UTM UNIVERSITI TEKNOLOGI MALAYSIA
INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING
ISSN 2180-4370
Journal Homepage : https://ijic.utm.my/

A User-Centric Multi-Context Hybrid Reasoning Information Retrieval Model

Samuel Ekene NNEBE[a], Efeosasere Okoro[b], Fatai Idowu SADIQ[a] & Benjamin A. ABARA[b]
[a]Department of Computer Science, Ambrose Alli University,
Ekpoma, Edo, Nigeria
[b]Department of Computer Science, National Institute of Construction Technology and Management,
Uromi, Nigeria

Submitted: 23/8/2021. Revised edition: 9/11/2021. Accepted: 16/11/2021. Published online: 16/5/2022
DOI: https://doi.org/10.11113/ijic.v12n1.337

***Abstract*—Information Retrieval (IR) has been in existence since the 1940s and is impossible to do without. However, research has shown that current information retrieval models do not consider sufficient contexts leading to users in different contexts retrieving the same results. Using a restaurant use case, we propose a user-centric multi-context hybrid reasoning Information Retrieval model to improve the accuracy of retrieved results. Our proposed model uses a hybrid reasoning model of ontology, rules and unsupervised machine learning, considers 14 contexts grouped into user, environmental and database-specific context and considers related domains to the food domain. The result shows that our proposed model outperformed the existing models (location and text-based IR) objectively by 33%. The results suggest that the consideration of a wider range of contexts, a hybrid reasoning model and the consideration of related domains would improve context-based information retrieval significantly.**

***Keywords*—Context-based Information Retrieval, Personalized Information Retrieval, Information Retrieval, Ontology, Rule-based reasoning, Clustering**

I. INTRODUCTION

Information Retrieval (IR) is a branch of Computer Science. It deals with the storage, indexing and accessing of information [1]. The discipline of Librarianship is said to have started modern information retrieval. It involved indexing hard resources using cataloguing schemes [2]. This approach was flawed with increasing information to manage. To overcome this challenge, machines were developed. To quickly go through catalogues for a quick view, but the introduction of computers presented a better prospect. In the 1960s, Computers enabled information to be ranked, sorted and retrieved from storage devices. But computer systems were limited to storage capability as the libraries in the pre-computer era were [3]. This limited the amount of information that could be stored hence reducing the information the user can retrieve. However, with the arrival of the web within the late 1990s, the quantity of data users created multiplied exponentially. As of the year 2018, the amount of data created was 2.5 quintillion per day and ninety per cent of the data in the world was created in the last few years [4]. This leads to a new need. A need not just to find information that is stored and relevant but, relevancy has been re-defined to information that is based on a wide range of contexts and preferences and not just based on query keywords entered alone [5].

A study done by [6], shows that current web search engines are only between 30 to 50% effective. This problem is compounded by the fact that traditional information retrieval is done using keyword-based matching algorithms where queries are matched to common terms. This simply means the document with a higher frequency of a term is more relevant. Due to a wide range of factors, keyword matching algorithms have low accuracy [7], [8].

Search experts are advocating for a more personalised and context-based algorithm that returns customised results based on the user information needs [9]. This is because the accuracy of what an information retrieval service returns is determined by the user and not the system. To do this, we need to ask for more information from the users such as what context are they in and what are their preferences [10]. This data is then used to improve information returned to them and the usability of the system [11].

The most cited description of context describes it as any knowledge that may be used to identify the situation of an individual (person, place and object) related to the relationship between the user and the application, including the user and the application [12]. The context in IR describes all "cognitive and social factors as well as the user's aims and intentions during a search session" [13]. Context affects all aspects of IR, how users interact with an information retrieval system, the result they expect and how decision making and sensemaking is done based on the information retrieved [13]. The context in Information Retrieval can be categorised broadly into three: user context (e.g., location), environmental context (e.g., time, weather, other user reviews, traffic etc.) and database-specific context (e.g., rating etc.) [14].

To integrate context in IR, the system has to identify the user context based on available information in other to deliver more relevant results [7], [15]. One way to infer the user's interest is to assume that they are interested in the information about their location and it is common to use location as the only context [14], [16]. While this is true, there are drawbacks to this. Users information interests are divided into 5: locally based, non-local based, location-based, non-location based and community-based information [7]. Although location-based IR is useful for answering local and location queries and information about physical activities, it does not satisfy a wide range of users non-local information needs.

In this paper, we propose a user-centric multi-context hybrid reasoning Information Retrieval model to improve the accuracy of retrieved results. We use a restaurant use case to show how this model could be applied to IR.

II. RELATED WORKS

In recent years, several solutions have been proposed and developed. Here, we analyse current context-based and personalised information retrieval systems solutions over the last decade using the restaurant domain as a use case. We analyse them based on the context considered, their reasoning models, the use of a user profile and system consideration of related fields that affect food choice such as health. The restaurant domain is used because it represents an active domain for context-based IR challenges where a significant amount of context and preference is needed. It also relates to other large domains like health, nutrition and tourism [17], [18].

