Poisoned texts report

Contents

T	Sun	nmary	4
	1.1	Basic info	2
	1.2	Shared or genuine lexical items	2
	1.3	File by file coincidence	2
2	Styl	le	3
	2.1	Heatmap	3
	2.2	Most frequent words	3
	2.3	Most frequent words first position in sentence	4
		2.3.1 Single words	4
		2.3.2 Two first words	4
		2.3.3 Three first words	4
		2.3.4 Four first words	4
	2.4	Most frequent upos first position in sentence	5
		2.4.1 Single upos	5
		2.4.2 Two first upos	5
		2.4.3 Three first upos	5
		2.4.4 Four first upos	5
3	Sim	nilar sentences	6
4	Anr	mex (texts)	8
F	ois	soned texts	

Reference: Cabedo, A. (2022) Poisoned texts. V.1.0 beta. Available at: https://github.com/acabedo/poisonedtexts

GNU General Public License v3.0

In this report you will find tables and charts covering the selected texts. Please, take into account that the interpretation it has not only one direction, and that this is a task for a specialist.

1 Summary

In this section, there are two brief summaries. The first one shows the number of paragraphs, sentences and ngrams for each text file. The second one presents the quantity of ngrams shared more than once, just once between at least two files or even ngrams appearing only at one of the texts being analysed.

The selected ngrams for the analysis were 5.

1.1 Basic info

filename	paragraphs	sentences	words
estegosaurio.txt	1	19	600
spinosaurio.txt	1	21	683
tiranosaurio.txt	1	15	422

1.2 Shared or genuine lexical items

ngram	type	freq	total
estegosaurio.txt	once	1	0.23
estegosaurio.txt	only	441	99.77
spinosaurio.txt	once	$_4$	0.76
spinosaurio.txt	only	522	99.24
tiranosaurio.txt	once	5	1.61
tiranosaurio.txt	only	306	98.39

1.3 File by file coincidence

filename	variable	value	total	percent
tiranosaurio.txt	spinosaurio.txt	4	311	1.29
spinosaurio.txt	tiranosaurio.txt	4	526	0.76
tiranosaurio.txt	estegosaurio.txt	1	311	0.32
estegosaurio.txt	tiranosaurio.txt	1	442	0.23
estegosaurio.txt	estegosaurio.txt	0	442	0.00
spinosaurio.txt	estegosaurio.txt	0	526	0.00
estegosaurio.txt	spinosaurio.txt	0	442	0.00
spinosaurio.txt	spinosaurio.txt	0	526	0.00
tiranosaurio.txt	tiranosaurio.txt	0	311	0.00

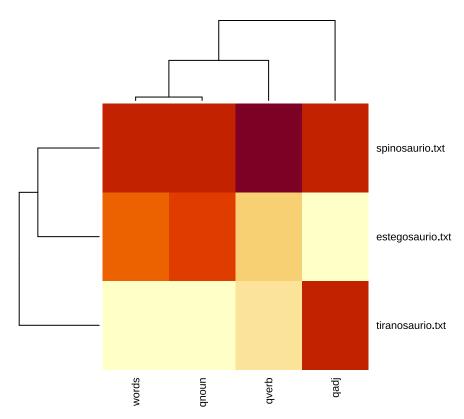
2 Style

This section covers several variables covering common styling characteristics for individuals. These variables must be selected previously on the app. Accordingly, they will be projected on a heatmap.

The frequencies have been automatically scaled with a reference to 100 words, because every sentence has a different quantity of words. Furthermore, the variable words refers to the mean of words by sentence for every text.

doc_id	words	qnoun	qpropn	qcconj	qadv	qprep	qadj	qverb	qdet	qcomma (qcolon
estegosaurio.txt	31.6	19.1	8.5	4.4	5.4	14.3	7.9	6.2	12.1	5.5	0.0
spinosaurio.txt	32.5	19.3	6.2	5.1	2.6	13.1	8.9	8.9	11.9	4.2	0.9
tiranosaurio.txt	28.1	17.8	4.6	3.9	4.4	15.0	8.9	6.0	13.5	3.6	0.1

2.1 Heatmap



2.2 Most frequent words

doc_id	ngram	freq
estegosaurio.txt	de	6
estegosaurio.txt	de	4
spinosaurio.txt	de	5

doc_id	ngram	freq
spinosaurio.txt	de	5
tiranosaurio.txt	de	5
tiranosaurio.txt	del	4

