

## DB structure

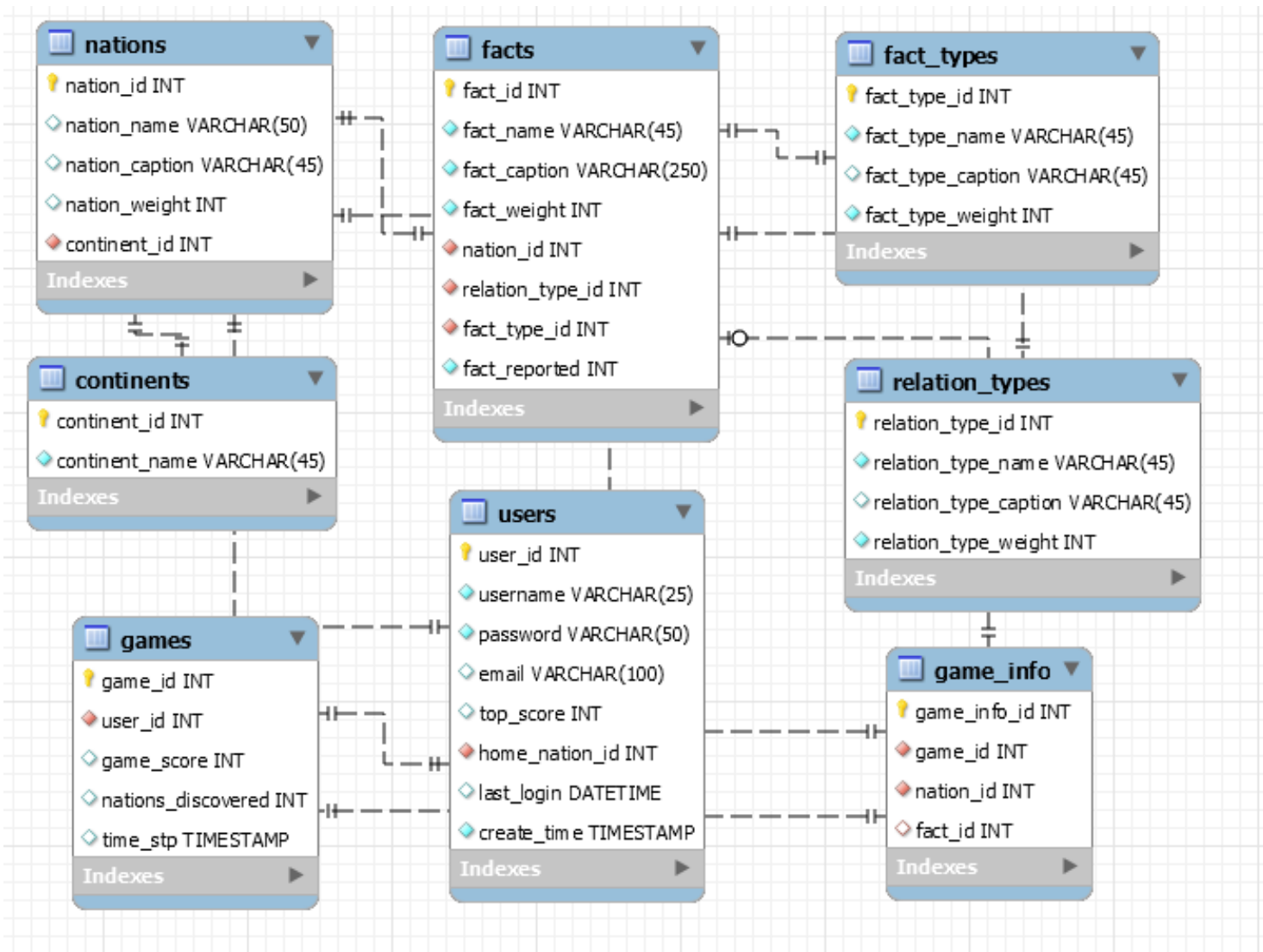
בסיס הנתונים שלנו מכיל 8 טבלאות.

תהליך תכנון המבנה כלל מחשבה על סוג השאילתות שישלחו אל בסיס הנתונים, מניעת עודפי מידע מיותרים, תמיכה מקסימאלית ומהירה לאפליקציה. בכל זאת כמו בכל פרוייקט טוב עברנו מספר גרסאות, שינינו, שיפרנו, הוספנו פיצ'רים, לכן דאגנו לתאימות לאחור ואמינות בסיס הנתונים, אך לא במחיר של הקרבת מהירות או עודפי מידע.

בחרנו ליצור לכל טבלה ID יחודי, מתקדם-אוטומטית (AI), אשר יהווה מפתח לטבלה.

שמים לב כי בחלק מהטבלאות ישנו שדה name וגם שדה caption. בכדי לעשות ניתוח מעמיק yago יש צורך להחזיק את השם כפי שמופיע ביאגו, לפני העיבוד וקיצוץ התווים המיותרים (מכיוון שלא הפיך).

