IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Revue d'Intelligence Artificielle
Vol. 34, No. 6, December, 2020, pp. 721-729
Journal homepage: http://iieta.org/journals/ria

AQG: Arabic Question Generator

Kheira Z. Bousmaha[1*], Nour H. Chergui[2], Mahfoud Sid Ali Mbarek[3], Lamia Belguith Hadrich[4]

[1] Department of Computer Science, FSEA, RIIR Laboratory, University Oran1, Oran 31005, Algeria
[2] Department of Computer Science, FSEA, University Oran1, Oran 31005, Algeria
[3] Department of Electronic, LSI Laboratory, University of Science and Technology Oran – Mohamed Boudiaf (USTO), Oran 31001, Algeria
[4] Department of Computer Science, Faculty of Economics and Management (FSEGS), University of Sfax, Sfax 3018, Tunisia

Corresponding Author Email: kzbousmaha@univ-oran1.dz

https://doi.org/10.18280/ria.340606

Received: 3 August 2020
Accepted: 16 December 2020

Keywords:
Arabic natural language process, question generation, semantic role labelling, semantic methods, model-based methods

ABSTRACT

The Arabic natural language process (ANLP) community does not have an automatic generator of questions for texts in the Arabic language. Our objective is to provide it one. This paper presents a novel automatic question generation approach that generates questions as a form of support for children learning through the platform QUIZZITO. Our approach combines the semantic role labelling of PropBank (SRL) and the flexibility of question models. It essentially relates to an approach of instantiation model of representation based on an analysis focused on the semantics. This allowed us to capture the maximum sense of sentence given the flexibility of the grammar of the Arabic language. This model was written in a set of Patterns and Templates based on the REGEX languages. Our goal is to enrich Quizzito's online quiz platform, which contains more than 254.5k quizzes, and to provide it with a generator of Arabic language questions for children's texts. Our Arabic Question Generator system (AQG) is functional and reaches up to 86% f-measure.

1. INTRODUCTION

There are many uses of an automatic question generator (QG): the creation of multiple-choice tests and quizzes for learning materials, human-machine dialogue systems, or interactive Question-Answer.

While this topic has been widely debated in other languages, for the Arabic language, it has never been discussed. Yet, Arabic is a language spoken by more than 300 million people in more than 22 countries.

The ANLP is considered difficult to apprehend because of its singular characteristics. His morphology has always been a challenge for the NLP specialists. Thus, its automatic processing must cope with its agglutinative nature, its inflectional richness and its diacritics bearing meaning and certain morphosyntactic traits such as declination, mode, and case; which leads to a large number of ambiguities lexical, morpho-syntactic and semantic thus curbing its treatment. In fact, the adaptation of existing models that have shown their effectiveness in treating other languages to the Arabic language is problematic [1].

Our contribution is about this theme. It focuses on an approach of instantiation of the representation model based on an analysis driven by the semantics of the sentences of a text. It allows you to automatically switch from text to generated questions. Building a question manually requires reading and understanding the text.

We have chosen to define this model using Semantic Role Labelling, i.e. the PropBank SRL [2]. SRL are based on intuitive reasoning, close to human reasoning. In addition, they provide a framework of representation and formal semantics that can guide the identification of the components of the sentence. Semantic role tagger attempts to identify the predicates (relationships and actions) and the semantic entities associated with each of these predicates. The set of semantic roles used in PropBank include both predicate-specific roles whose precise meaning is determined by their predicate (Rel) and multipurpose adjunct-type modifier roles whose meaning is consistent for all predicates(C) [2]. Our task is the identification of the components of the sentence such as: المبتدأ (mubtada, beginning), الخبر (elkhabar, predicate), الجار و المجرور (jar wamajrur, prepositional phrase), الفعل (fiel, verb), الفاعل(faeil, subject) to facilitate the generation of questions.

The SRL can be used to identify the appropriate question words in the absence of named entities, but the replacement of question words is not the biggest challenge to solve in QG from text. The challenge is rather the understanding of the meaning of the sentence.

We chose the semantic method based on models to capture the maximum meaning of the sentence, regardless of the syntactic order of its words. Knowing that during the syntactic annotation, it must be taken into account that the Arabic language is a pro-drop zero-subject language and that it systematically omits the morphological realization of the subject pronoun; the verb contains a pronominal index subject in its flexion. It is also important to know that the Arabic language has a mixed order of words, that is to say, neither completely free nor completely fixed. Concerning the verbal sentence, the order of the standard Arabic sentence usually follows the VSO order (that is, with a verb (V) followed by its subject (S) and object (O)). Nevertheless, one can find the

existence of other structures including the order SVO (subject-verb-object). The VOS (verb-object-subject) structure also exists, but with a lower frequency. It serves exclusively to emphasize the subject. Finally, the rare OVS (object-verb-subject) structure is sometimes used to express focus on the subject [1]. The complexity and flexibility of the grammar of the Arabic language poses a real problem in identifying the different components of the sentence.

This paper presents a new approach to generating questions from an Arabic language text combining the semantics of SRL and the flexibility of question models. We wrote a set of patterns and templates based on the REGEX language [3]. Our evaluation methodology was focused on evaluating issues from an educational perspective.

This article is composed of 4 sections. Section 2 presents a literature review on question generation methods and deals with related work. Section 3 details our contribution. Section 4 and 5 make an assessment of AQG. Finally, a conclusion of this work and some perspectives are given in section 6.

2. LITERATURE REVIEW

Efforts in QG from text can be regroup into three categories: syntax-based, semantics-based, and model-based. These three categories are not totally disjoint. Whichever approach is chosen, systems must perform at least four tasks:

(1) Content Selection: selection of source text ranges (usually single sentences) from which questions can be generated.

(2) Target identification: determines which specific words and / or phrases should be queried.

(3) Formulation of the question: determines the appropriate question (s) based on the identified content.

(4) Surface shape generation: final production of the surface shape.

These tasks are not always discrete and do not necessarily occur in that order. Identifying the target can guide the formulation of the question and vice versa.