TABLE 1. LITERATURE ANALYSIS TABLE

Author	Context considered	Reasoning models	Model Achievements	Model Limitation
[19]	Location	Probabilistic	Proposed a conceptual framework for modern Geographic information retrieval system.	• They did not recognize the context required in IR, such as the environmental context and database-specific context. This leaves them inefficient in answering context-based questions beyond location-based queries • The lack of a user profile meant little personalization. • They did not recognize food-related domains that affect eating preferences such as diet, exercise and wellness.
[20]	Location	Ontology	Provided a dynamic location-based service and increase the information retrieve accuracy especially on the limited mobile screen by accessing cloud application.	• They were limited in the range of context used and personalisation • No user profile was used. • They did not recognise related domains. • Since it was ontologically based, it made uncertainty handling poor
[21]	Location and database-specific	Ontology	Proposed methods of providing Personalized Mobile Information Retrieval System using NFC (Near Field Communication).	• They did not consider the environmental context. • Their reasoning model did not resolve uncertainties well • Their system did not recognise related domains.
[22]	Location	Machine learning	Presented a personalized location-based restaurant recommendation system that integrates with mobile technology.	• Other information retrieval contexts such as environmental contexts and database-specific contexts were not considered. • They also did not consider food-related domains.
[18]	Location and database-specific	Ontology	Integrated food, health, and nutrition domains ontologies to personalized information retrieval systems to retrieve food and health recommendations based on the user's health conditions and food preferences.	• They did not consider environmental context and user location • Their reasoning model handled uncertainty poorly.
[23]	Location and database-specific	Rules	Introduced a novel opinion-based restaurant search engine.	• Reduced personalisation • They did not take into account the environmental context • They did not consider related domains.

Author	Context considered	Reasoning models	Model Achievements	Model Limitation
[24]	Database-specific	Rules	Developed a probabilistic factor analysis framework for exploiting multi-source information for personalized restaurant recommendations.	• They did not consider the environmental context and location context • They did not consider related fields
[25]	Location and database-specific	Ontology	Developed an innovative approach for finding restaurants based on the dishes a user would like to taste	• They did not consider the environmental context • A user profile was absent, their reasoning engine did not handle uncertainties well • They did not consider related fields.
[26]	Location and database-specific	Rules	Proposed a location, time, and preference-aware restaurant recommendation method using the user location, historical check-in data, and the time of the recommendation request.	• Personalisation was minimised with the absence of a user profile • Environmental contexts and related fields were not considered.
[27]	Location and database-specific	Rules	Proposed a personalized restaurant recommendation approach that combines group and customer preferences	• They also did not consider the environmental context and related fields.
[28]	Location	Machine learning	Presented a personalized Point of Interest recommendation system that learns user preferences based on user transaction history and restaurants' locations.	• They did not consider the environmental and database-specific context as well as related domains.

The gaps of previous solutions in the literature analysis table are obvious (1) all the context defined in IR, user, environment and database-specific context are not used. At most, the previous works considered two and none of them considered the environmental context. (2) The lack of user profiles of some reduced personalisation. (3) The limitations of using one reasoning model is as obvious, a combination of two reasoning models is better than one and (4) they did not consider related fields that influence user choice of food such as health and fitness. These made the recent works not effective.

A. Analysis of Existing Context-based Reasoning Models

The core components of a context-based IR system are context acquisition, context modelling and reasoning and context dissemination or representation [29]. Our emphasis in this work is on the context reasoning component. Context reasoning models focus on the deduction of context from homogenous or heterogeneous sources [30]. From the literature analysis table of works done on Context-based Information Retrievals in the food domain (see Table 1) and works from [31]–[33], there are several context reasoning models. They are ontology-based, rule-based, distributed models, Probabilistic, Fuzzy logic, Bayesian networks, Key-based representation, Multi-Entry Bayesian Network (MEBN), Machine Learning (ML), Case-based, Direct sensor, Hidden Markov model etc. The most common context reasoning models are ontology, rule-based and machine learning models.

1) Rule-based System

Rule-based models are a model representation and reasoning support. It allows the assertion of new facts to a knowledge base, which can later be used as input for other rules, making the knowledge base more dynamic. It consists of three key elements: the knowledge base, the fact base and the inference engine. The knowledge base is commonly considered to be a collection of rules in the form of production rules, IF {THIS} THEN {DO THIS}. The fact base includes details used to check if the requirements of the rules are met. The inference engine incorporates a method that allows rules to be processed within the knowledge base. The main downside is ineffective with search and slower to implement such as in the case of ontologies.

2) Ontologies

This is described as a formal explicit specification of a shared conceptualisation [34]. It supports a set of modelling primitives to define classes, their attributes and their relations. It is commonly used to model context hierarchy and dependencies with a context space. The reasoning for ontological models is commonly supported by description logic. They are useful for formalisation and hierarchy generation of knowledge. The main downside is that it does not have ways to infer complex information from existing data, such as in the case of rule-based models.

3) Machine Learning

This is described as a branch of artificial intelligence (AI) that focus on improving computer-based algorithms automatically through the use of input data, events and experiences [35], [36]. ML is used in several applications [37]. In IR it is used in text categorisation, query formulation, information filtering etc. [38]. It is divided into three paradigms: supervised, unsupervised and reinforcement learning [39], [40]. Common supervised ML subtypes are classification and regression [41]. Common unsupervised ML subtypes are clustering or cluster analysis and association.

In the next section, we discuss the methodology used in the designing of the conceptual user-centric multi-context hybrid reasoning Information Retrieval model to improve the accuracy of retrieved results.

III. METHODOLOGY

In this section, we proposed a model that integrates our user-centric multi-context Information Retrieval model with a hybrid reasoning framework to improve the accuracy of retrieved results using a restaurant use case and discussed how our model was evaluated.

A. Proposed Model

To address the limitations of the solutions given in the literature analysis table on Context-based Information Retrievals in the domain of food/restaurant, we propose a user-centric multi-context Information Retrieval model with a hybrid reasoning framework using a restaurant use case. Three phrases stand out here, user-centric, multi-context and hybrid reasoning. User-centric means taking a user profile into account. Multi-context implies the consideration of a broader variety of contexts from different heterogeneous sources, such as environmental, spatial-temporal, social, physical activities, etc. Hybrid here is a synthesis of the strengths of more than one context reasoning model to overcome the limitation that they face individually[42]. We use a mixture of ontology, rule-based and machine learning reasoning because of the individual advantages of these models. The ontology portion of the model would create a better knowledge structure for search efficiency of the content of the domain, the rule-based part will allow for prioritisation of results based on certain conditions and the ML part will allow for user segmentation.

