

Welcome Gitana

SQL-based Project Activity Inspector tool

רני מקלדה ניל חמזה

?Gitana אה זה

כלי פיקוח פרויקטים לפיתוח תוכנה תומך ב:

GIT • פותה בשפת PYTHON ועובד עם מסד נתונים MYSQL. BUGZILLA

יודע לשלב נתונים ממספר כלי פיתוח •

תוכנה שונים

מספק גישה מבוססת SQL לניתוח קל •

SLACK

GITHUB

ECLIPSE FORUMS

היסטוריה ופיתוח הכלי

- SOM RESEARCH LAB פותח ב-2010 על ידי
 - מטרתו לגשר בין פלטפורמות/כלים שונות.
 - מציע פתרון לאתגרים בשאילתות ושילוב נתונים.
 - NETWORK ANALYSIS-הוא תומך ב-

פיתוח והתחזוקה של הכלי

- פותח על ידי צוות של חוקרים ומהנדסים המשתפים פעולה הן עם האקדמיה והן עם התעשייה.
- בתחילה היה מתוחזק היטב עם עדכונים ושיפורים שוטפים
- נכנס לארכיון ב-09/09/2020 מסיבות שונות. עדכון אחרון Samir Savla on December 13, 2022 at 11:22 pm בכלי היה ב-2017



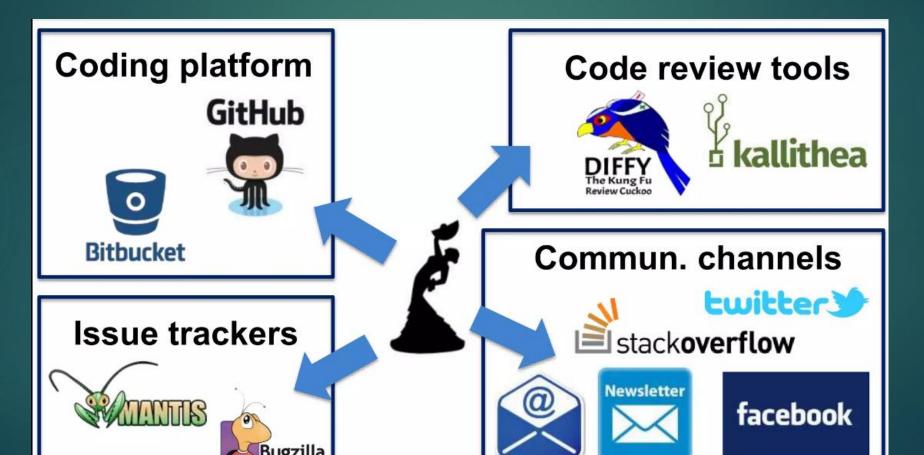
Just curious as why you decided not to continue to maintain/productize this. It would be a great tool for most enterprises looking to get empirical data from github.



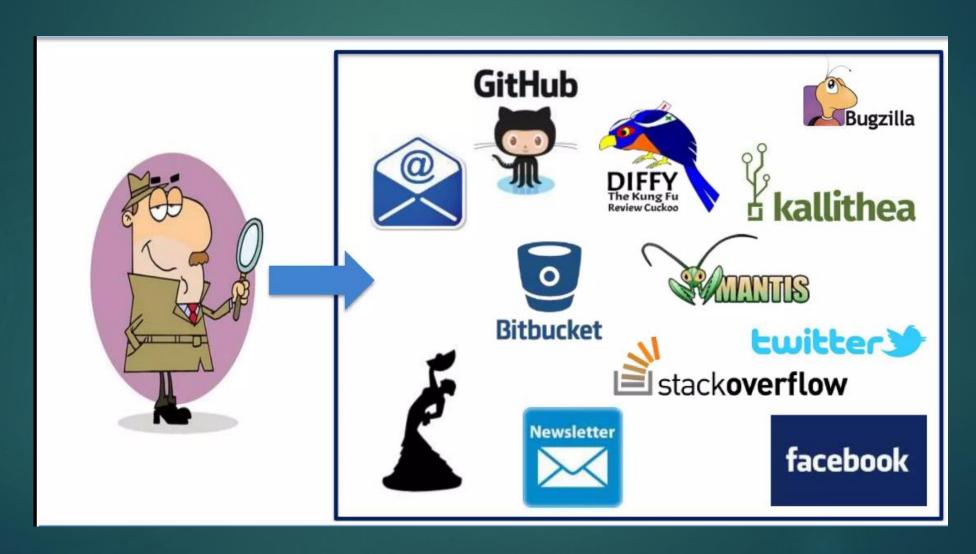
lordi Cabot on December 14, 2022 at 8:12 am

agree it is/was a promising tool but as a research team there is so much we can do it if there are no companies supporting the work. And at the time, we couldn't find anybody interested.