2.3 Most frequent words first position in sentence

2.3.1 Single words

doc_id	token	freq
estegosaurio.txt	Stegosaurus	4
estegosaurio.txt	A	2
spinosaurio.txt	La	3
spinosaurio.txt	El	2
tiranosaurio.txt	Su	2
tiranosaurio.txt	Aunque	1

2.3.2 Two first words

doc_id	ngram	freq
estegosaurio.txt	Debido a	2
estegosaurio.txt	A el	1
spinosaurio.txt	Los huesos	2
spinosaurio.txt	A esta	1
tiranosaurio.txt	Aunque otros	1
${\it tiranosaurio.txt}$	Como otros	1

2.3.3 Three first words

doc_id	ngram	freq
estegosaurio.txt	A el menos	1
estegosaurio.txt	A pesar de	1
spinosaurio.txt	Los huesos de	2
spinosaurio.txt	A esta estructura	1
tiranosaurio.txt	Aunque otros terópodos	1
${\it tiranosaurio.txt}$	Como otros tiranosáuridos	1

2.3.4 Four first words

$\overline{\mathrm{doc}_{-}\mathrm{id}}$	ngram	freq
estegosaurio.txt	A el menos tres	1
estegosaurio.txt	A pesar de su	1
spinosaurio.txt	A esta estructura se	1
spinosaurio.txt	El cráneo del de	1

doc_id	ngram	freq
tiranosaurio.txt	Aunque otros terópodos rivalizan	1
tiranosaurio.txt	Como otros tiranosáuridos ,	1

2.4 Most frequent upos first position in sentence

2.4.1 Single upos

doc_id	upos	freq
estegosaurio.txt	PROPN	5
estegosaurio.txt	DET	4
spinosaurio.txt	DET	11
spinosaurio.txt	ADP	2
tiranosaurio.txt	DET	7
${\it tiranosaurio.txt}$	ADP	2

2.4.2 Two first upos

doc_id	ngram	freq
estegosaurio.txt	DET NOUN	3
estegosaurio.txt	ADP DET	2
spinosaurio.txt	DET NOUN	10
spinosaurio.txt	PRON VERB	2
tiranosaurio.txt	DET NOUN	6
tiranosaurio.txt	ADP DET	1

2.4.3 Three first upos

doc_id	ngram	freq
estegosaurio.txt	DET NOUN ADP	2
estegosaurio.txt	VERB ADP DET	2
spinosaurio.txt	DET NOUN ADJ	3
spinosaurio.txt	DET NOUN ADP	3
tiranosaurio.txt	DET NOUN ADP	4
$\underline{\text{tiranosaurio.txt}}$	ADP DET NOUN	1

2.4.4 Four first upos

doc_id	ngram	freq
estegosaurio.txt	ADP DET ADV NUM	1
estegosaurio.txt	ADP DET NOUN SCONJ	1
spinosaurio.txt	DET NOUN ADP DET	3
spinosaurio.txt	DET NOUN VERB SCONJ	2
tiranosaurio.txt	DET NOUN ADP NOUN	2

doc_id	ngram	freq
tiranosaurio.txt	ADP DET NOUN PUNCT	1

3 Similar sentences

This last section presents the common verbatim concordances for the coinciding ngrams among at least two texts.