1) Model Description

Fig. 1 depicts the framework for the proposed User-Centric Multi-Context Hybrid Reasoning Information Retrieval Model.

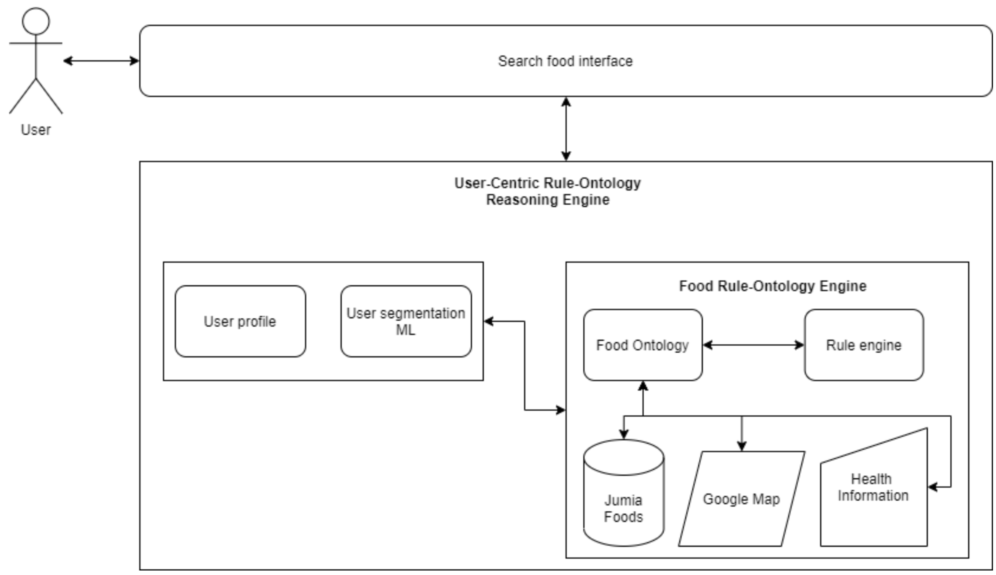

Fig. 1. The framework for the User-Centric Multi-Context Hybrid Reasoning Information Retrieval Model

The proposed framework as depicted in Fig. 1 consists of two major components, a Search food interface and a user-centric rule-ontology reasoning engine. The search food interface component interfaces between the user and the user-centric rule-ontology reasoning engine. The User-centric rule-ontology reasoning engine consists of a user profile component, user segmentation machine learning component and a food rule-ontology engine. The user profile component includes user preferences and history. User preferences such as user dietary requirements, personalised rating, if they are vegetarians or not, allergies, budget, illness etc. The user segmentation machine learning component contains an algorithm to cluster users into various segments based on their demography. The food rule-ontology engine is the engine that houses the framework reasoning i.e., the integration between the ontology and rules. The food rule-ontology engine is made up of the food ontology and the rule engine. Individually, both reasoning models have their limitations so we are combining them to reduce their limitations. The food ontology is an ontology that consists of the taxonomy of restaurants and the foods they sell. This can be accessed from sites such as web-based food applications like Jumia Food [43]. Jumia food is a web-based food delivery application that provides information from a wide variety of restaurants across most Nigerian states [43]. The Jumia Food website is considered because its data is accessible manually and automatically, and has a wide variety of restaurant-related attributes. Environmental information can be accessed from the Google map. Google map provides the functionality to track traffic around an environment, infer road conditions and determine distance from one address to another. Health-based information can be accessed from the web either manually or automatically.

The food rule-ontology reasoning engine is a synthesis of the two reasoning models. It is used to improve the accuracy of IR by considering the user query entered, their preference and various context such as user context, environmental or physical context, database-specific context, context history.

This taxonomy in Fig. 2 can be combined with rules. For example, if user activity is driving and a search is made, we can say, consider a restaurant where the weather is rainy or sunny and if the user activity is walking, only consider restaurants with sunny weather. This can be expressed with rules as:

IF user activity: driving
    RestaurantTemperature = Sunny OR Rainy
ELSE IF user activity = walking
    RestaurantTemeperature = Sunny

1.1.1.1 Ontologies

We adopted the restaurant-food ontology that is based on ontology web language (OWL) and implemented it in Protégé ontology editor [44].

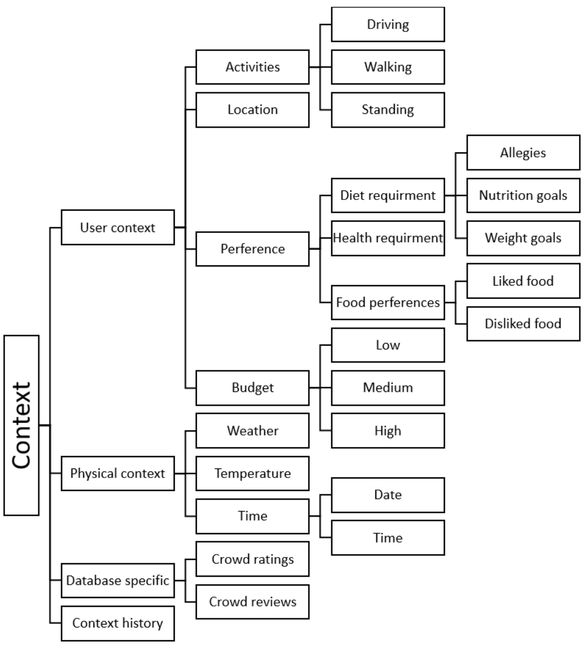

Fig. 2. A taxonomy of food context concepts

1.1.1.2 Rules

We used a simple weighting multiple criteria decision analysis (MCDA) approach system as a ranking system to return the best results. An MCDA approach is considered because ranking in multi-context-based information retrieval requires multiple criteria. Using an MCDA approach for ranking is not new. It was used for the ranking of usability attributes of mobile health applications in Nigeria [45]. We used different weighting for the various criteria assuming that the user context is the most important IR context category in this study.