Papasalouros and Chatzigiannakou [4] provides an overview of the existing research of the subject of applying Semantic Web technologies for automatic question generation. The review provides a classification based on technological as well as on pedagogical aspects of the works presented. According to recent research [5, 6], the majority of QG systems focus on generating questions for the purpose of assessment. The template-based approach was the most common method employed in the reviewed literature. The focus was on the generation of questions from text and for the language domain. Duan et al. [7] explored how to generate questions from given passages using neural networks. They explore two ways to generate questions for a given passage: one is a method based on the recovery using a convolutional neuron network (CNN), the other is a method based on the generation using a network of recurrent neurons (RNN). De Viron et al. [8] presented a system of automatic generation of questions for French. The generation system proceeds by transforming declarative sentences into interrogatives and is based on a prior syntactical analysis of the basic sentence. Ciguene et al. [9] focus more specifically on the automatic generation of test subjects, to assist the teacher in the development of his subjects. Their approach consists of an iterative construction of subjects "on the fly", starting from a source exercise database, built beforehand by the teacher. This work presents the design of an automatic generator based on multiple choice questions exercises called DIFAIRT-G (Different FAIR Tests Generator). Yao and Zhang [10] presented a QG approach based on minimal semantic recursion semantics (MRS), a superficial semantic analysis framework. Their method uses an eight-step pipeline to convert the input text into a set of questions. Key elements of their approach are MRS decompositions and MRS transformations. Decompositions convert complex sentences into simple sentences from which MRS transformations can generate questions.

All these researches have shown great efficiency. The question that arises is how to adapt existing methods to the Arabic language, knowing that no research work has been devoted to it in the field of automatic question generation.

3. DESCRIPTION OF AQG

The design steps for our question generation system are illustrated in Figure 1. It begins with the analysis of a text in Arabic using the online morphological analyzer MADAMIRA [11]. MADAMIRA performs segmentation into words, not sentences. To make up for that, we integrated the STAR (Arabic Texts Slicer) tool [12] as a module for the segmentation of the text into sentences. The output will be a segmented file and annotated in XML.

Then, AQG proceeds to the generation and the assignment of roles to the different actors of the sentence. We used Propbank SRLs that we needed, and we adapted them to the Arabic language, a task that was not easy given the complexity of the language. Once this phase was completed, the Pattern design modules and Template modules were designed to generate the questions. The method used is based on the semantic model method. These models, written in the form of regular expressions, required a very profound knowledge of the Arabic language grammatical rules.

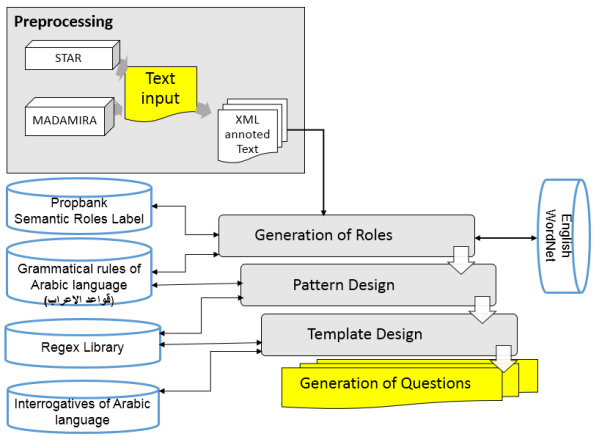


Figure 1. Architecture of AQG

3.1 Preprocessing

For the preprocessing and annotation of our texts, we have chosen MADAMIRA, a system for morphological and orthographic disambiguation of Arabic. It produces a rich feature set for each word, containing more than 14 morphological and lexical features. It includes the full diacritization of the word, identification of its stem and morphemes, part-of-speech, identification of the lemma of the

word and its English gloss as well [13]. It also provides different tokenization schemes. MADAMIRA has an implementation that is more robust, portable, and extensible and is faster than its ancestors by more than an order of magnitude [11].

The performance of MADAMIRA is quite competitive, scoring 96% accuracy for lemma and stem as over 99% accuracy for word segmentation. Diacritization is also high (96%) when excluding the case markers. MADAMIRA's speed is close to 1000 words per second in a server-client mode [13]. MADAMIRA also provides XML and HTTP support: input and output text can be supplied as plain text or in XML [11].

3.2 Generation of roles

Semantic roles describe how words are used in sentences and the functions they fulfill. For their adaptation, we were inspired by the work [14-18]. Table 1 summarizes the set of roles we utilized and the grammatical category or categories that we assigned to each of these roles.

Table 1. Adaptation of semantic roles to the Arabic language

ARG	Description	Grammatical category assigned
ARG0/A0	agent	(faeil,subjectالفاعل) (mubtada,beginning) المبتدأ
ARG1/A1	patient	(mafeul-bih,object)المفعول به (sifa,adjective) الصفة (elkhabar,predicate)الخبر
AM_LOC	locative	ظرف المكان (zarf almakan, adverb of place)
AM_TMP	temporal	ظرف الزمان (zarf alzaman, temporal adverb)
AM_PRP	purpose	جار ومجرور (jar wamajrur,prepositional phrase)
Predicate	verb	الفعل (الماضي / المضارع / الأمر/ المبني للمجهول) (Verb (past/present/imperative / passive voice))

Writing SRL must start by recognizing their grammatical categories. It is therefore important to recognize the type of sentence. In the nominal sentence, we detect its subject (\'/ خبر'\'(mubtada (ARG: A0)) and its predicate/'مبتدأ\ elkhabar (ARG: A1)) and in the verbal sentence, we identify مفعول به' \'\(faeil (ARG:A0)) and its object/ فاعل' \'\(its subject .mafeul-bih (ARG: A1)) and this in all their possible forms/

For the other argument that the ARG= {AM_LOC, AM_TMP, AM_PRP} we need to identify the locative, temporal, and the purpose expression in the sentence, respectively. We have written a set of patterns in the form of regular expressions based on the research work of linguists [19-23]. We used the syntax of Regex from Python NLTK [24].

3.2.1 SRL design of AM_TMP

This role corresponds to the presence of temporal adverb (ظرف الزمان (zarf zamen)). We tried to enumerate them in a list so that we could ask questions of temporal types. We remark that their meaning is different; it will be a duration or a precis time. Like illustrated in Table 2, we designed two sub-roles AM_TMP_K to ask a question about duration with the interrogative كم (how many) and AM_TMP_M to ask a question about a precis time with the interrogative متى (when). (_K) means كم(how long, how much time) and (_M) means متى(when).