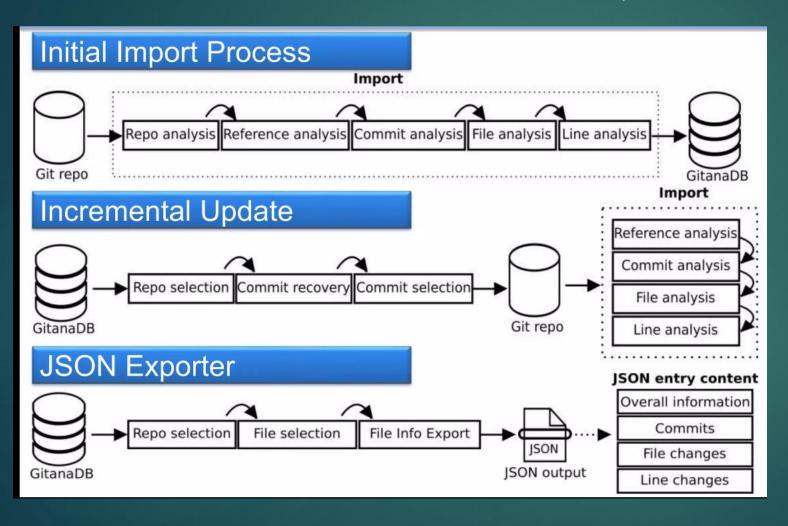
קלטים



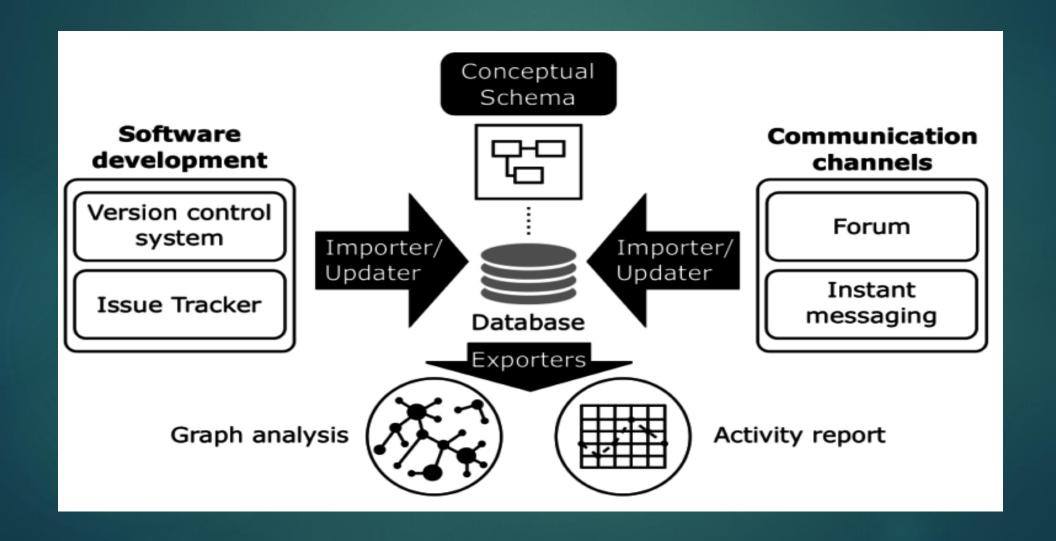
מאגר קלטים



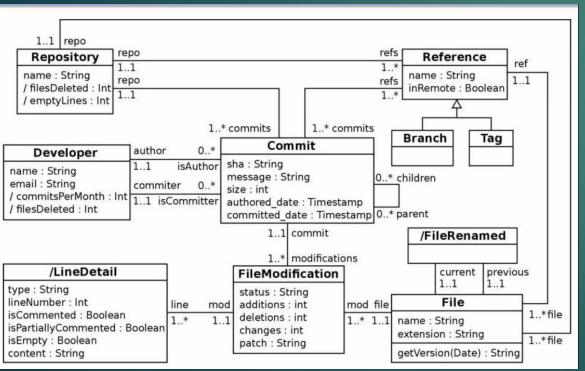
פונקציות עיקריות

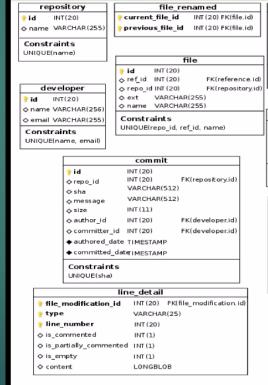


מודל נתונים



מסד נתונים







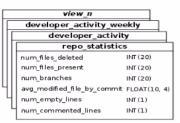
reference

commit_idINT(20)	FK(commit.i
parent_id INT(20)	FK(commit.i
	5)
♦ parent_sha VARCHAR(25	5)
Constraints UNIQUE(repo_id, commit_id)	d. parent_shal

fil	file_modification				
💡 id	INT(20)				
commit_ic	FK(commit.id)				
file_id	INT(20)	FK(file.id)			
status	VARCHAR(20)			
additions	INT(20)				
deletions	INT(20)				
changes	INT(20)				
patch	LONGBLOB				

comr	commit_in_reference			
💡 ref_id	INT (20)	FK(reference.id)		
💡 commit	idNT(20)	FK(commit.id)		

Tables



Views

:				
:	FUI	NCTION		
:	soundex_match			
:	Paramet	ers		
:	IN str_1	VARCHAR(25)		
:	IN str_2	VARCHAR(255		
:	Returns			
:	INT(1)			

FUNCTION			
levenshtein_distance			
Pέ	aramet	ters	
IN	str_1	VARCHAR(255)	
IN	str_2	VARCHAR (255)	
Re	eturns		
IN	(255)		

FUNCTION	STORED PROCEDURE
get_file_history	get_file_version
Parameters Parameters	Parameters
IN _file_id INT	IN _file_id INT
Returns	IN _date TEXT
TEXT	IN by_line BOOL

STORED PROCEDURE	STORED PROCEDURE