doc_sidntengramid to	doc_idntergramid totalentence2		
spinosauriteltxt 2 período cretá- cico hace aprox- i- mada- mente	Spinosaurus aegyptiacus es la única especie conocida del género extinto Spinosaurus gr «lagarto de espina» de dinosaurio terópodo espinosáurido que vivió a mediados del período cretácico hace aproximadamente 99 a 935 millones de años desde el Cenomaniense hasta el Turoniense en lo que es ahora África12345 Este género se conoció inicialmente por los restos fósiles descubiertos en Egipto en la década de 1910 y descritos por el paleontólogo alemán Ernst Stromer6 Los restos originales fueron destruidos durante los bombardeos de la Segunda Guerra Mundial pero se ha recuperado material adicional a principio del siglo XXI		
tirandsauriel.txt 2 período cretá- cico hace aprox- i- mada- mente	Tyrannosaurus rex del griego latinizado tyrannus 'tirano' y saurus 'lagarto' y el latín rex 'rey'1 es la única especie conocida del género fósil Tyrannosaurus de dinosaurio terópodo tiranosáurido que vivió a finales del período cretácico hace aproximadamente entre 68 y 66 millones de años23 en el Maastrichtiense en lo que es hoy América del Norte		
spinosaurės.tat 2 única es- pecie cono- cida	Spinosaurus aegyptiacus es la única especie conocida del género extinto Spinosaurus gr «lagarto de espina» de dinosaurio terópodo espinosáurido que vivió a mediados del período Cretácico hace aproximadamente 99 a 935 millones de años desde el Cenomaniense hasta el Turoniense en lo que es ahora África12345 Este género se conoció inicialmente por los restos fósiles descubiertos en Egipto en la década de 1910 y descritos por el paleontólogo alemán Ernst Stromer6 Los restos originales fueron destruidos durante los bombardeos de la Segunda Guerra Mundial pero se ha recuperado material adicional a principio del siglo XXI		
tiranosaueiolaxt 2 única es- pecie cono- cida	Tyrannosaurus rex del griego latinizado tyrannus 'tirano' y saurus 'lagarto' y el latín rex 'rey'1 es la única especie conocida del género fósil Tyrannosaurus de dinosaurio terópodo tiranosáurido que vivió a finales del período Cretácico hace aproximadamente entre 68 y 66 millones de años23 en el Maastrichtiense en lo que es hoy América del Norte		
estegósaucio.txt 2 uno de los más	Debido a sus distintivas espinas de la cola y placas Stegosaurus es uno de los más reconocidos dinosaurios		
tiran 3 sau e io.txt 2 uno de los más	El debate de si Tyrannosaurus fue un depredador dominante o un carroñero es uno de los más largos en la paleontología		

doc <u>sidntengram</u> id	l tot	asentence2
spinosauria.txt única es- pecie cono- cida del	2	Spinosaurus aegyptiacus es la única especie conocida del género extinto Spinosaurus gr «lagarto de espina» de dinosaurio terópodo espinosáurido que vivió a mediados del período Cretácico hace aproximadamente 99 a 935 millones de años desde el Cenomaniense hasta el Turoniense en lo que es ahora África12345 Este género se conoció inicialmente por los restos fósiles descubiertos en Egipto en la década de 1910 y descritos por el paleontólogo alemán Ernst Stromer6 Los restos originales fueron destruidos durante los bombardeos de la Segunda Guerra Mundial pero se ha recuperado material adicional a principio del siglo XXI
tirandsaulio.txt única es- pecie cono- cida del	2	Tyrannosaurus rex del griego latinizado tyrannus 'tirano' y saurus 'lagarto' y el latín rex 'rey'1 es la única especie conocida del género fósil Tyrannosaurus de dinosaurio terópodo tiranosáurido que vivió a finales del período Cretácico hace aproximadamente entre 68 y 66 millones de años23 en el Maastrichtiense en lo que es hoy América del Norte
estegosausco.txt han en- con- trado restos de	2	Al menos tres especies han sido identificadas en la porción superior de la Formación Morrison del oeste de Estados Unidos y se han encontrado restos de por lo menos 80 individuos2 En 2006 se descubrió un espécimen de Stegosaurus en Portugal lo que indica que este animal también estaba presente en Europa3
estegosausio.txt han en- con- trado restos de	2	se han encontrado restos de más de 80 individuos de este género
	uBus	s Debido a sus distintivas espinas de la cola y placas stegosaurus es uno de los más reconocidos dinosaurios
	u ß us	s Debido a su combinación distintiva de placas anchas y verticales y cola con puntas en la punta stegosaurus es uno de los tipos de dinosaurios más reconocibles
esteg d9 au rtegos a es uno	u ß us	s stegosaurus es uno de los dinosaurios más conocidos y ha aparecido en películas sellos postales y muchos otros tipos de medios
de los spino \$2 ur ima xt es-tructura similar a una	2	Las espinas neurales distintivas del Spinosaurus que eran largas extensiones de las vértebras crecían hasta al menos 165 metros de largo y probablemente tenían piel que las conectaba formando una estructura similar a una vela como la de Dimetrodon aunque algunos autores han sugerido que las espinas estaban cubiertas de grasa y formaban una joroba

