران و محققین برخی کشورها از جمله ایران، محدودیت دسترسی ایجاد کرده بودند. با توجه به افزایش تعداد رکورد ثبت‌شده درخصوص مهم‌ترین سموم مورد استفاده به عنوان سلاح‌های زیستی که حاکی از اهمیت این سموم است، پیشنهاد می‌شود تحقیقات گسترده‌تری در این زمینه انجام شود.

### نتیجه‌گیری

اولویت‌های حوزه پیگیری بیوتروریسم و سلاح‌های زیستی، تشخیص زودرس، سلامت عمومی و کنترل این عوامل است، در نتیجه سیاست‌گذاری‌های خرد و کلان می‌بایست معطوف به این زمینه‌ها شود.

تشکر و قدردانی: بدین‌وسیله بر خود لازم می‌دانم از دانشگاه مالک اشتر، پژوهشکده فناوری زیستی به خاطر حمایت‌های مادی و معنوی سپاسگزاری نمایم.

با جستجوی اختصاصی نام هر یک از مهم‌ترین سموم مورد استفاده به عنوان سلاح زیستی، تعداد کل رکوردهای ثبت‌شده در تمامی مقالات نمایه‌شده در پایگاه اطلاعاتی PubMed بررسی شد. بیشترین رکورد نوروکسین‌ها از میان ۲۱ مورد بررسی‌شده مربوط به تترودوتوکسین (Tetrodotoxin) با مجموع ۱۸۹۷۰ رکورد، بیشترین رکورد سابتوتوکسین‌ها از میان ۱۵ مورد بررسی‌شده مربوط به پرتوسیس‌توکسین (Pertussistoxin) با مجموع ۱۴۳۹۰ رکورد و بیشترین رکورد سابتوتوکسین‌های خطرناک پوستی از میان ۱۰ مورد بررسی‌شده مربوط به زیرالنون (Zearalenone) با مجموع ۲۴۵۶ رکورد بود (جدول ۱).

با جستجوی کلیدواژه‌های متن‌کاوی و متن‌کاوی پزشکی در ابزار آنلاین و خوشه‌بندی نتایج این جستجو مشخص شد که بیشترین استفاده از متن‌کاوی در حوزه پزشکی برای پیشگویی سرطان و عوامل مؤثر در آن بود. همچنین با کاربرد متن‌کاوی و الگوریتم‌های جستجوی متن در حوزه پزشکی، بیشترین کاربرد این علم در ثبت وقایع و رویدادهای پزشکی مرتبط با بیماری‌ها و کاربرد ابزارهای الکترونیک برای ثبت اطلاعات پزشکی به منظور کشف اطلاعات خاص از میان داده‌ها بود (شکل ۲).



شکل ۳) خوشه‌بندی نتایج جستجوی کلیدواژه‌های ردیابی بیوتروریسم (بالا) و ردیابی جنگ بیولوژیک (پایین)

در ادامه خوشه‌بندی نتایج با جستجوی آنلاین کلیدواژه‌های ردیابی بیوتروریسم و جنگ بیولوژیک و بررسی نتایج خوشه‌بندی مشخص شد که مهم‌ترین نکته در پیگیری بیوتروریسم توسط متن‌کاوی، تشخیص زودرس آن بود. همچنین موضوع حائز اهمیت در پیگیری سلاح‌های زیستی به عنوان بیوتروریسم، سلامت عمومی، تشخیص زودرس و کنترل این عوامل بود (شکل ۳).