get_file_version_compact	get_file_version_by_line
Parameters	Parameters
IN _file_ids TEXT	IN _file_ids TEXT
IN _date TEXT	IN _date TEXT

Functions & Stored Procedures

פלט לדוגמה

Branch	Project Name	Extraction Process			Update Process		
Interval	username/project	NB	NC	ET	Δ NB	Δ NC	\mathbf{ET}
	sahat/hackathon-starter	7	2,027	511.3	0	+39	56.6
1 - 9	robbyrussell/oh-my-zsh	1	4,170	1,482.5	0	0	30.7
	neovim/neovim	8	6,562	1,554.4	0	+402	225.4
	select2/select2	36	2,631	266.4	0	0	32.6
10 - 49	google/material-design-lite	47	3,058	1,050.2	0	+92	112.5
	tensorflow/tensorflow	23	8,594	3,587.9	0	+3,216	4,987.1
	nylas/N1	63	4,195	2,272.9	0	+396	248.0
50 - 99	driftyco/ionic	93	9,797	2,178.8	0	+546	291.2
	Microsoft/vscode	77	10,446	4,761.3	0	+2,828	3,512.2
> 99	antirez/redis	214	10,372	5,446.2	0	+119	218.3
	electron/electron	256	11,191	7,056.8	0	+999	536.3
	zurb/foundation-sites	153	12,705	8,541.6	0	+738	443.9

NB = Number of Branches NC = Number of commits ET = Extraction time (secs)

הדגמת תרחיש שיפוסי לשימוש בכלי

סרטון בקישור הבא:

video of running gitana

סיכום

חסרונות	יתרונות
ייבוא נתונים איטי	git-מאכסן את כל המידע ב
עדכונים פוגעים בתצוגה במסד נתונים	גישה למידע SQL שאילות
קבוע json פורמט	Json קבצי
	חינם

kahoot