$R_{\text{with profile}} = \Sigma\,(C)$
$C = \Sigma\,(D, E, U)$
$C = 1$
$D = 15\%$, $E = 15\%$ and $U = 70\%$

Where R is the ranking score, C is context, D is database-specific context and E is environmental context and U is user context.

Here, database-specific context is expressed as 3 subcontexts: restaurant rank, opened and closed status and nearness. Environmental context is expressed as 2 subcontexts: weather and traffic while user context is expressed as 9 subcontexts: user health, taste, price, delivery time, friend and family preferences, religion, vegetarian belief, food quality and prior ranking.

1.1.1.3 User Segmentation

In other to perform user segmentation, an extensive review was done aimed at identifying the various factors that users of online food delivery services prioritise when choosing food or restaurant. These variables were measured using a web-based questionnaire. It contained 14 questions which were organised into two sections: the first contained information related to demography while the second focused on food choice preferences and their priority in food choices. 10 Variables that included food choice preferences were used and they were: past experiences, hunger, health, the taste of food, price, delivery time, advice and suggestions from friends and family, religion, food quality, and vegetarian belief. Each statement measured the priority placed on each variable. A five-point Likert's scale was used to record the responses. 121 persons who had previously ordered food online took part in the study. The participants ranged from ages 19 – 45. Participants were recruited randomly.

A cluster analysis was done using K-means unsupervised clustering to segment the users based on their demography. Cluster analysis is a technique used for the grouping of subjects into groups based on certain criteria [46]. Clustering allows for the grouping of data according to certain traits. After examining the output of two to eight cluster solutions based on the respective distances, an eight-cluster solution was selected.

2) Model Evaluation

We are hypothesising that given the significance of multiple contexts and user preferences; the user-centric multi-context hybrid reasoning IRS would return more relevant results than existing systems such as location-based IRS and traditional text-based algorithms.

To evaluate the hypothesis, a study was set up with IRS as the independent variable and effectiveness expressed as objective and subjective relevance as the dependent variable. Effectiveness here deals with the retrieving of the most relevant documents to a user's need [50].

Since the core part of IR is for a user to enter a query and the system returns results that are relevant to them, standard IR measures such as precision and recall can be used as metrics for objective relevance and asking users to rate the search output of the various IRS could be used as the metric for subjective relevance.

In this study, participants are recruited to perform a task where they are asked to search for selected information using the user-centric multi-context hybrid reasoning IRS and other existing IRS (location and text-based) and then to rank them. The user-centric multi-context hybrid reasoning IRS is built with a wide range of context and association to related domains. Location-based IRS considers the user location and database context and text-based IR only considers the database context.

3.1.1.1 Evaluation Design

The study used a wholly repeated measure experimental design. Here, each participant uses all the IRS under evaluation. To prevent the likelihood of order, practice or fatigue effect, a counterbalance of the task was used.

There was one independent variable: IRS (with three levels: multi-context-based, location-based and text-based). The dependent variable was effectiveness measured in terms of objective and subjective relevance. Participants were asked to search for a range of queries using the various IRS.

We used a natural experimental method. Here, the user is left to take part in the experiment in their natural environment. The benefit of this experimental method is we can assume that the user's behaviour in this experiment is more likely to reflect real-life behaviours (very high ecological validity).

3.1.1.2 Measure of Relevance

Relevance denotes how well a retrieved document meets the user information need. There are two main classes of relevance, they are objective and subjective [51]. The objective class of relevance measures system performance using standard measures such as precision and recall [52]. The subjective class of relevance measures the user's perceived relevance of a system results or answer to their information needs and preferences.

3.1.1.3 Participants

34 participants took part in the study. All participants voluntarily accepted to take part in the study. Participants were recruited through social media groups and personal chat rooms. 35% were female and 65% male with an age range of 19 to 35.

3.1.1.4 Material

The dataset was acquired from Jumia Food [43] and Google Maps. It contained information of 21 restaurants manually extracted from Jumia food's web application after a filter search of Lagos City was applied. Jumia food is a food delivery web application that houses the information of a wide range of restaurants across most states in Nigeria [43]. Google Map information of the various restaurants was used to add nearness and weather information to the database. The general database attributes were: restaurant name, image, location, rating, price, description, category, suitability, open and close time, traffic, nearness and delivery duration.

An application built to illustrate the algorithms of the three levels of the independent variable was used. The application is available at foodme.glideapp.io. The application also contains a questionnaire for measuring subjective relevance.

3.1.1.5 Procedure

Participants were given four tasks to complete. The tasks were to open the application, enter two single word queries using the various IRS of evaluation and rank which IRS returned results reflecting their choices based on the queries they entered. They were also asked to enter the reason for their choice.

IV. System implementation

The model was implemented using GlideApp visual programming application [47].

GlideApp uses Google sheets Application Programming Interface (API) in Node.js to connect the GlideApp client view with Google Sheets (Fig. 3). Google API allows the developer to build applications that can read, write and update a Google spreadsheet [48], [49]. It can also act as a database where an application features data and renders it to a user interface (UI) template.

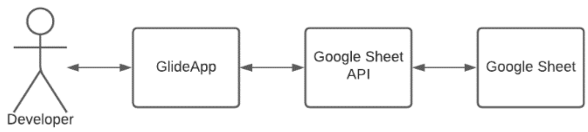

Fig. 3. Illustrating the relationship between the developer, GlideApp, Google Sheet API and Google Sheet

A. Ranking Restaurant Algorithm

Two broad cases are considered in the ranking of restaurants: If the user has a profile and if they do not. If the user does not have a profile, the database-specific and environment contexts category is considered excluding the user context category. However, if they have a profile, the user context category is included. The ranking of a restaurant when the user has a user profile is based on the user cluster.
Where cluster priority (5 highest to 1 lowest priority)

Priority 1 (P1) = 5
Priority 2 (P2) = 4
Priority 3 (P3) = 3
Priority 4 (P4) = 2
Priority 5 (P5) = 1

The weightings of 5%, 7.5% and 7.78% are given to each database specific, environmental and user context respectively (see the section on rules).