Table 2. AM_TMP_M and AM_TMP_K

AM_TMP_M (corresponding to the question when)	AM_TMP_K (corresponding to the question How long/ How much time)
الآن , قبل , بعد , عند , لدى , لدن, بينما , عندما , قط , إذا , أيان , جانفي , فيفري , مارس , أفريل , ماي , جوان , جويلية , أوت , سبتمبر ,أكتوبر ,نوفمبر . ديسمبر ,السبت , الأحد , الاثنين , الثلاثاء , الأربعاء , الخميس , الجمعة , محرم , صفر , ربيع الأول , ربيع الثاني , جمادى الأولى , جمادى الثانية , شعبان , رمضان , شوال , ذو القعدة , ذو الحجة , مساءا , صباحا , ليلا , فجرا , ظهرا , عصرا , صيفا , شتاءا , خريف , ربيع , مغرب , أمس , البارحة , غدا قديما , كانون , زوالا, حديثا , الأسبوع الشهر, العام , اليوم, الثاني ,شباط , آذار نيسان , أيار , حزيران , تموز, آب , أيلول ,تشرين الأول , تشرين الثاني,كانون الأول.	لحظة، برهة، مدة، دهرا، وقت، زمن، أسبوعا ,شهرا , عاما , يوم , ثانية , ساعة , دقيقة ,حولا ,قرنا, فترة ,هنيهة.

Sometimes the same word can be interpreted differently. For example, in the sentence (1) the word "**أسبوعا**" (week) that indicates here a duration, is an AM_TMP_K. While in the sentence (2), the same word "**الأسبوع**" (the week) must be recognized as AM_TMP_M because it signifies a moment although the adverb is listed in AM_TMP_K.

(1) (I stayed in the Levant for a week) بقيت في الشام أسبوعا
AM_TMP_K

(2) (I visited Baghdad last week) زرت بغداد الأسبوع الماضي
AM_TMP_M

In the Arabic language, the problem is solved by introducing the morphological characteristics of the word into the sentence. Therefore, we include the part of speech in the SRL design of AM_TMP.

3.2.2 SRL design of AM_LOC

This role corresponds to the presence of adverb of place (ظرف المكان (zarf almakan)) in the sentence. We recorded the adverbs places in a list like { شرق , غرب , شمال , جنوب, خلف , تحت , ذراع , ميل , مكان , ناحية , جانب , ذات , فوق , وراء, قدام يمين , يسار أين , هنا , ثم , حيث , شطر , حول , قبل , بعد ,لدى , لدن , حوالي , نحو , أسفل , بين, أعلى } so that we could ask questions of types: أين (where).

Example:
(Maldives is located north of the western coast of the Republic of Sri Lank)
تقع جزر الملديف **شمال** الساحل الغربي لجمهورية سيريلانكا.
AM_LOC

At the meantime, we have distinguished a list grouping temporal and place adverbs (ظرف (zarf)) common to both. Common temporal and place adverbs = { عند, بين, مع, قبل, بعد ,ذات, لدى}. Their processing will be different because it will depend on the word categorization that follows it. This categorization can be organization, locative, Date (Temporal), Person or Numeric expression. It is extracted from English WordNet from the "gloss" tag of MADAMIRA, which gives the translation of these words into English. Therefore, we include this categorization in the SRL design of AM_LOC and AM_TMP. The following examples illustrate the categorization of some of these adverbs.

Temporal	يتمرن الرياضي **بين العصر والمغرب.**	(The athlete exercises between the al-asr and the Maghrib)
Locative	تقع الجزائر **بين الشرق والغرب.**	(Algeria is located between east and west.)
Locative	يطوف الحجاج **بين الصفا والمروة.**	(Pilgrims roam between Safa and Marwah.)
Temporal	وقت حفظ الطلاب في المسجد **بين 15سا و17 سا**.	(The time for students to be kept in the mosque is between 15 hrs and 17 hrs)

3.2.3 SRL design of AM_ PRP

In Arabic, there are nearly 20 prepositions; only 8 of them are the most used, each its function. We have decomposed the AM_PRP SRL into subtypes listed in Table 3 in order to find the corresponding interrogative in writing patterns.

Table 3. Subtypes of AM_PRP

AM-PRP	Example
AM_PRP_M (من/From)	Layla's book is a gift from Master كتاب ليلى هدية **من المعلمة** AM_PP_M
AM_PRP_I (إلى/to)	Back to the country a great joy الرجوع **الى البلد** فيه فرحة كبيرة AM_PRP_I
AM_PRP_N (عن/about)	Deep talk about studying abroad الحديث معمق **عن الدراسة** في الخارج AM_PRP_N
AM_PRP_AL (على/on)	The phone is on the chair in the room الهاتف **على الكرسي** في الغرفة AM_PRP_AL
AM_PRP_F (في/in)	The passengers in the plane as the King كل مسافر **في الطائرة** كالملك AM_PRP_F
AM_PRP_B (الباء/in,with)	Oranges are rich in vitamins البرتقال غني **بالفيتامينات** AM_PRP_B
AM_PRP_K (الكاف/as)	The girl's face is as beautiful as the moon وجه الفتاة جميل **كالقمر** AM_PRP_K
AM_PRP_L (اللام/for)	Fund for poor money مال الصندوق **للفقراء** AM_PRP_L

Example of a regular expression detected AM_PRP in the AQG program.

(1) {<IN\+DT\+NN?>}

(2) {<IN><DT\+NN?><DT\+NNS>?}

(3) {<IN><DT\+JJ><DT\+NNS>?}

(4) {<IN><NN?>(<NN?>|<NN\+PRP\$>|<JJ>|<DT\+NN>)?}

For example, the SLR AM_PRP: {<IN><NN\+PRP\$>} signifies preposition followed by a name. In the sentence " المعلمون المخلصون **من صفاتهم** الصدق **في عملهم**" (Teachers loyal to their honesty in their work), and according to this SLR, the AM_PRP (underlined) are recognized. The NNP, DT, NN, JJ, IN, NNP, etc. are the Part of speech of MADAMIRA [10].

3.2.4 SRL design of A0 and A1 in nominal sentence

We have tried to write a set of regular expressions covering all possible cases for the nominal sentence (NS) [21]. Each SLR is denoted by an index ranging from 1 to 9. The Part of speech of MADAMIRA (NNP, DT, NN, JJ, IN, NNP,….) are the morphological details that we need for designing the A0 and the A1 in order to remove ambiguities as we have shown in paragraph 3.2.1. The detection of the previous SLRs allowed us to reduce considerably the writing of the recognition SLRs of A0 and A1 while generating the maximum of possible cases in nominal sentence. We give in the following some examples on SRL which determines the A0 and the A1 in a NS.