- 1986;56(2):103-18.
- 12- Jacobs MK. The history of biologic warfare and bioterrorism. *Dermatol Clin*. 2004;22(3):231-46.
- 13- Subirana M, Solá I, Garcia JM, Gich I, Urrútia G. A nursing qualitative systematic review required MEDLINE and CINAHL for study identification. *J Clin Epidemiol*. 2005;58(1):20-5.
- 14- Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvermin V, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*. 2007;35:5-12.
- 15- Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, web of science, and Google scholar: Strengths and weaknesses. *The FASEB J*. 2008;22(2):338-42.
- 16- Lodhi H, Saunders C, Shawe-Taylor J, Cristianini N, Watkins C. Text classification using string kernels. *J Mach Learn Res*. 2002;2:419-44.
- 17- Ponmuthuramalingam P, Devi T. Effective term based text clustering algorithms. *Int J Comput Sci Eng*. 2010;2(5):1665-73.
- 18- Xiaojun Z, Michael W, Berry and Jacob Kogan, editors. Text mining: Applications and theory. *Inf Retr*. 2011;14(2):208-11.
- 19- Chaurasia N, Dhakar M, Tiwari A, Gupta RK. A survey on terrorist network mining: Current trends and opportunities. *Int J Comput Sci Eng Surv*. 2012;3(4):59-66.
- 20- Schmitt K, Zacchia NA. Total decontamination cost of the anthrax letter attacks. *Biosecur Bioterror: Biodef Strategy Pract Sci*. 2012;10(1):98-107.
- 21- Kortepeter MG, Parker GW. Potential biological weapons threats. *Emerg Infect Dis*. 1999;5(4):523-7.
- 22- Jernigan JA, Stephens DS, Ashford DA, Omenaca C, Topiel MS, Galbraith M, et al. Bioterrorism-related inhalational anthrax: The first 10 cases reported in the United States. *Emerg Infect Dis*. 2001;7(6):933-44.
- 23- Grönvall GK. A new role for scientists in the Biological Weapons Convention. *Nat Biotechnol*. 2005;23(10):1213-6.
- 24- Kulkarni Sh, Baban Pawar S, Satpute G. Terror Detection Using Text Mining. *Imp J Interdiscip Res*. 2016;2(5):1882-5.
- ۵۰- رضا معظمیان فر و حمیده روحانی نژاد: سهم نویسندگان: رضا معظمیان فر (نویسنده اول)، نگارنده مقدمه/پژوهشگر کمکی/تحلیل‌گر آماری (۳۰٪): حمیده روحانی نژاد (نویسنده دوم)، روش‌شناس/پژوهشگر اصلی/تحلیل‌گر آماری/نگارنده بحث (۷۰٪).
- تأییدیه اخلاقی: موردی از طرف نویسندگان گزارش نشده است.
- تعارض منافع: موردی از طرف نویسندگان گزارش نشده است.
- منابع مالی: موردی از طرف نویسندگان گزارش نشده است.

### منابع

- Henderson DA. The looming threat of bioterrorism. *Science*. 1999;283(5406):1279-82.
- Bellamy RJ, Freedman AR. Bioterrorism. *Int J Med*. 2001;94(4):227-34.
- Christopher LG, Cieslak LT, Pavlin JA, Eitzen EM. Biological warfare: A historical perspective. *Jama*. 1997;278(5):412-7.
- Karwa M, Bronzert P, Kvetan V. Bioterrorism and critical care. *Crit Care Clin*. 2003;19(2):279-313.
- DaSilva EJ. Biological warfare, bioterrorism, biodefence and the biological and toxin weapons convention. *Electron J Biotechnol*. 1999;2(3):3-4.
- Barras V, Greub G. History of biological warfare and bioterrorism. *Clin Microbiol Infect*. 2014;20(6):497-502.
- Krallinger M, Leitner F, Valencia A. Analysis of biological processes and diseases using text mining approaches. *Bioinform Methods Clin Res*. 2010;593:341-82.
- Rodriguez-Esteban R. Biomedical text mining and its applications. *PLoS Comput Biol*. 2009;5(12):e1000597.
- Sirmakessis S, Mining for gems of information. In: Sirmakessis S, editor. Text mining and its applications: Results of the NEMIS Launch Conference. New York: Springer; 2004. pp. 1-6.
- Hu X, Zhang X, Wu D, Zhou X, Rumm P, Text mining the biomedical literature for identification of potential virus/bacterium as bio-terrorism weapons. In: Chen H, Reid E, Sinai J, Silke A, Ganor B, editors, *Terrorism Informatics*. Hamburg: Springer; 2008. pp. 385-406.
- Swanson DR. Undiscovered public knowledge. *Libr Q*.