LET CF = SUM (BT, BU, BV, BW, BX)
LET BT = M * 5%
// Where M is database restaurant rating expressed as number 1 to 5. 1 denotes low rating and 5 denotes high rating//

LET BU = AJ * 5%
// Where AJ is restaurants open and closed status expressed as number 1 to 5. 1 denotes closed and 5 denotes open//

LET BV = AP * 5%
// Where AP is restaurant nearness expressed as number 1 to 5. 1 denotes far and 5 denotes near//

LET BW = AR * 7.5%
// Where AR is traffic to restaurant expressed as number 1 to 5. 1 denotes heavy traffic and 5 denotes no traffic//

LET BX = AT * 7.5%
//Where AT is weather expressed as number 1 to 5. 1 denotes sunny and 5 denotes heavy rain//

LET BY = BL * 7.78%
// Where BL is food taste. In the absence of further data, we assumed food taste equals rating. //

LET BZ = BM * 7.78
//Where BM is food price expressed as number 1 to 5. 1 denotes low price and 5 denotes high price//

LET CB = BP * 7.78%
//Where BP is religion expressed as number 1 to 5. 1 denotes no religious preferences and 5 denotes high religious preferences. //

LET CC = BQ * 7.78%
//Where BQ is vegetarian beliefs expressed as number 1 to 5. 1 denotes no vegetarian preferences and 5 denotes high vegetarian preferences. //

LET CA = BO * 7.78%
//Where BO is friend and family influence expressed as number 1 to 5. 2 denotes no family and friends influence and 5 denotes high family and friends' influence. //

LET CD = BR * 7.78%
//Where BR is food quality. In the absence of further data, we assumed food quality equals rating. //
LET CE = BS * 7.78%
//Where BS is delivery time expressed as number 1 to 5. 1 denotes slow time and 5 denotes fast time. //

```
IF userprofile = 0
        Restaurant_ranking = CF
ELSE
        Detect the user_cluster
                IF user_cluster = 0
                        Restaurant_ranking = CF +
P2(SUM (BZ, CE, CD)))
                ELSE IF user_cluster = 1
                        Restaurant_ranking = CF +P1
(SUM (BY, CE, CD)) +P2(SUM(BZ))
                ELSE IF user_cluster = 2
                         Restaurant_ranking =
CF+P1(SUM (BY, CD)) +P2(SUM(CE)
                ELSE IF user_cluster = 3
                        Restaurant_ranking =CF+P1(SUM
(CB)) +P2(SUM(CD))
                ELSE IF user_cluster = 4
                        Restaurant_ranking =CF+P1(SUM
(BY, BZ, CE, CD)) +P2(SUM(CC))
                ELSE IF user_cluster = 5
                        Restaurant_ranking = CF+P2(SUM
(CC, CB, CE))
                ELSE IF user_cluster = 6
                        Restaurant_ranking =CF+P1(SUM
(CD, BY)) +P2(SUM (CC, CB))
                ELSE
                        Restaurant_ranking =CF+P1(SUM
(BY, CA, CD)) +P2(SUM(CE))
                END IF
END IF
```

Code 1. Restaurant ranking pseudocode

The pseudocode in code 1 checks if the user has a user profile or not. If the user has a profile, the algorithm checks the user cluster. Based on the user cluster labelled from 0 -7, the algorithm multiplies the rating of the context by the priority assigned to each context. Each cluster places different levels of priority on each of the contexts. Grouping them into priorities 1 to 5.

V. RESULTS

An analysis of variance (ANOVA) and Pearson's Chi-square test for independence test was performed to test for statistical significance of the objective and subjective relevance result respectively. The significance level used was 0.01.

1) *Objective Relevance*

This is relevance based on system performance in terms of precision and recall.

a) Precision at n

Precision is one of the standard measures for information retrieval. Precision at n is the ratio of documents properly identified as relevant in the top 3 retrievals. Users were asked to perform two single word queries. Table 3 shows the precision at n score for the two queries across the models. The average precision was calculated.

TABLE 2. PRECISION VALUE ACROSS ALL MODELS FOR VARIOUS QUERIES

	LOCATION	**MULTI-CONTEXT**	**TEXT**
Query 1	49%	80.4%	60%
Query 2	45%	80%	38%
Average Precision	47%	81%	49%

Table 2 shows that the user-centric multi-context hybrid reasoning IRS had a better precision (81%) than the location-based model (47%) and the text-based model (49%) on average. The ANOVA test reported a significant difference between the multi-context, location-based and text-based IRS ($F(2, 48) = 9.688637$, $p < 0.01$).

b) Recall at n

Recall is also a standard measure for information retrieval. Recall at n measures the proportion of the total number of relevant documents identified among the total number of relevant documents in the document population for the top 3 retrieved items. Users were asked to perform two single word queries. Table 3 shows the recall at n scores for the two queries across the models. The average recall was calculated.

TABLE 3. RECALL VALUE ACROSS ALL MODELS FOR VARIOUS QUERIES

	LOCATION	**MULTI-CONTEXT**	**TEXT**
Query 1	100%	100%	100%
Query 2	7%	100%	7%
Average Recall	54%	100%	54%

Table 3 shows that the multi-context-based IRS had a better recall (100%) than the location-based model (54%) and the text-based model (54%). The ANOVA test reported a significant difference between the multi-context-based, location-based and text-based IRS ($F(2, 96) = 17$, $p < 0.01$).

c) F Measure Analysis

The F measure is the harmonic mean of precision and recall. It measures the effectiveness of retrieval for both precision and recall [53].