doc_sidntengramic	doc_idntengranid totalentence2	
spino \$5 ur im\$ xt es- truc- tura simi- lar a una	2	Su cola larga y estrecha estaba profundizada por espinas neurales altas y delgadas y cheurones alargados formando una aleta flexible o una estructura similar a una paleta
spinosaur ínte at es- pecie cono- cida del género	2	Spinosaurus aegyptiacus es la única especie conocida del género extinto Spinosaurus gr «lagarto de espina» de dinosaurio terópodo espinosáurido que vivió a mediados del período Cretácico hace aproximadamente 99 a 935 millones de años desde el Cenomaniense hasta el Turoniense en lo que es ahora África12345 Este género se conoció inicialmente por los restos fósiles descubiertos en Egipto en la década de 1910 y descritos por el paleontólogo alemán Ernst Stromer6 Los restos originales fueron destruidos durante los bombardeos de la Segunda Guerra Mundial pero se ha recuperado material adicional a principio del siglo XXI
tirandsau ńioita t es- pecie cono- cida del género	2	Tyrannosaurus rex del griego latinizado tyrannus 'tirano' y saurus 'lagarto' y el latín rex 'rey'1 es la única especie conocida del género fósil Tyrannosaurus de dinosaurio terópodo tiranosáurido que vivió a finales del período Cretácico hace aproximadamente entre 68 y 66 millones de años23 en el Maastrichtiense en lo que es hoy América del Norte

4 Annex (texts)

In this section all the texts appear

sentence_id

estegosaurio.txt_1: Stegosaurus (gr. "lagarto con tejado" o "lagarto cubierto"1) es un género con 3, probablemente 4, especies conocidas de dinosaurios tireóforos estegosáuridos, que vivieron a finales del período Jurásico, hace aproximadamente 156 y 144 millones de años, entre el Kimmeridgiense y el Titoniense, en lo que hoy es América del Norte y posiblemente Europa.

estegosaurio.txt_2: Coloquialmente en español llamado estegosaurio.

estegosaurio.txt_3: Debido a sus distintivas espinas de la cola y placas, Stegosaurus es uno de los más reconocidos dinosaurios.

estegosaurio.txt_4: Al menos tres especies han sido identificadas en la porción superior de la Formación Morrison, del oeste de Estados Unidos, y se han encontrado restos de por lo menos 80 individuos.2 En 2006, se descubrió un espécimen de Stegosaurus en Portugal, lo que indica que este animal también estaba presente en Europa.3

estegosaurio.txt_5: De las especies que se han clasificado en la Formación Morrison superior del oeste de los EE.

estegosaurio.txt_6: UU., solo tres son universalmente reconocidas, Stegosaurus stenops, Stegosaurus ungulatus y Stegosaurus sulcatus, con una cuarta Stegosaurus armatus siendo dudosa.

estegosaurio.txt 7: Se han encontrado restos de más de 80 individuos de este género.

estegosaurio.txt_8: Stegosaurus habría vivido junto a dinosaurios como Apatosaurus, Diplodocus , Brachiosaurus , Ceratosaurus, Allosaurus y Torvosaurus, siendo estos últimos sus posibles predadores. estegosaurio.txt_9: Siendo un gran herbívoro cuadrúpedo, pesadamente constituido, Stegosaurus tenía una distintiva e inusual postura, con un lomo fuertemente arqueado, los miembros anteriores cortos, la cabeza cerca del suelo y la cola rígida sostenida en el aire.

sentence id

estegosaurio.txt_10: Debido a su combinación distintiva de placas anchas y verticales y cola con puntas en la punta, Stegosaurus es uno de los tipos de dinosaurios más reconocibles.

estegosaurio.txt 11: Su arsenal de placas y de púas ha sido el tema de muchas conjeturas.

estegosaurio.txt_12: Las espinas fueron utilizadas muy probablemente para la defensa, mientras que las placas también se han propuesto como mecanismo defensivo, y como parte de la exhibición y de las funciones de termorregulación.

estegosaurio.txt_13: Stegosaurus tenía una proporción de masa cerebral a cuerpo relativamente baja. estegosaurio.txt_14: Tenía un cuello corto y una cabeza pequeña, lo que significa que probablemente comía arbustos y plantas bajos.

estegosaurio.txt_15: Una de sus especies, Stegosaurus ungulatus, fue uno de los más grandes estegosaurianos, más grande que Kentrosaurus y que Huayangosaurus y aunque aproximadamente del tamaño de un autobús, compartía muchas características anatómicas, incluyendo a las espinas y las placas traseras, con los otros géneros estegosaurianos.