### EER diagram



## בקצרה על הטבלאות:

nations - הטבלה מכילה את שמות המדינות השונות, משקל של כל מדינה (רמת קושי) ומידע באיזו יבשת המדינה.

nation_id	nation_name	nation_caption	nation_weight	continent_id
52	<Swaziland>	Swaziland	10	2
53	<Malawi>	Malawi	13	2
54	<Namibia>	Namibia	19	2
55	<Rwanda>	Rwanda	13	2

continents - טבלה קטנה המכילה את היבשות המוכרות בעולם (ועוד כמה שקולומבוס עוד לא גילה).

continent_id	continent_name
1	El mundo
2	Africa
3	Asia

facts - טבלה זו מכילה את שם הנושא לעובדה - שם דמות, שם איזור, שם הקרב וכו'. לדוגמא: Robert Brom, Kiryat Malakhi, Sky Blue Pink. בנוסף מכילה הטבלה גם את משקל העובדה ואם דווח על עובדה זו. כמובן גם מוחזק id של סוג העובדה וסוג הקשר.

fact_id	fact_name	fact_caption	fact_weight	nation_id	relation_type_id	fact_type_id	fact_reported
1	<Robert_Brom>	Robert Brom	439	159	1	1	0
2	<CJ_CheilJedang>	CJ CheilJedang	6	115	2	2	0
3	<Battle_of_Ebelsberg>	Battle of Ebelsberg	12	170	3	3	0
4	<Laura_Hickman>	Laura Hickman	439	159	1	4	0

fact types - טבלה זו מפרטת את התואר המיוחס לנושא העובדה, לדוגמא: movie, bishop, region, וכו'. כמו כן מכילה טבלה זו גם משקל אופציונאלי לסוג עצמו.

fact_type_id	fact_type_name	fact_type_caption	fact_type_weight
1	<wordnet_bishop_109857200>	bishop	1
2	<wordnet_company_108058098>	company	1
3	<wordnet_conflict_100958896>	conflict	1
4	<wordnet_writer_110794014>	writer	1

relation types - טבלה זו מפרטת את הקשר בין הנושא לסוגו. היא מכילה את רוב הטקסט אשר יוצג למתמודד (אליו ישורשרו הנושא והסוג, במיקומם הנכון, כדי ליצור משפט תקין). כמו כן הטבלה מכילה גם משקל עבור הקשר, הכרחי לקבלת עובדות מעניינות.

relation_type_id	relation_type_name	relation_type_caption	relation_type_weight
1	<isCitizenOf>	A citizen of this nation: The %s	2
2	<isLocatedIn>	A place in this nation: The %s	5
3	<happenedIn>	An event of this nation: The %s	8
4	<isPoliticianOf>	A politician in this nation: The %s	6

users - טבלה זו מכילה מידע על המשתמשים (שם משתמש, סיסמה, דואל, ניקוד, חיבור אחרון, תאריך יצירה, יבשת בה גר השחקן)

user_id	username	password	email	top_score	home_nation_id	last_login	create_time
1	alon	d8578edf8458ce06fbc5bb76a58c5ca4	alon/k@l.com	500	5	2015-06-18 16:27:46	2015-06-09 19:22:28
2	ogoun	30242c80ce0c971bc3a44d0cdff14d66	ogoun.d@gmail.com	1474	191	2015-06-18 17:03:32	2015-06-12 16:31:43
4	alonalon	d4b9967b6c3f9cdb6210cb408e82a5be	alon.elmaliyah@gmail.com	0	43	2015-06-16 11:59:00	2015-06-16 11:59:02
5	the dude	a4589a60ea90b98f8f75780b4c829e9a	vuvalzuaretz@gmail.com	698	191	2015-06-18 16:35:09	2015-06-17 18:52:08

games - מכילה מידע על המשחקים ששחקו (בכללותם, מספר סיבובים). מתמודד, ניקוד, כמה מדינות התגלו, ומתי שחק המשחק.

game_id	user_id	game_score	nations_discovered	time_stp
534	9	0	0	2015-06-18 17:14:33
533	2	1474	17	2015-06-18 17:12:56
532	2	100	1	2015-06-18 17:07:22
531	2	348	4	2015-06-18 17:05:56

game\_info - מכילה מידע על משחקים, כגון איזה מדינה שוחקה בכל סיבוב, ואיזה עובדות ראה המתמודד בכל סיבוב.

game_info_id	game_id	nation_id	fact_id
16	2	37	65666
19	2	37	281231
21	2	60	105620

### באיזה מידע מיוצג YAGO השתמשנו:

כל המידע שנמצא במסד הנתונים מבוסס על מידע מיאגו.

#### 1. מדינה מוגדרת להיות עובדה מהצורה (מתוך **yagoTypes.tsv**):

`<id> <nation_entity> rdf:type <wikicat_Countries>`

כך נאספה רשימת מדינות וניתן לגשת לאיסוף עובדות על מדינות

#### 2. עובדה על מדינה מוגדרת להיות עובדה מיאגו בה מדינה היא הנושא או הנשוא:

מתוך **yagoFacts.tsv**:

`<id> <nation_entity> <relation> <fact_name>`

`<id> <fact_name> <relation> <nation_entity>`

מתוך **yagoDateFacts.tsv & yagoLiteralFacts.tsv**:

`<id> <nation_entity> <relation> "Literal"`

#### 3. וסוג העובדה מוגדר להיות הסוג של `<fact_name>` מתוך **yagoTypes.tsv**:

`<id> <fact_name> rdf:type <fact_type>`

אלרגוריתם איסוף העובדות עובר תחילה על **yagoTypes.tsv** כדי לאסוף רשימה של כל המדינות, לאחר מכן עובר על **yago\*Facts.tsv** בכדי לאסוף את כל העובדות הקשורות למדינות (ובמקביל לנהל רישום של נפוצות לכל עובדה) ולבסוף עובר שוב על **yagoTypes.tsv** בכדי להכריע את הסוג של כל עובדה.

בסופו של התהליך המידע מוזן למסד הנתונים כאשר שמות העובדות והמדינות עוברים תהליך לכדי טקסט קריא.

### באילו חבילות חיצוניות השתמשנו:

SWT 3.6.1 – לבניית ממשק המשתמש.

JDBC-5.1.35 – לתקשר עם מסד הנתונים.

java-mail-1.4.4 – לשלוח מייל למשתמש חדש.