Examples of SRL design A0

(1) {^<NNP>+(<DT\+NN>|<DT\+JJ>)*<AM_PRP>*}: the sentence must begin with a proper name (one or more) followed by a name or an adjective (zero or more) followed by a prepositional sentence (zero or more).

(2) {^<DT\+NN>+<DT+JJ>*<AM_PRP>*}

(3) {^<AM_LOC>+<DT+JJ>*<AM_PRP>*}: the sentence must begin with an adverb of place (one or more) followed by an adjective (zero or more) followed by a prepositional sentence (zero or more).

(4) {^<AM_TMP>+<DT+JJ>*<AM_PRP>*}

Examples of SRL design A1

(1) {<NN>+(<DT\+NN>|<DT\+JJ>|<AM_LOC|AM_TMP\+(PRP\$)?>|<NNP>)*<AM_PRP>*}

(2) {<JJ>+(<DT\+NN>|<DT\+JJ>|<AM_LOC|AM_TMP\+(PRP\$)?>|<NNP>)*<AM_PRP>*}

(3) {<AM_PRP>+}

(4) {<NN\+PRP\$>+(<DT\+NN>|<DT\+JJ>|<AM_LOC|AM_TMP\+(PRP\$)?>|<NNP>)*<AM_PRP>*}

The NS sentence"اللهو عن الدراسة أمر سيء" (Having fun from school is bad) is recognized by AQG as illustrated in Figure 2.

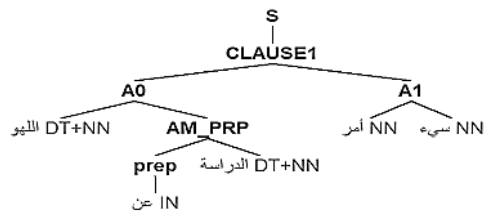


Figure 2. Decomposition of the sentence
اللهو " عن الدراسة أمر سيء "

3.2.5 SRL Design of A0 and A1 in verbal sentence

We have designed a set of regular expressions covering all possible cases for the verbal sentence (VS) [21]. We have adopted the same process to those of the nominal sentence. We give some examples.

Example of SRLs that determines the A0 and the A1 in VS.

Examples of SRL$_i$_ A0

(1) {<WP>(<CC>(<NN\+(PRP\$)?>|<DT?\+NN>|<DT?\+JJ>|<NNP>|<AM_PRP>)+)*}

(2) {<VB.?(\+PRP)?>(<CC>(<NN\+(PRP\$)?>|<DT?\+NN>|<DT?\+JJ>|<NNP>|<AM_PRP>)+)*}

(3) {<DT\+NN?><DT\+JJ>?(<CC>(<NN\+(PRP\$)?>|<DT?\+NN>|<DT?\+JJ>|<NNP>|<AM_PRP>)+)*}

(4) {<NN(\+PRP\$)?><DT\+NN>?(<CC>(<NN\+(PRP\$)?>|<DT?\+NN>|<DT?\+JJ>|<NNP>|<AM_PRP>)+)*}

For example the SLR_8 A0: {<DT?\+NN>(<DT\+NN>*|<num_noun>|<num_adj>) (<,|و> (<NN\+(PRP\$)?> | <DT?\+ NN> |<DT ?\+ JJ>| <AM_LOC|AM_TMP \+(PRP\$) ?> | <NNP> | <digit> |< noun_num> | <adj_num>|<AM_PRP>)*)* } detects A0 in the sentence " تهدف **برامج النظافة** إلى جعل بيئتنا نظيفة" (Hygiene programs aim to make our environment clean) where A0 is underlined.

Examples of SRL_i_A1

(1) {<VB.?><.*>*(<CC>(<NN\+(PRP\$)?>|<DT?\+NN>|<DT?\+JJ>|<NNP>|<AM_PRP>)+)*}
(2) {<VB.?\+PRP><.*>*(<CC>(<NN\+(PRP\$)?>|<DT?\+NN>|<DT?\+JJ>|<NNP>|<AM_PRP>)+)*}
(3) {<NN><DT\+NN><DT\+JJ>?(<CC>(<NN\+(PRP\$)?>|<DT?\+NN>|<DT?\+JJ>|<NNP>|<AM_PRP>)+)*}

For example the SLR_1_A1 :{ < DT ?+\NNP >< DT?\+NN> ?} detects A1 in the sentence "و كلم الله **موسى تكليما**" (and to Moses Allah spoke directly) where A1 is underlined.

The VS sentence"غزا العالم الفضاء في القرن العشرين"(The world invaded space in the twentieth century) is recognized by AQG as illustrated in Figure 3.

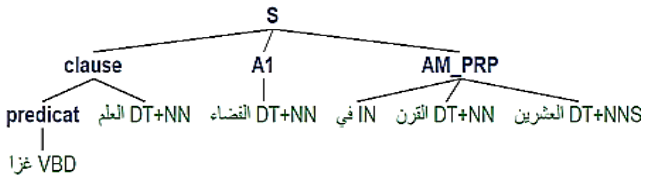


Figure 3. Decomposition of the sentence "غزا العالم الفضاء في القرن العشرين "

3.3 Pattern design

Once the SRLs of each sentence are recognized, we move on to the design of the patterns. Patterns help us recognize the structure and meaning of the sentence. We will cite the different cases of patterns in NS and VS.

The nominal sentence can be written in the form A0 + A1 or A1 + A0. Both possible cases cited in the literature have been taken into account by our system [22].

Patterns for the form A0 + A1 are all the combinaisons:

- SRL_i _A0 < SRL_7_A1 for i ranging from 1 to 4
- SRL_5_A0 < SRL_j_A1 for j ranging from 8 to 9

Patterns for the form A1 + A0 are all the combinaisons:

- SRL_i_A_1 < SRL_6_A0 for i ranging from 1 to 6

These two examples show the application of the patterns for the recognition of A0 and A1 in a NS placed in brackets in the example.

(1) [الله لطيف بعباده] يرزق من يشاء.
A1 A0
(Allah is Subtle with His servants; He gives provisions to whom He wills)

(2) وأطيعوا الله وأطيعوا الرسول فإن توليتم [فإنما على رسولنا البلاغ المبين].
A 1 A0
(So obey Allah, and obey His Messenger but if ye turn back, the duty of Our Messenger is but to proclaim (the Message) clearly and openly)

According to the grammar of the Arabic language, the verbal sentence can have four possible forms. The four forms were taken into account by our AQG system. We have summarized them in a set of reduced patterns that are able to recognize the A0 and A1 (underlined in the examples below) of a verbal sentence whatever its syntactic structure [23].