estegosaurio.txt_16: Los restos de Stegosaurus fueron identificados por primera vez durante las "Guerras de los Huesos" por Othniel Charles Marsh en el Monumento Nacional Dinosaur Ridge.

estegosaurio.txt_17: Los primeros esqueletos conocidos eran fragmentarios y los huesos estaban dispersos, y pasarían muchos años antes de que se entendiera bien la verdadera apariencia de estos animales, incluida su postura y la disposición de las placas.

estegosaurio.txt_18: A pesar de su popularidad en libros y películas, los esqueletos montados de Stegosaurus no se convirtieron en un elemento básico de los principales museos de historia natural hasta mediados del siglo XX, y muchos museos han tenido que ensamblar exhibiciones compuestas de varios especímenes diferentes debido a la falta de esqueletos completos.

estegosaurio.txt_19: Stegosaurus es uno de los dinosaurios más conocidos y ha aparecido en películas, sellos postales y muchos otros tipos de medios.

spinosaurio.txt_1: Spinosaurus aegyptiacus es la única especie conocida del género extinto Spinosaurus (gr. «lagarto de espina») de dinosaurio terópodo espinosáurido, que vivió a mediados del período Cretácico, hace aproximadamente 99 a 93,5 millones de años, desde el Cenomaniense hasta el Turoniense, en lo que es ahora África.12345 Este género se conoció inicialmente por los restos fósiles descubiertos en Egipto en la década de 1910 y descritos por el paleontólogo alemán Ernst Stromer.6 Los restos originales fueron destruidos durante los bombardeos de la Segunda Guerra Mundial, pero se ha recuperado material adicional a principio del siglo XXI.

spinosaurio.txt_2: No existe claridad de si se trata de una o dos especies representadas por los fósiles reportados en la literatura científica.

spinosaurio.txt 3: La especie más conocida es S. aegyptiacus descubierta en Egipto;

spinosaurio.txt 4: sin embargo, una posible segunda especie S. maroccanus se descubrió en Marruecos.

spinosaurio.txt_5: También se han encontrado restos en Argelia, Túnez y posiblemente de Níger y Libia.

spinosaurio.txt_6: La mayoría de las últimas publicaciones consideran que la segunda especie e un sinónimo menor de la primera.

spinosaurio.txt_7: El género contemporáneo de espinosáuridos Sigilmassasaurus también ha sido sinonimizado por algunos autores con S. aegyptiacus, aunque otros investigadores proponen que sea un taxón distinto . Otro posible sinónimo menor es Oxalaia de la Formación Alcântara en Brasil, lo que extendería el rango del génereo hasta Sudamérica . Spinosaurus pudo ser el más grande de todos los dinosaurios carnívoros, más grande aún que Tyrannosaurus rex y Giganotosaurus, a pesar de no compartir la contextura robusta de estos anteriores.

spinosaurio.txt_8: Estimaciones publicadas en 2005, 2007 y 2008 sugieren que tenía 12,6 a 18 metros de longitud y 7 a 20,9 toneladas de peso.378 Nuevas estimaciones publicadas en 2014 y 2018, basadas en un espécimen más completo, respaldaron la investigación anterior y encontraron que el Spinosaurus podría alcanzar longitudes de 15 a 16 metros.910>11

spinosaurio.txt_9: Las últimas estimaciones sugieren un peso de 6,4 a 7,5 toneladas.1011 spinosaurio.txt_10: El cráneo del Spinosaurus era largo, bajo y angosto, similar al de un cocodrilo moderno y tenía dientes cónicos rectos sin estrías.

sentence id

spinosaurio.txt_11: Habría tenido extremidades anteriores grandes y robustas con manos de tres dedos, con una garra agrandada en el primer dígito.

spinosaurio.txt_12: Las espinas neurales distintivas del Spinosaurus, que eran largas extensiones de las vértebras, crecían hasta al menos 1,65 metros de largo y probablemente tenían piel que las conectaba, formando una estructura similar a una vela, como la de Dimetrodon, aunque algunos autores han sugerido que las espinas estaban cubiertas de grasa y formaban una joroba.

spinosaurio.txt_13: A esta estructura se le han atribuido múltiples funciones, incluyendo termorregulación y exhibición.

spinosaurio.txt_14: Los huesos de la cadera de Spinosaurus estaban reducidos y las piernas eran muy cortas en proporción al cuerpo.

