Aplicație bazată pe microservicii pentru identificarea persoanelor cu interese similare

Radu-Valentin Cornea  
UNIVERSITATEA TEHNICĂ,,Gheorghe Asachi” din IAȘI  
FACULTATEA DE AUTOMATICĂ  
ȘI CALCULATOAREIași, România[radu-valentin.cornea@student.tuiasi.ro](mailto:radu-valentin.cornea@student.tuiasi.ro)  
Coordonator științific:  
Ș.l. dr. inf. Tiberius Dumitriu

*Abstract*—În această lucrare se propune prezentarea unor metode de recomandare a anumitor persoane, în funcție de preferințele pe care le au în comun cu ceilalți utilizatori, folosind metode din sfera inteligenței artificale. Accentul cade pe grupuri de interes din sfera educațională, dar poate fi extins la orice alte domenii.

Studiul redă câteva aspecte generale ale aplicațiilor bazate pe microservicii, modalități de securizare a aplicației, tehnologiile folosite, motivația realizării unei aplicații diferite de cele existente, arhitectura realizată și rezultatele obținute. Experimentele efectuate presupun aplicarea unor algoritmi de inteligență artificială, precum k-nearest neighbors (KNN) pentru a i se recomanda unui nou utilizator, pe baza preferințelor sale, un grup de persoane (cunoscute sau nu) ce au interese similare. Pentru aceasta se utilizează valori diferite ale parametrului K, precum și mai multe metrici (Jaccard, euclidiană, cosine etc.) cu scopul de a detemina cea mai potrivită recomandare. Rezultatele obținute sunt prezentate sub formă de tabele și grafice care evidențiază avantajele acestei abordări.

Pe baza experiențelor anterioare obținute în urma recomandărilor, unui utilizator nou i se poate asocia rapid un grup de persoane cu interese similare.

Cuvinte cheie—Aplicații software, Aplicații educaționale, Inteligență artificală

# Introducere

Proiectul constă într-o aplicație pentru interconectarea oamenilor, în funcție de preferințele lor, utilizând diverși algoritmi. Unii dintre algoritmii utilizați sunt mai simpli (căutari directe pe utilizatori utilizând paradigma funcțională), iar alții folosesc inteligență artificială. Se dorește și recomandarea utilizatorilor în funcție de alte filtre precum distanța geografică, dar momentan recomandările se fac doar în funcție de preferințe. Scopul proiectului este de a uni oamenii mai ușor în scop educațional în funcție de preferințele legate de tehnologii sau concepte teoretice, însă proiectul ar putea fi folosit și în alte arii decât cele educaționale.

Aplicația se folosește de modelul client-server, iar arhitectura este bazată pe servicii/microservicii. Câteva dintre serviciile implementate până acum cuprind: IDM (Identity Management), Profile, Algorithms și Gateway-ul ce leagă serviciile. Alte exemple de servicii utile aplicației, dar neimplementate încă, sunt cele pentru identificarea locației și pentru suport-ul utilizatorului (în cazul în care acesta are nevoie de ajutor, să poată depune cereri). Pentru o bună funcționalitate, dar și dezvoltare mai facilă a aplicației, părțile de cod au fost despărțite pe module, unele dintre ele având propriul proiect și server, tocmai pentru a mări decuplarea aplicației.

Pentru a fi siguri că aplicația este una sigură în ceea ce privește datele utilizatorilor, s-au utilizat divere mecanisme de protecție a datelor, precum JWT-urile, păstrate la nivelul clientului în format criptat (tocmai pentru a se evita schimbarea câmpurilor din interiorul lor), ele fiind mai întâi decriptate la nivelul server-ului, apoi se validează formatul și semnătura acestora, urmând abia apoi să se realizeze validările pe câmpurile lor (de exemplu, există un câmp expiry), autorizarea realizându-se abia la sfârșit, în funcție de autoritatea pe care o deține utilizatorul respectiv. Pentru a persista parolele, acestea sunt ținute în baza de date criptate cu ajutorul funcției BCrypt, iar în cazul în care un client se autentifică, la verificare, parola introdusă este criptată și ea și comparată cu valoarea din baza de date. S-ar fi putut folosi alte metode pentru autentificare și autorizare, precum OAuth2, SAML, OpenID, dar în scop demonstrativ s-au folosit doar JWT-urile.