TABLE 4. F MEASURE VALUE DISTRIBUTION ACROSS ALL IRS

	IRS		
	LOCATION	MULTI-CONTEXT	TEXT
Precision	41.00%	82.00%	60.00%
Recall	54.00%	100.00%	54.00%
F measure	46.61%	90.11%	56.84%

Table 4 shows that the multi-context-based IRS has a better F-measure (90.11%) than the location-based (46.61%) and the text-based IRS (56.84).

2) Subjective Relevance

This is relevance based on human performances in terms of user rating

a) User Rating

Fig. 4 shows that participants consider the user-centric multi-context hybrid reasoning IRS to return more relevant items (73%) to them than the one returned by the location-based IRS (20%) and the text-based IRS (7%).

The Chi-Squared test reported a significant difference between the multi-context and location-based IR (X-squared = 30, df = 2, p-value < 0.01) and multi-context and text-based IR (X-squared = 30, df = 2, p-value < 0.01). However, there was no significant different between the location-based IR and text-based IR (X-squared = 6.8805e-30, df = 1, p-value > 0.01).

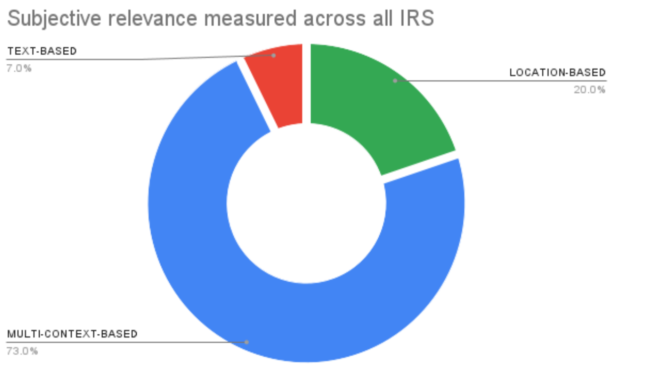

Fig. 4. Pie chart showing the percentage of subjective relevance measured across all IRS

b) Reason for the Preferred Choice

The user responses were also analysed to determine the various reasons for their preferred choice of system.

TABLE 5. PERCENTAGE FREQUENCY OF THE VARIOUS REASONS FOR CHOICE ACROSS ALL IRS

	IRS		
Reason for choice	LOCATION	MULTI-CONTEXT	TEXT
Close to context and preference	22.00%	78.00%	0.00%

Table 5 shows that 22%, 78% and 0% of users selected the Location-based, multi-context-based and text-based IRS respectively because of context closeness.

VI. Conclusion

This paper proposed a User-Centric Multi-Context Hybrid Reasoning Information Retrieval Model using the restaurant domain as a case study. It intends to extract many contextual considerations relevant to a particular user search for food and restaurant-based information to improve their overall information retrieval. The model incorporates ontology, rule and machine learning reasoning techniques, multiple contexts and a user profile.

The overall result shows a user-centric multi-context hybrid reasoning IRS that outperforms the existing system (location and text-based IR) objectively by 33%. These results are consistent with the assumptions that IRS that are user-focused, considers user preferences and a wide range of context will outperform existing IRS. The user-centric multi-context hybrid reasoning IRS had an F measure score of 90% while the existing IRS had between 47 – 57%. This finding is consistent with the finding of [6] that shows that current IRS are only between 30 to 50% effective. Also, Multi-context-based IR had a subjective relevance score of 73% while the existing IRS has between 7-20%. The reason for user choices explains this statistic; the multi-context-based IR was more aligned with user preferences and expectations. This is also in line with the thoughts of [54] that users prefer applications that consider more context.

A. Limitation

The limitation of this study is the distribution of participants in the identification and analysis of user information retrieval requirements/behaviours and user evaluation of the system did not reflect the aged and people with medical conditions

B. Further Works

We intend to test our approach in other domains such as taxi-hauling and eLearning to determine its level of generalisability.

Acknowledgment

The authors acknowledge Ambrose Alli University and the National Institute of Construction Technology and Management for providing support for this research.

VII. References

[1] S. Ranwez, V. Ranwez, M.-F. Sy, J. Montmain, and M. Crampes. (2012). User Centered and Ontology-based Information Retrieval System for Life Sciences. *BMC Bioinformatics*, 13(Suppl 1), S4. Doi: 10.1186/1471-2105-13-S1-S4.

[2] M. Sanderson and W. B. Croft. (2012). The History of Information Retrieval Research. *Proceedings of the IEEE*, 100(Special Centennial Issue), 1444-1451. Doi: 10.1109/JPROC.2012.2189916.

[3] T. Cuff. (2019). *Computer Memories, A History, Revision 2*. Doi: 10.13140/RG.2.2.36686.41282.

[4] B. Marr. (2018). How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. *Forbes*, May 21, https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/ (accessed Jan. 08, 2019).

[5] S. Gong. (2012). The Personalized Information Retrieval Model Based on User Interest. *Physics Procedia*, 24, 817-821. Doi: 10.1016/j.phpro.2012.02.122.

[6] D. Lewandowski. 2008. The Retrieval Effectiveness of Web Search Engines: Considering Results Descriptions. *Journal of Documentation*, 64(6), 915-937. Doi: 10.1108/00220410810912451.

[7] D. Lee. (2008). To Find or to be Found, That is the Question in Mobile Information Retrieval.

[8] R. Zhang, K. Zettsu, Y. Kidawara, and Y. Kiyoki. (2012). Web Service Ranking Based on Context. *2012 Second International Conference on Cloud and Green Computing*, Xiangtan, Hunan, China, 375-382. Doi: 10.1109/CGC.2012.11.

