

Description of the keys and values of each JSON object in `pol_062016-112019_labeled.ndjson` file

Key	Value Type	Appears	Description	Possible values
<code>no</code>	integer	always	The numeric post ID	any positive integer
<code>resto</code>	integer	always	For replies: this is the ID of the thread being replied to. For OP: this value is zero	0 or any positive integer
<code>sticky</code>	integer	OP only, if thread is currently stickied	If the thread is being pinned to the top of the page	1 or not set
<code>closed</code>	integer	OP only, if thread is currently closed	If the thread is closed to replies	1 or not set
<code>now</code>	string	always	MM/DD/YY(Day)HH:MM (:SS on some boards), EST/EDT timezone	string
<code>time</code>	integer	always	UNIX timestamp the post was created	UNIX timestamp
<code>name</code>	string	always	Name user posted with. Defaults to Anonymous	any string
<code>trip</code>	string	if post has tripcode	The user's tripcode, in format: !tripcode OR !!securetripcode	any string
<code>id</code>	string	if post has ID	The poster's ID	any 8 characters
<code>capcode</code>	string	if post has capcode	The capcode identifier for a post	Not set, mod, admin, admin_highlight, manager, developer, founder

Key	Value Type	Appears	Description	Possible values
country	string	if country flags are enabled	Poster's ISO 3166-1 alpha-2 country code	2 character string Or XX if unknown
country_name	string	if country flags are enabled	Poster's country name	Name of any country
sub	string	OP only, if subject was included	OP Subject text	any string
com	string	if comment was included	Comment (HTML escaped)	any HTML escaped string
tim	integer	always if post has attachment	Unix timestamp + microtime that an image was uploaded	integer
filename	string	always if post has attachment	Filename as it appeared on the poster's device	any string
ext	string	always if post has attachment	Filetype	.jpg, .png, .gif, .pdf, .swf, .webm
fsize	integer	always if post has attachment	Size of uploaded file in bytes	any positive integer
md5	string	always if post has attachment	24 character, packed base64 MD5 hash of file	
w	integer	always if post has attachment	Image width dimension	any positive integer
h	integer	always if post has attachment	Image height dimension	any positive integer
tn_w	integer	always if post has attachment	Thumbnail image width dimension	any positive integer
tn_h	integer	always if post has attachment	Thumbnail image height dimension	any positive integer

Key	Value Type	Appears	Description	Possible values
filedeleted	integer	if post had attachment and attachment is deleted	If the file was deleted from the post	1 or not set
spoiler	integer	if post has attachment and attachment is spoiled	If the image was spoiled or not	1 or not set
custom_spoiler	integer	if post has attachment and attachment is spoiled	The custom spoiler ID for a spoiled image	1-10 or not set
replies	integer	OP only	Total number of replies to a thread	0 Or any positive integer
images	integer	OP only	Total number of image replies to a thread	0 Or any positive integer
bumplimit	integer	OP only, only if bump limit has been reached	If a thread has reached bumplimit, it will no longer bump	1 or 0 or not set
imagelimit	integer	OP only, only if image limit has been reached	If an image has reached image limit, no more image replies can be made	1 or not set
semantic_url	string	OP only	SEO URL slug for thread	string
since4pass	integer	if poster put 'since4pass' in the options field	Year 4chan pass bought	any 4 digit year
unique_ips	integer	OP only, only if thread has NOT been archived	Number of unique posters in a thread	any positive integer
m_img	integer	any post that has a mobile-optimized image	Mobile optimized image exists for post	1 or not set

Key	Value Type	Appears	Description	Possible values
archived	integer	OP only, if thread has been archived	Thread has reached the board's archive	1 or not set
archived_on	integer	OP only, if thread has been archived	UNIX timestamp the post was archived	UNIX timestamp
extracted_poster_id	string	If post has ID	The poster's ID	any 8 characters Or Null or not set
troll_country	string	If post has troll country	The name of the troll country selected by the poster	string
troll_country	string	If post has troll country	The name of the troll country selected by the poster	string

Description of the “perspectives” key

Key	Value Type	Appears	Description	Possible values
perspectives	dictionary	always	The perspective scores of the post	<p>7 keys:</p> <ol style="list-style-type: none"> 1. "TOXICITY" 2. "SEVERE_TOXICITY" 3. "INFLAMMATORY" 4. "PROFANITY" 5. "INSULT" 6. "OBSCENE" 7. "SPAM" <p>each has a float value (0 to 1) or empty list</p>

Description of the "entities" key

Key	Value Type	Appears	Description	Possible values
entities	List of dictionaries	always	The named entities mentioned in the post	Each dictionary holds 4 keys: 1. "entity_text" - value: string OR integer 2. "entity_start" - value: positive integer 3. "entity_end" - value: positive integer 4. "entity_label" - value: string (see https://spacy.io/api/annotation#named-entities) OR empty list