The Seven Layers of Complexity of Recommender Systems for Children in Educational Contexts

Emiliana Murgia
emiliana.murgia@unimib.it
Università degli Studi di Milano-Bicocca
Milano, Italy

Monica Landoni
monica.landoni@usi.ch
Università della Svizzera Italiana
Lugano, Switzerland

Theo Huibers
t.w.c.huibers@utwente.nl
University of Twente
Enschede, The Netherlands

Jerry Alan Fails
jerryfails@boisestate.edu
Boise State University
Boise, Idaho

Maria Soledad Pera
solepera@boisetstate.edu
PIReT - Dept. of Computer Science
Boise State University, Boise, Idaho

ABSTRACT

Recommender systems (**RS**) in their majority focus on an average target user: adults. We argue that for non-traditional populations in specific contexts, the task is not as straightforward–we must look beyond existing recommendation algorithms, premises for interface design, and standard evaluation metrics and frameworks. We explore the complexity of **RS** in an *educational context* for which *young children* are the target audience. The aim of this position paper is to spell out, label, and organize the specific layers of complexity observed in this context.

CCS CONCEPTS

• **Social and professional topics** → **Children**; • **Applied computing** → **Education**; • **Information systems** → *Recommender systems.*

KEYWORDS

children, recommender systems, education, roles, guidance, interface, algorithm, teachers

1 INTRODUCTION

In general, the recommendation process is a complex one, as it does not occur in isolation [5]. Algorithmic design and evaluation demand that multiple dimensions coexist, if effectiveness is to be achieved from the perspectives of users, systems, and companies/organizations deploying the systems. We argue that focusing on non-traditional populations in specific contexts makes the process even more complex as it must look beyond existing algorithms, premises for interface design, and standard evaluation frameworks.

Using diverse lenses (from industry versus academia to visions from researchers in education, information retrieval and human-computer interaction, to name a few), we explore and discuss our views on the extended complexity of recommender systems (**RS**) that are used in an **educational context** with **young children** as the main users. The aim of this position paper is to spell out the many elements and facets of the recommendation process that contribute towards the complexity and richness of the design space for the production of **RS** that target children in an educational context.

We take a user-centred approach that puts children at the core of the experience. Thus, consider different roles children can play when they face the need for information in a learning context, as well as the level of guidance each of these requires from teachers. **RS** design must then match the needs and preferences defined by these roles, including the differences in teaching and learning practices for children and teachers (as the expert in the loop). We pay particular attention to the interface design, as it enables and supports interactions between the **RS** and the main actors in this process, and therefore directly impacts user engagement and influences perception of what makes the **RS** visible, trustworthy, as much as useful, usable and used within educational environments. We also consider the other various stakeholders involved in this educational scenario: parents, publishers, content and recommendation engine providers, classmates. Naturally, the need for criteria that can summarize **RS** performance emerges as another layer worth exploring, as assessment metrics need to encapsulate the various perspectives, goals, and motivations of the stakeholders involved.

2 RELATED WORK

We offer a brief overview of **RS** design and evaluation, as a starting point for us to compare and contrast with respect to the complexity layers later outlined for our domain of interest. We would be remiss if we did not mention existing literature in the context of education and children, as it serves as foundation to understand the gaps in the area of **RS** for children in a educational setting, along with the manifold requirements needed to fill those gaps.

Background on RS. There are a number of data sources **RS** used to produce either a list of ranked suggestions (top-N) or score for a given item (predictive). Among these data sources we find (i) user profiles and contextual parameters for personalization, (ii) community data, to identify popular and/or similar preference patterns–with respect to other users, (iii) product metadata, to connect to other items that exhibit similar traits to those already favoured by a user, (iv) knowledge models, to infer needs applicable to a particular domain. Moreover, on an algorithmic level, **RS** are evaluated using metrics like RMSE (predictive) or nDCG (top-N), as well as novelty, serendipity, or user satisfaction, to name a few. The majority of the

ComplexRec 2019, 20 September 2019, Copenhagen, Denmark

assessment is conducted off-line (depending on the availability of existing benchmarks) or via (short-term) user studies.

From literature exploration it emerges that solutions to problems in **RS** focus only on one specific part. For example, they are either concerned with satisfying multiple stakeholders [6], explaining suggestions [33], promoting user engagement [29], or ensuring content suitability for a given domain [28]. We question the applicability of existing solutions to simultaneously address the needs of the different stakeholders (along with their goals and expectations) involved in context of educational **RS** for children.

Technology-enhanced Learning. The body of literature in the area of technology-enhanced learning is rich [24, 32]. Unfortunately, when specifically looking at educational **RS**, Bodily and Verbert [4] state that literature is less prominent. Further analysis of the articles surveyed reveal that while their focus is on supporting learning, proposed strategies were rarely, if at all, evaluated with young users as target audience.

Children. There have been some attempts by the research community at large to take a look at challenges and needs of **RS** when the target audience are children [26]; as well as solutions to open problems in the area [8, 22, 28, 31, 34]. Similarly to what we previously discussed, solutions address a particular problem (e.g., query suggestions, aesthetic relevance, music preference patterns) and a single stakeholder (i.e., a child). For this reason, while serving as foundation for understanding the context of **RS** for children, they do not fully address the complexities we foresee for this area.

3 THE PROPOSED COMPLEXITY LAYERS

Building upon a literature exploration (Section 2), along with our experience on information retrieval systems for which children are the target audience [20, 21], in the following subsections we outline *seven layers* of *complexity* that impact **RS** for children in an educational setting.