(1) (predicate < A0 < A1)
Example: تهدف برامج النظافة إلى جعل بيئتنا نظيفة (Cleaning programs aim to make our environment clean).
(2) (predicate < A0)
Example: انتصر المسلمون على عدوهم في غزوة بدر (Muslims defeated their enemy in the Battle of Badr).
(3) (A0 < predicate < A1)
Example: الأم أرضعت طفلها (The mother breastfed her child).
(4) (A1 < predicate < A0)
Example:إياك نعبد (Thee do we worship). We notice that in this case the subject (A0) is null pro-drop.

We don't have implemented the case of the passive voice Verbal sentence =" نائب فاعل + / faeal/verb + nayib faeil(فعل" Pro-agent) although we have made it design because MADAMIRA often detects the passive voice (VBN) as a noun, noun, or an adjective or a proper noun.

3.4 Template design

In Arabic, there are more ways to ask questions than in English. Arabic uses the same interrogatives WH of English in addition to several other interrogatives specific to it. For instance, interrogatives like: من(Who), لمن (Whose), ما/ماذا/مما (What), أي (Which), أين (Where), متى/أيّان (When), لماذا (Why), كم (How much, How many) and كيف (How), are some ways of asking questions in Arabic. In addition, Arabic Interrogative nouns include other particles that are related to the expected question types and/or scopes [25].

For the generation of templates, we used the interrogatives of the Arabic language. And for each recognized pattern, we wrote a set of templates. The templates were written taking into account the type of sentence as well as the two forms of the nominal sentence and the 4 forms of the verbal sentence.

If we take the sentence" محمد كالأسد في القتال" (Muhammad is like a lion in combat). When templates are recognized the questions are asked. For this sentence, two templates are triggered.

Template 1: ؟A1 (A0 – AM_PRP_K) ما IN
Template 2: ؟ A0 (A1 - AM_PRP_K) ماذا يشبه / تشبه
The questions posed are **Q1:** في ما محمد كالأسد؟ (u and **Q2:**؟ ماذا يشبه / تشبه محمد)

Table 4 illustrates some examples of templates (S: sentence and Q: question).

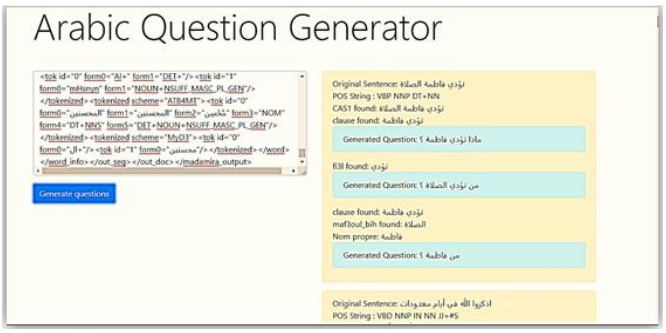
Example of templates in AQG program
Template('من $fi3l $maf3oul_bih $jar ؟')
Template('ماذا $target $jar ؟')
Template('$PP ما $target $maf3oul_bih ؟')
Template('من $nom ؟')

There are other questions that our AQG system may ask, like "اشرح الكلمات التالية" (Explain the following words, " أعط وظف الكلمة" ,"عنوانا مناسبا للنص" (Give a suitable title for the text) "التالية في جملة مفيدة" (Use the following word in a useful sentence), we choose 3 words randomly from the text.

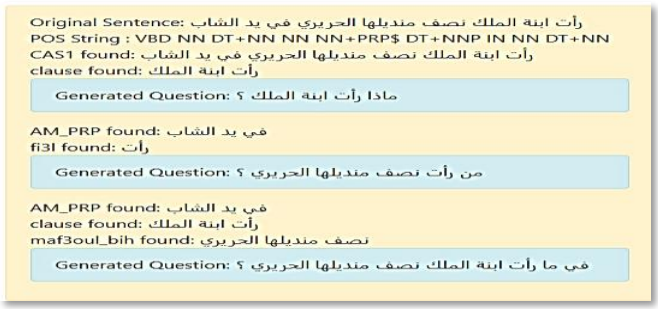
Some examples of questions generated by AQG system are shown in Figure 4.

Table 4. AQG templates

Template		Example
ماذا AM-PRP-M ؟	S	من أقوال هتلر : لا تقاتل من ليس لديه شيء يخسره
	Q	ماذا من أقوال هتلر؟
أعط مثال عن A0	S	تطوير الهوايات الجميلة كالرسم والزخرفة والموسيقى والنحت
	Q	أعط مثال عن تطوير الهوايات الجميلة.
علل لماذا A0 A1؟	S	السفر قطعة من العذاب لأن الرجل يشتغل فيه عن صيامه وصلاته وعبادته.
	Q	علل لماذا السفر قطعة من العذاب.
من الذي/التي predicate *AM_PRP A1؟	S	تسير ابنة الملك في حديقة قصرها.
	Q	من التي تسير في حديقة قصرها؟
أين predicate AM_PRP* (A1 ?) (A0 ?) ؟	S	وضعت المعلمة الكتب فوق الخزانة
	Q	أين وضعت المعلمة الكتب؟
مع كم (A1 ?) predicate(? A0) ؟	S	تحدثت مع سبعة طلاب
	Q	مع كم طالب تحدثت؟
من NNP ؟	S	كان الثعلب 'ثعلوب' يمشي قرب بيت الديك 'كوكو'
	Q	من ثعلوب؟ / من كوكو؟
كمpredicat (A0 ?)	S	بكى ثعلوب كثيرا
	Q	كم بكى ثعلوب؟
predicat (A1 ?) (A0 ?) كيف	S	يبدو منزل صديقي كالقصر
	Q	كيف يبدو منزل صديقي؟
ما (PP-AM_PRP_M) هي	S	الفروسية والسباحة والقفز من أفضل الرياضات
	Q	ما هي أفضل الرياضات
ما اللون الذي A0 predicate ؟ (AM_PRP?)	S	استعملت الفنانة اللون الأحمر في لوحتها
	Q	ما اللون الذي استعملت الفنانة في لوحتها؟



(a) AQG principal interface



(b) Questions generated

Figure 4. AQG interface

4. EXPERIMENTS

We have integrated our system into the Quizzito online game [26, 27]. Quizzito is a platform with a set of texts in the Arabic language for children aging between 6 and 14 years old (see Figure 5). In order to evaluate the understanding of these texts, questions are asked about them and the children respond on this platform. The generation of these questions is quickly becoming automatic for other languages, but regarding the Arabic language, the designers of the platform are currently soliciting Internet workers to prepare a set of questions related to each text of their corpus.