[9] A. Bouramoul, M.-K. Kholladi, and B.-L. Doan. (2010). PRESY: A Context Based Query Reformulation Tool for Information Retrieval on the Web, *Journal of Computer Science*, 6(4), 470-477. Doi: 10.3844/jcssp.2010.470.477.

[10] M. F. P. Coutinho, M. K. Asnani, and M. D. A. Caeiro, 'Context Based Information Retrieval. (2012). *International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE)*, 1(7), 35-38.

[11] M. Hadjouni, M. R. Haddad, H. B. Zghal, M.-A. Aufaure, and H. B. Ghezala. (2009). Personalized Information Retrieval Approach. *Sixth International Workshop on Web Information Systems Modeling (WISM'09), Jun.* Accessed: Jul. 08, 2019. [Online]. Available: https://www.researchgate.net/publication/261511032_Personalized_Information_Retrieval_Approach

[12] A. K. Dey. (2000). Providing Architectural Support for Building Context-aware Applications. PhD Thesis. Georgia Institute of Technology, Atlanta, GA, USA.

[13] D. Boughareb and N. Farah. (2014). Context in Information Retrieval. *2014 International Conference on Control, Decision and Information Technologies (CoDIT)*, 589-594. Doi: 10.1109/CoDIT.2014.6996961.

[14] M. F. Mokbel and J. J. Levandoski. (2009). Toward Context and Preference-aware Location-based Services. *Proceedings of the Eighth ACM International Workshop on Data Engineering for Wireless and Mobile Access*, New York, NY, USA, 25-32. Doi: 10.1145/1594139.1594150.

[15] D. K. Limbu, A. M. Connor, and S. G. MacDonell. (2018). A Framework for contextual information Retrieval from the WWW. Accessed: Sep. 18, 2018. [Online]. Available: https://arxiv.org/abs/1407.6100.

[16] S. Ilarri, A. lllarramendi, E. Mena, and A. Sheth. (2011). Semantics in Location-Based Services. *IEEE Internet Computing*, 15(6), 10-14. Doi: 10.1109/MIC.2011.156.

[17] B. Agarwal, N. Mittal, P. Bansal, and S. Garg. (2015). Sentiment Analysis Using Common-Sense and Context Information. *Computational Intelligence and Neuroscience*. https://www.hindawi.com/journals/cin/2015/715730/ (accessed Aug. 31, 2020).

[18] T. Helmy, A. Al-Nazer, S. Al-Bukhitan, and A. Iqbal, Health. (2015). Food and User's Profile Ontologies for Personalized Information Retrieval. *Procedia Computer Science*, 52, 1071-1076. Doi: 10.1016/j.procs.2015.05.114.

[19] X. Lin. (2020). Modern GIR Systems: Framework, Retrieval Model and Indexing Techniques. Accessed: Aug. 26, [Online].

Available: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-34787.

[20] K. S. Shetty and S. Singh. (2011). Cloud Based Application Development for Accessing Restaurant Information on Mobile Device using LBS. *arXiv:1111.1894 [cs]*. Doi: 10.5121/iju.2011.2404.

[21] O. Choi, K. Kim, D. Wang, H. Yeh, and M. Hong. (2012). Personalized Mobile Information Retrieval System. *International Journal of Advanced Robotic Systems*, 9(1), 11. Doi: 10.5772/50910.

[22] A. Gupta and K. Singh. 2013. Location based Personalized Restaurant Recommendation System for Mobile Environments. *2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 507-511. Doi: 10.1109/ICACCI.2013.6637223.

[23] J. Choi *et al.* (2015). Smith Search: Opinion-based Restaurant Search Engine. *Proceedings of the 24th International Conference on World Wide Web*, Florence, Italy, May, 187-190. Doi: 10.1145/2740908.2742829.

[24] J. Sun, Y. Xiong, Y. Zhu, J. Liu, C. Guan, and H. Xiong. (2015), Multi-source Information Fusion for Personalized Restaurant Recommendation. *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, New York, NY, USA, 983-986. Doi: 10.1145/2766462.2767818.

[25] G. Barlacchi, A. Abad, E. Rossinelli, and A. Moschitti. (2016). Appetitoso: A Search Engine for Restaurant Retrieval based on Dishes. *CEUR Workshop Proceedings*, 1749. Accessed: Feb. 24, 2020. [Online]. Available: https://qfrd.pure.elsevier.com/en/publications/appetitoso-a-search-engine-for-restaurant-retrieval-based-on-dish.

[26] Md. A. Habib, Md. A. Rakib, and M. A. Hasan. (2016). Location, Time, and Preference Aware Restaurant Recommendation Method. *2016 19th International Conference on Computer and Information Technology (ICCIT)*, 315-320. doi: 10.1109/ICCITECHN.2016.7860216.

[27] C. Zhang, H. Zhang, and J. Wang. (2018). 'Personalized Restaurant Recommendation Method Combining Group Correlations and Customer Preferences. *Information Sciences*, 454-455, 128-143. Doi: 10.1016/j.ins.2018.04.061.

[28] M. Du, R. Christensen, W. Zhang, and F. Li. (2019). Pcard: Personalized Restaurants Recommendation from Card Payment Transaction Records. *The World Wide Web Conference*, New York, NY, USA, 2687-2693. Doi: 10.1145/3308558.3313494.

[29] X. Li, J.-F. Martínez, and G. Rubio. (2017). Towards a Hybrid Approach to Context Reasoning for Underwater Robots. *Applied Sciences*, 7(2), Art. 2. Doi: 10.3390/app7020183.

[30] S. Jie and W. ZhaoHui. (2006). Context Reasoning Technologies in Ubiquitous Computing Environment. *Embedded and Ubiquitous Computing*, Berlin, Heidelberg, 1027-1036. Doi: 10.1007/11802167_103.