spinosaurio.txt $_15$: Su cola larga y estrecha estaba profundizada por espinas neurales altas y delgadas y cheurones alargados , formando una aleta flexible o una estructura similar a una paleta.

spinosaurio.txt_16: Se cree que pudo haberse alimentado de pescado;

spinosaurio. txt_17 : la evidencia sugiere que permanecía tanto en el agua como en tierra, como un cocodrilo moderno.

spinosaurio.txt_18: Se sabe que Spinosaurus comía pescado, y la mayoría de los científicos creen que cazaba presas tanto terrestres como acuáticas.

spinosaurio.txt_19: La evidencia sugiere que era altamente semiacuático y vivía tanto en la tierra como en el agua al igual que los cocodrilos modernos.

spinosaurio.txt_20: Los huesos de las piernas de Spinosaurus tenían osteosclerosis, alta densidad ósea, lo que permitía un mejor control de la flotabilidad y la cola en forma de paleta probablemente se usaba para la propulsión bajo el agua.

spinosaurio.txt_21: Vivió en un ambiente húmedo de planicies intermareales y manglares y bosques junto con muchos otros dinosaurios, así como peces, crocodilomorfos, lagartos, tortugas, pterosaurios y plesiosaurios.

tiranosaurio.txt_1: Tyrannosaurus rex (del griego latinizado tyrannus 'tirano' y saurus 'lagarto', y el latín rex, 'rey'),1 es la única especie conocida del género fósil Tyrannosaurus de dinosaurio terópodo tiranosáurido, que vivió a finales del período Cretácico, hace aproximadamente entre 68 y 66 millones de años,23 en el Maastrichtiense, en lo que es hoy América del Norte.

tiranosaurio.txt_2: Su distribución en el continente fue mucho más amplia que otros tiranosáuridos.

tiranosaurio.txt_3: Comúnmente abreviado como T. rex, y castellanizado como tiranosaurio4 es una figura común en la cultura popular.

tiranosaurio.txt_4: Fue uno de los últimos dinosaurios no avianos que existieron antes de la extinción masiva del Cretácico-Terciario.

tiranosaurio.txt_5: Como otros tiranosáuridos, T. rex fue un carnívoro bípedo con un enorme cráneo equilibrado por una cola larga y pesada.

tiranosaurio.txt_6: En relación con sus largos y poderosos miembros traseros, los miembros superiores de Tyrannosaurus eran pequeños, pero sorprendentemente fuertes para su tamaño, y terminaban en dos dedos con garras.

tiranosaurio.txt_7: Aunque otros terópodos rivalizan o superan a Tyrannosaurus rex en tamaño, todavía es el mayor tiranosáurido conocido y uno de los mayores depredadores conocidos de la Tierra, midiendo entre 12 a 13 metros de largo,5 4 metros de altura hasta las caderas,6 y con pesos estimados entre 6 a 9 toneladas.5 Durante mucho tiempo fue el mayor carnívoro de su ecosistema;

tiranosaurio.txt_8: debió de haber sido el superpredador, cazando hadrosáuridos y ceratópsidos, aunque algunos expertos han sugerido que era principalmente carroñero.

tiranosaurio.txt_9: El debate de si Tyrannosaurus fue un depredador dominante o un carroñero es uno de los más largos en la paleontología.

tiranosaurio.txt_10: Hay más de 30 especímenes de Tyrannosaurus rex identificados, algunos de los cuales son esqueletos casi completos.

tiranosaurio.txt_11: Se han encontrado tejido conjuntivo y proteínas en por lo menos uno de estos especímenes.

sentence_id

tiranosaurio.txt_12: La abundancia de material fósil ha permitido investigar en detalle muchos aspectos de su biología, incluyendo su ciclo de vida y su biomecánica.

tiranosaurio.txt_13: Los hábitos de alimentación, la fisiología y la velocidad potencial de Tyrannosaurus rex son objeto de controversia.

tiranosaurio.txt_14: Su taxonomía es también polémica, con algunos científicos que consideran a Tarbosaurus bataar de Asia como una segunda especie de Tyrannosaurus mientras otros mantienen a Tarbosaurus como género separado.

tiranosaurio.txt_15: Varios otros géneros de tiranosáuridos norteamericanos también han sido sinonimizados a Tyrannosaurus.