Figure 5. Quizzito interface

Quizzito is composed by 4,712 quizzes in Arabic, each quiz having either 5 or 10 multiple-choice questions, which make 40,435 questions. The questions are derived from children books and books summaries, and have all been generated by humans. Specifics instructions are given to them, such as: "read the book before writing the questions", and “use a simple language, that children can understand”.

For the evaluation of AQG, we selected a subset of ten documents (texts) consisting of 600 sentences taken from the corpus of Quizzito which counts 374 documents**.** In Figure 6, we show an example of text taken from this corpus.

يُحْكَى أَنَّهُ كَانَتْ هُنَاكَ عَائِلَةٌ بِهَا سَبْعُ شَقِيقَاتٍ. وَفِي أَحَدِ الْأَيَّامِ، خَرَجَ الْأَبُ لِجَمْعِ الْأَخْشَابِ وَوَجَدَ سَبْعَ بَيْضَاتِ بَطٍّ بَرِّيٍّ. أَخَذَهَا الْأَبُ إِلَى الْمَنْزِلِ، وَلَكِنَّهُ لَمْ يُفَكِّرْ فِي إِعْطَاءِ أَيٍّ مِنْهَا إِلَى بَنَاتِهِ؛ فَقَدْ كَانَ يُرِيدُ أَكْلَهَا مَعَ زَوْجَتِهِ. فِي الْمَسَاءِ، اسْتَيْقَظَتِ الِابْنَةُ الْكُبْرَى مِنَ النَّوْمِ وَسَأَلَتْ أُمَّهَا: مَاذَا تَطْبُخُ؟ قَالَتِ الْأُمُّ: «أَنَا أَطْبُخُ بَيْضَ الْبَطِّ الْبَرِّيِّ. سَأُعْطِيكِ وَاحِدَةً وَلَكِنْ يَجِبُ أَلَّا تُخْبِرِي أَخَوَاتِكِ بِالْأَمْرِ.» وَأَعْطَتْهَا الْأُمُّ بَيْضَةً. ثُمَّ اسْتَيْقَظَتِ الِابْنَةُ الثَّانِيَةُ وَسَأَلَتِ الْأُمَّ: مَاذَا تَطْبُخُ؟ قَالَتْ: «بَيْضَ الْبَطِّ الْبَرِّيِّ. إِذَا لَمْ تُخْبِرِي أَخَوَاتِكِ فَسَأُعْطِيكِ وَاحِدَةً.» وَأَعْطَتْهَا بَيْضَةً بِالْفِعْلِ. وَفِي النِّهَايَةِ، أَكَلَتِ الْفَتَيَاتُ الْبَيْضَ كُلَّهُ وَلَمْ يَتَبَقَّ شَيْءٌ.

Figure 6. Example of text in Quizzito

5. RESULTS AND DISCUSS

Two common types of evaluation methodologies have been put into practice [14]. The first is based on a direct human evaluation of the output of the system; the people directly involved in the research have evaluated the results of their own system. The second common methodology is based on a comparative evaluation between the output of the system (Qg) and the questions that man has generated (Qh) from the same corpus. We could consider this an indirect human evaluation. The metrics used are given in equations Eq. (1), Eq. (2) and Eq. (3).

$$\Pr ecision = \frac{Qg \cap Qh}{Qg} \quad (1)$$

$$\text{Re}\,call = \frac{Qg \cap Qh}{Qh} \quad (2)$$

$$F - measure = \frac{(2 * \text{Re}\,call * \text{Pr}\,ecision)}{(\text{Re}\,call + \text{Pr}\,ecision)} \quad (3)$$

For a given document or corpus, the precision and the recall can be calculated as indicated by Eqns. (1) and (2), where the questions generated with a hyper generation are considered as the gold standard. Obviously, this type of assessment is only possible if human-generated questions are already available or if it is possible to have humans generate questions. For example, we choose a text from our corpus

This is questions generated by our system as illustrated by the Figure 7.

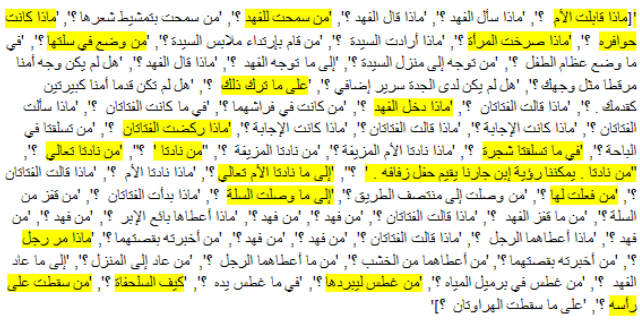


Figure 7. Questions generated by AQG from the text "الفهد" (lion)

Below in Figure 8, we display the questions generated by the Internet users from the same text.

```
<table name="quiz_question_tmp">
    <column name="id">43930</column>
    <column name="Question">؟أَيْنَ وَضَعَتِ الفَتَيَانِ القُرْبَ</column>
    <column name="DocTitle">الفهد</column>
    <column name="URL">https://www.hindawi.org/books/18528352</column>
</table>
<table name="quiz_question_tmp">
    <column name="id">43931</column>
    <column name="Question">؟عَمَّاذَا حَدَثَ لِلْفَهْدِ عِنْدَمَا غَطَسَ يَدَهُ فِي بِرْمِيلِ المَاءِ</column>
    <column name="DocTitle">الفهد</column>
    <column name="URL">https://www.hindawi.org/books/18528352</column>
</table>
<table name="quiz_question_tmp">
    <column name="id">43932</column>
    <column name="Question">؟كَيْفَ كَانَتْ نِهَايَةُ الفَهْدِ</column>
    <column name="DocTitle">الفهد</column>
    <column name="URL">https://www.hindawi.org/books/18528352</column>
</table>
```

Figure 8. Questions writing by the Internet users from the text "الفهد" (lion)

Therefore, in this result we have 66 (Qg) questions generate automatically from the AQG system with the text "الفهد" (lion). 47 (Qg∩Qh) from them are linguistically correct and the others are incorrect (we put them in yellow in the Figure 7) and we have five questions manually writing by the Internet users.