Tehnologiile și conceptele teoretice utilizate până acum pentru rezolvarea problemei sunt:

* Backend: Spring, Spring Security, JWTs, REST, Kotlin, MariaDB, MongoDB, criptare, decriptare;
* Frontend: Thymeleaf, JavaScript, CSS, HTML;
* AI: KNN, Python, sklearn, pandas;
* Aplicații software: IntelliJ IDEA, PyCharm, Visual Studio Code, Postman, MongoDBCompass, DBeaver.

În urma unui studiu de piață, s-a constatat că, în ciuda faptului că există aplicații care încearcă să recomande oameni după anumite criterii, acestea nu realizează în totalitate obiectivul dorit. Multe dintre aplicațiile găsite pe Play Store includ opțiunea de alegere de preferințe (de exemplu, mâncare, hobby-uri, muzică), însă filtrele de utilizatori sunt inexistente, neputând primi utilizatori similari nouă, cel puțin din punctul de vedere al preferințelor. Singura aplicație care s-a constatat că ar face o parte dintre aceste funcționalități de recomandare este Meetup, dar acolo au loc recomandări de evenimente, nu de persoane. Panion ar fi fost un exemplu bun, dar acum nu mai funcționează publicului larg.

Pe baza acestui studiu de piață, s-a ajuns la concluzia că o astfel de aplicație ar fi necesară pieței. Scopul principal este acela de a găsi și de a filtra cât mai mulți utilizatori potriviți cu ținta în cauză, dar și de a avea un produs funcțional, sigur și securizat, care să fie ușor de folosit, plăcut și de înțeles de oricine.

Tocmai de aceea, s-au folosit diverși alogirtmi de recomandare. Câțiva dintre acești algoritmi ce pot fi folosiți în acest sens sunt: KNN, SVM, Random Forest, filtrul colaborativ, filtrul bazat pe conținut. Până în momentul de față, s-a realizat o comparare între câteva dintre metricile algoritmului KNN, acestea fiind Jaccard, euclidiană și cosine. Dar pe lângă aceste comparări, s-au întocmit și niște rapoarte care ar sugera diverse situații de utilizare mai potrivite acelor metrici.

# Ease of Use

## Selecting a Template (Heading 2)

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the A4 paper size. If you are using US letter-sized paper, please close this file and download the Microsoft Word, Letter file.

## Maintaining the Integrity of the Specifications

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

# Prepare Your Paper Before Styling

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

## Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

## Units

* Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
* Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
* Do not mix complete spellings and abbreviations of units: “Wb/m2” or “webers per square meter”, not “webers/m2”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.

Identify applicable funding agency here. If none, delete this text box.

* Use a zero before decimal points: “0.25”, not “.25”. Use “cm3”, not “cc”. (*bullet list*)

## Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop. To make your equations more compact, you may use the solidus ( / ), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

*a**b* 

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

## Some Common Mistakes

* The word “data” is plural, not singular.
* The subscript for the permeability of vacuum **0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
* In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
* A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
* Do not use the word “essentially” to mean “approximately” or “effectively”.
* In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
* Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
* Do not confuse “imply” and “infer”.
* The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
* There is no period after the “et” in the Latin abbreviation “et al.”.
* The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [7].

# Using the Template

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Authors and Affiliations

**The template is designed for, but not limited to, six authors.** A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

### For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

### For papers with less than six authors: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

#### Deletion: Delete the author and affiliation lines for the extra authors.

## Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named “Heading 1”, “Heading 2”, “Heading 3”, and “Heading 4” are prescribed.

## Figures and Tables

#### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

1. Table Type Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

1. Sample of a Table footnote. (*Table footnote*)
2. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

##### Acknowledgment *(Heading 5)*

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

##### References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*
2. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
3. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
4. K. Elissa, “Title of paper if known,” unpublished.
5. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
6. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
7. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**