[31] A. Bikakis and G. Antoniou. (2010). Rule-based Contextual Reasoning in Ambient Intelligence. *Semantic Web Rules*, Berlin, Heidelberg, 74-88. Doi: 10.1007/978-3-642-16289-3_8.

[32] E. Gilman. 2015. Exploring the use of Rule-based Reasoning in Ubiquitous Computing Applications. University of Oulu, Finland, [Online]. Available: http://urn.fi/urn:isbn:9789526209586.

[33] G. Nalepa and S. Bobek. (2014). Rule-based Solution for Context-aware Reasoning on Mobile Devices. *ComSIS*, 111, pp. (171-193). Doi: 10.2298/CSIS130209002N.

[34] T. R. Gruber. (1995). Toward Principles for the Design of Ontologies UsedFor Knowledge Sharing? *International Journal of Human-Computer Studies*, 43(5), 907-928. Doi: 10.1006/ijhc.1995.1081.

[35] V. Sindhu, S. Nivedha, and Prakash. (2021). An Empirical Science Research on Bioinformatics in Machine Learning – Journal. *Journal of Mechanics of Continua and Mathematical Sciences*, 7. Accessed: Aug. 02, 2021. [Online]. Available: https://www.journalimcms.org/special_issue/an-empirical-science-research-on-bioinformatics-in-machine-learning/.

[36] K. Sultan, H. Ali, and Z. Zhang. (2021). Big Data Perspective and Challenges in Next Generation Networks', *Future Internet*, vol. 10, no. 56, 2018, Accessed: Aug. 02, 2021. [Online]. Available: https://www.researchgate.net/publication/325928183_Big_Data_Perspective_and_Challenges_in_Next_Generation_Networks.

[37] L. Si and R. Jin. (2021). Machine Learning for Information Retrieval. *Conference on Research and Development in Information Retrieval, Beijing, China*, Jan. 2011. Accessed: Aug. 02, 2021. [Online]. Available: https://www.researchgate.net/publication/221299750_Machine_learning_for_information_retrieval.

[38] S. J. Cunningham, J. Littin, and I. H. Witten. (1997). Applications of Machine Learning in Information Retrieval.

[39] T. Jo. (2021). Introduction. Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning. T. Jo, Ed. Cham: Springer International Publishing, 3-22. Doi: 10.1007/978-3-030-65900-4_1.

[40] A. Dasgupta and A. Nath. (2016). Classification of Machine Learning Algorithms. *International Journal of Innovative Research in Advanced Engineering*, 3, 6-11. Doi: 10.6084/M9.FIGSHARE.3504194.V1.

[41] E. Alpaydin. 2014. *Introduction to Machine Learning*. Third Edition. MIT Press.

[42] E. Okoro, B. Abara, U. Alex, A. Ajonye, and Z. Isa. (2018). A Hybrid Approach to Fake News Detection on Social Media. *Nigerian Journal of Technology*, 37. Doi: 10.4314/njt.v37i2.22.

[43] Jumia Food, 'Jumia Food Nigeria', *Jumia Food*, 2020. https://food.jumia.com.ng/ (accessed Jan. 04, 2021).

[44] E. S. Nnebe, E. Okoro, and B. A. Abara, 'Restaurant Multi-Context-Based Information Retrieval System Ontological Model', *Covenant Journal of Informatics and Communication Technology*, 9(1), Art. no. 1. Accessed: Jul. 04, 2021. [Online]. Available: https://journals.covenantuniversity.edu.ng/index.php/cjict/article/view/2614.

[45] F. Kasali, O. Adekola, I. Akinyemi, I. Ebo, and J. Balogun. (2021). Model for Ranking the Usability Attributes of Mobile Health Applications in Nigeria (MCDM Approach). *Covenant Journal of Informatics and Communication Technology*, 8(2), Art. no. 2. Accessed: May 31, 2021. [Online]. Available: https://journals.covenantuniversity.edu.ng/index.php/cjict/article/view/2440.

[46] S. Prashar, T. S. Vijay, and C. Parsad. (2016). Segmenting Online Shoppers: A Study of Consumers' Web Portal Selection Motivations for E-shopping, 21 27-46.

[47] GlideApp. (2021). Glide. https://www.glideapps.com/ (accessed Sep. 22, 2021).

[48] Code Labs. 2021. Use Google Sheets as Your Application's Reporting Tool. *Google Codelabs*. https://codelabs.developers.google.com/codelabs/sheets-api (accessed Jul. 19, 2021).

[49] Google API. (2021). *googleapis/google-api-nodejs-client*. Google APIs. Accessed: Jul. 19, 2021. [Online]. Available: https://github.com/googleapis/google-api-nodejs-client.

[50] R. Blanco and F. Silvestri. (2008). Efficiency Issues in Information Retrieval Workshop. *Advances in Information Retrieval*, Berlin, Heidelberg, 711-711. Doi: 10.1007/978-3-540-78646-7_84.

[51] P. Borlund. (2003). The Concept of Relevance in IR. *Journal of the American Society for Information Science and Technology*, 54(10), 913-925, Doi: https://doi.org/10.1002/asi.10286.

[52] P. Clough and M. Sanderson. (2013). Evaluating the Performance of Information Retrieval Systems Using Test Collections. *Information Research*, 18(2), Art. no. 2. Accessed: Apr. 23, 2021. [Online]. Available: http://www.informationr.net/ir/18-2/paper582.html#.U2unLPldXTp.

[53] C. J. van Rijsbergen. (1981). Retrieval Effectiveness. *Information Retrieval Experiment*, Butterworths.

[54] M. Lee and J.-D. Cho. (2014). Logmusic: Context-based Social Music Recommendation Service on Mobile Device. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, New York, NY, USA, 95-98. Doi: 10.1145/2638728.2638749.