The evaluation of the results obtained by our system is given by the histogram in Figure 9, which graphically illustrates the values of the f-measure obtained for each text. We note that the results obtained are satisfactory and reach up to 86%.

The unsatisfactory f-measures obtained are due to either:

(1) Patterns not defined by AQG. The processing of personal pronouns is in progress.

(2) Incorrect morphological analysis made by MADAMIRA. The authors speak of an error rate of 6%.

A non-fine morphological analysis made by MADAMIRA. This fine analysis will allow us to reach a better understanding of the meaning of the word especially for names. MADAMIRA does not recognize اسم جامد (static noun, aism jamid), اسم علم (proper noun, aism eilm) ,اسم فاعل (Agent noun, aism faeil), اسم مفعول (passive participle, aism mafeul).

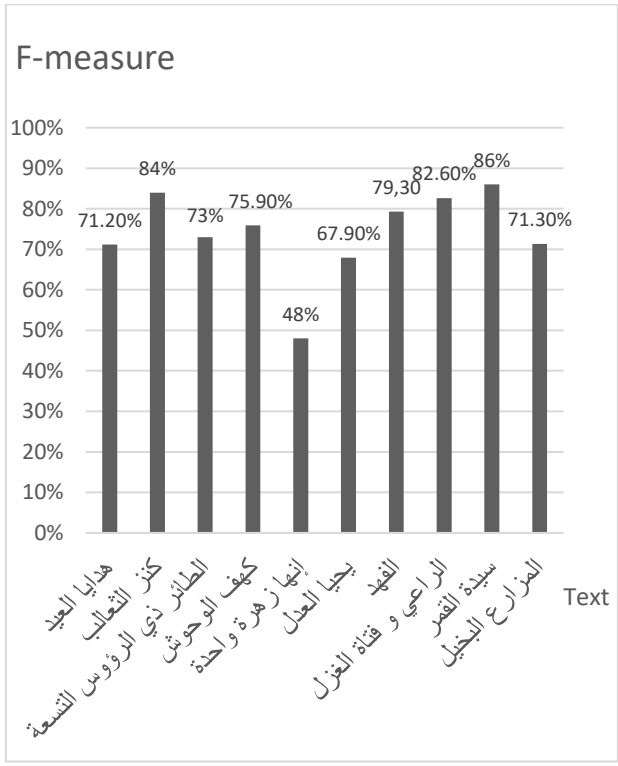


Figure 9. Results of f-measures obtained for evaluation of AQG

Another human-based evaluation of metrics examines several dimensions of the results of issue generation systems. This evaluation proves to be necessary for a correct evaluation of QG results. These metrics measure the possible deficiencies of the questions [14]. Applying these metrics to the AQG corpus, we obtained the results reported in Table 5.

Table 5. Results of the evaluation of incorrect questions generated for 10 texts

Category	Result
Ungrammatical	77.3%
has no sense	8.7%
Vague	7.4%
Bad interrogative tool	6.69%

Ungrammatical questions are grammatically incorrect questions that do not respect the grammar rules of the Arabic language: الصرف و النحو (grammar and conjugation/ sarf w nahw). They represent the highest percentage of the incorrect questions in the evaluation of AQG. This is explained by the fact that we have not yet rewritten the questions generated by the templates by respecting the times and pronouns used in the patterns. The rewrite module is presented as a perspective of this work.

The 8.7% of the category that has no meaning are mainly due to the fact that the input vowels in MADAMIRA are not taken into account. The grammatical category cannot

distinguish, on its own, the A1 and the A0 in VS.

As for the use of the bad interrogative tool, 6.69% is explained by the temporal and place adverbs, which have not been recognized as such. There will, therefore, be a bad attribution of the SRL, the interrogative will be badly chosen especially between متى (when) and أين (where).

The Quizzito designers requested another evaluation of questions that is important for children learning through their platform. The evaluation of the percentage of use of the different Arabic interrogative tools by internet users and by our AQG system. The results are illustrated in Table 6.

Another comparison given in Table 7 clearly shows the large number of questions generated by our system compared to those of voluntary Internet users on the 10 previous texts.

Table 6. Use of interrogative tools by Internet users and by AQG

Interrogatives	Internet users	AQG
الهمزة{ألف الاستفهام} Inversing of verb and subject	0.00%	0.00%
هل /Do you	1.68%	6.00%
أم /Or	0.00%	0.00%
من /who	16.23%	41.00%
ما /what	53.60%	43.00%
كم /how many (how much)	4.63%	1.00%
متى /when	3.82%	4.00%
كيف /how	12.67%	5.00%
أين /where	6.82%	3.00%
أي /which	0.54%	0.00%
أيان /(to ask about time)	0.00%	0.00%
أنى /(to ask about time)	0.00%	0.00%

Table 7. Comparison between the numbers of questions generated by AQG and Internet users

Text	Internet users	AQG
هدايا العيد (hadaya aleid)	10	29
كنز الثعالب (kanz althaealib)	10	31
الطائر ذي الرؤوس التسعة (alttayir dhi alruwuws altse)	5	79
كهف الوحوش (kahf alwuhush)	5	59
إنها زهرة واحدة (iinaha zahrat wahida)	10	48
يحيا العدل (yahya aleadl)	10	46
الفهد (alfahd)	5	87
الراعي وفتاة الغزل (alraaei wafatat alghazl)	10	55
سيدة القمر (sayidat alqamar)	5	56
المزارع البخيل (almazarie albakhil)	5	48

6. CONCLUSIONS

Generating questions can be a time consuming and effortful process. In this research, we work toward automating that process. In particular, we focus on the problem of automatically generating factual questions from individual texts. The comparison between the user-generated questions and the AQG-generated questions is according to three criteria:

(1) Number of questions generated
(2) The interrogatives tools used
(3) Quality and rationality of the generated questions

These criteria allowed us to compare qualitatively and quantitatively the questions generated. The results obtained confirm the high quality of the questions generated by the AQG system.

At the end of this research work, we estimate to have submitted a contribution for the ANLP. Our system proves its effectiveness when the texts are long and where a human does not have the patience to read the text several times to understand it in order to generate questions. The morphological analyzer MADAMIRA has helped us a lot in linguistic preprocessing as well as in the disambiguation of words. Although we have designed all possible cases of verbal and nominal sentences, we have not been able to implement all the cases conceived for the reasons mentioned above concerning the morphological analyzer MADAMIRA.

The results achieved in this paper are very promising for the company Quizzito, as the objective is to automate the questions generation process in the Arabic language. By implementing this method, the company will be able to assist the quiz creators in the process of questions generation, letting them rather focus on the distractor generation task. Adding to that, it will reduce the costs of each quiz in terms of both time and money, allowing a rapid growth of the numbers of quizzes on Arabic content for children.

As perspectives, we plan to improve the grammatical structure of the generated questions, to filter the questions, generate answers, generate multiple-choice questions and validate the answers.

REFERENCES

[1] Bousmaha, K.Z., Rahmouni, M.K., Kouninef, B., Belguith Hadrich, L. (2016) A hybrid approach for the morpho-lexical disambiguation of Arabic. Journal of Information Processing Systems (JIPS), 12(3): 358-380. https://doi.org/10.3745/JIPS.02.0041

[2] Kingsbury, P., Martha P. (2002). From TreeBank to PropBank. Proceedings of the Third International Conference on Language Resources and Evaluation (LREC'02), Las Palmas, Canary Islands, Spain, pp. 1989-1993.

[3] Freydenberger, D. (2013). Extended regular expressions: Succinctness and decidability. Theory of Computing Systems, 53(2): 159-193. http://dx.doi.org/10.1007/s00224-012-9389-0

[4] Papasalouros, A., Chatzigiannakou, M. (2018). Semantic web and question generation: An overview of the state of the art. In: The International Conference E-Learning, pp. 89-192.

[5] Dhawaleswar, R., Saha, K.S. (2018). Automatic multiple choice question generation from text: A survey. IEEE Transactions on Learning Technologies, 13(1): 14-25. https://doi.org/10.1109/tlt.2018.2889100

[6] Kurdi, G., Leo, J., Parsia, B., Sattler, U., Al-Emari, S. (2020). A systematic review of automatic question generation for educational purposes. International Journal of Artificial Intelligence in Education, 30(1): 121-204. https://doi.org/10.1007/s40593-019-00186-y

[7] Duan, N., Tang, D., Chen, P., Zhou, M. (2017). Question generation for question answering. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pp. 866-874. http://dx.doi.org/10.18653/v1/D17-1090

[8] De Viron, L., Bernhard, D., Moriceau, V., Tannier, X. (2011). Génération automatique de questions à partir de

textes en français. Traitement Automatique des Langues Naturelles.
[9] Ciguene, R., Joiron, C., Dequen, G. (2015). Vers la génération automatique de tests d'évaluations différenciés et équitables en contexte universitaire. Environnement Informatique pour l'Apprentissage Humain, Agadir.
[10] Yao, X., Zhang, Y. (2010). Question generation with minimal recursion semantics. In Proceedings of QG, the Third Workshop on Question Generation, pp. 68-75. Citeseer,
[11] Pasha, A., Al-Badrashiny, M., Diab, M., El Kholy, A., Eskander, R., Habash, N., Pooleery, M., Rambow, O., Roth, R. (2014). Madamira: A fast, comprehensive tool for morphological analysis and disambiguation of Arabic. Conference on Language Resources and Evaluation (LREC). 14: 1094-1101.
[12] Belguith, L., Baccour, L., Mourad, G. (2005). Segmentation de textes arabes basée sur l'analyse contextuelle des signes de ponctuations et de certaines particules. in Actes de la 12éme Conférence annuelle sur le traitement Automatique des Langues Naturelles, Dourdan, France, pp. 451-456.
[13] Alja'am J.M., El Saddik A., Sadka A. (2018). Recent Trends in Computer Applications. Best Studies from the 2017 International Conference on Computer and Applications, Dubai, UAE. Springer. http://dx.doi.org/10.1007/978-3-319-89914-5
[14] Lindberg, D.L. (2013). Automatic question generation from text for self-directed learning. Doctoral dissertation, Applied Sciences: School of Computing Science.
[15] Meguehout, H., Bouhadada, T., Laskri, M.T. (2017). Semantic role labeling for Arabic language using case-based reasoning approach. International Journal of Speech Technology, 20(2): 363-372. https://doi.org/10.1007/s10772-017-9412-6
[16] Bonial, C., Babko-Malaya, O., Choi, J.D., Hwang, J., Palmer, M. (2012). English PropBank annotation guidelines. Center for Computational Language and Education Research Institute of Cognitive Science University of Colorado at Boulder, 48.
[17] Punyakanok, V., Roth, D., Yih, W.T. (2008). The importance of syntactic parsing and inference in semantic role labeling. Computational Linguistics, pp. 257-287. https://doi.org/10.1162/coli.2008.34.2.257
[18] Zaghouani, W. (2015). Le développement de corpus annotés pour la langue arabe. PhD thesis, University Paris ouest Nanterre la défense.
[19] Abdul Rahman Tawfiq Al-Omani (2008). Interrogative Tools. Master of Arabic Language, Jordan University.
[20] Tayel Muhammad Ahmad Al-Sarira. (2018). Interrogative Tools in Fluent Arabic. Doctoral thesis, Mutah University, Jordan.
[21] Aziz, B.H.A. (2015). Patterns of conversion in the interrogative style of the Noble Qur'an as a model. Master of Arabic Language. Al Albayt University, Jordan.
[22] Khalid, Q. (2017). The Nominal Sentence in the Last Quarter of the Holy Quran - A Study of grammar. Doctoral dissertation.
[23] El Naeemi, H. (2009). Patterns of Transformation in The Verbal Sentence an Applied Study of the Holy Quran (AAL– E– IMRAN Sura as an example). Magister, Al al-Bayt University, Jordan.
[24] Hardeniya, N., Perkins, J., Chopra, D., Joshi, N., Mathur, I. (2016). Natural Language Processing: Python and NLTK. Packt Publishing Ltd.
[25] Al-Shawakfa, E. (2016). A rule-based approach to understand questions in Arabic question answering. Jordanian Journal of Computers and Information Technology (JJCIT), 2(3). https://doi.org/10.5455/jjcit.71-1467325836
[26] https://magnitt.com/startups/quizzito. Last accessed on 23 March 2020.
[27] https://www.quizzito.com/main/quizzito/fr. Last accessed on 23 March 2020.