3.1 The child as the protagonist

Collaborative design is a popular approach when producing interactions for children. It is important to note, however, the paucity of research in the child-computer interaction (CCI) community in dealing with this in the area of educational **RS** except for the recent introduction of the KidRec workshops in 2018 and 2019 [16, 26]. One possible reason for this is the complexity in modeling both users and tasks **RS** should be designed to support. Each child is unique in terms of skills, needs, personalities, and attitude towards learning. With that in mind, one promising starting point to get to know this user group would be to revisit and adapt to the educational setting, a typology of seven different roles children play in the search process [9]. These roles include: developing, domain specific, power user, non-motivated, distracted, visual, and rule-bound users. With the caveat that the same child could play different roles in different circumstances.

We acknowledge that these roles account for a number of heterogeneous, intertwined factors, each of which closely impacts the others. They can also help inform the design of **RS** that serve children in each role who vary from personality traits and cognitive development to the level of engagement and the degree of interest generated by the task. The amount of experience related to interactions with **RS**, or lack thereof, as well as the level of freedom and independence children experience when relying on suggestions for information discovery for the classroom, as opposed to the need for specific rules to scaffold their interactions with the **RS**, are important factors in the study of children's behaviour with respect to the use of **RS**. Personal preferences define the more visual user, mostly looking for non-textual information. Different levels of experience or familiarity with the recommendation process result in the developing versus the power user. While personality has an impact on the distracted user. The task the children want to accomplish also plays an important role as if it fits into the specific interests of the child user then the **RS** needs to respond to a domain-specific user behaviour otherwise it may encounter a non-motivated user. The rule-bound child user adheres more the influence and guidance provided by older and more experienced stakeholders, such as parents, siblings and teachers. In this sense the rule-bound user has an additional social dimension – an element of complexity specific of this young group of users. Personality and motivation of users together with the intricacy and nature of tasks are also elements accounted for in models of information seeking behaviour for adults.

In our user studies [20, 21], we observed how children's behaviour differs from that of adults. The same elements listed above have a stronger impact on young users' behaviour and cause more extreme reactions. For instance, children lacking experience often fail to engage with **RS**, as opposed to adults who are more used to simply resorting to other similar interaction experiences with **RS** in different contexts, e.g., Amazon, YouTube. Faced with a non-interesting task the child often decides not to engage with it at all, adults are more likely to take advantage of suggestions and deliver a result anyway. Finally, when presented with recommended items without explicit information about their source, children tend to assume the rigid rule-bound role and mistrust them.

3.2 The perspectives from multiple stakeholders

There is a broad spectrum of children that must be accounted for by the **RS** (see Section 3.1). There is also a large diversity of thought (some of which is highly opinionated) among the adults that directly or indirectly participate in the design, development, deployment, and adoption of **RS**. Over the past few years, researchers in **RS** have taken an interest and highlighted the complexity of involving multiple stakeholders into the recommendation process, which is a common occurrence in a real-world contexts for **RS** [1, 7].

In the education domain the literature is far less rich [6, 11, 15]; still there is a consensus on the complexity required to simultaneously maximize utility and meet the expectations of target users, with those of more satellite stakeholders. As discussed in Section 3.3, the role of the teacher remains crucial in transmitting confidence and trust in the **RS**, giving explanations on its potential use and on the way it works, which is why teachers, along with children, are the major stakeholders. Nonetheless, they are not the only ones: perspectives from content providers, parents, as well as organizations (non-profit/commercial) also influence and have an impact on this convoluted space of educational **RS** for children.

3.3 The concept of relevant recommendations that also foster learning

When children feel the need to discover information and have access to a device, usually they proceed by copying what the adults around them do. Children think of devices as “magical boxes that can do anything” which seems both amazing and a little bit scary. When it comes to using the device to retrieve information in an educational environment, the task is more complex: the students are asked to complete their assignments, in a reasonable amount of time, with no distractions and, of course, in a safe on- and off-line environment. The teachers, on the other hand, have to organize their interventions in order to give to all the students the possibility to achieve the given objectives, in the way that best match their personal way to learn. Many studies outline how the simple introduction of any kind of technology in a classroom is not sufficient to impact the level of learning [13]. The benefits become visible when the tool is specifically designed to take into count the complexity of the school environment, teachers are aware of the potentials of the technology and are able to organize the lessons allowing students to explore and learn the technology as well. A **RS** should meet the four dimensions of teaching outlined by Fadel [12] if its output is to be deemed relevant to the children (and teachers). A **RS** should:

(1) Facilitate and improve the acquisition of new knowledge
(2) Encourage the development of skills helping students to learn, be critical and thrive to widen their knowledge
(3) Support meta cognition by enabling reflections on the actions taken and their effects
(4) Boost the development of confidence in young learners by promoting the emotional side of learning (positive mood implies more effective learning experience) and reinforce children’s autonomy and self esteem

A **RS** that meets the aforementioned four dimensions becomes the “never without" technology in every classroom as it helps teachers in: (i) the personalization of the learning path, (ii) reduction of the time taken to find the correct information, (iii) avoiding or lowering the frustration from failing the task, and (iv) empowering children’s sense of self-consciousness and thus the ability to learn to learn. Accommodating these constraints into the notion of “relevance”, that given the context already must consider reading levels, alignment to curriculum, user interest, etc., translates into a complex component of **RS** for children in the educational setting.

3.4 The quest for interaction, engagement, and learning

From our experience [27], we see that children hardly use the recommended resources if they cannot trust them but also that they are enthusiastic adopters of new technological solutions if these provide enough guidance and challenge from them to engage with. The tension between feeling safe and at the same time going out to explore unknown territories is what makes it worth engaging with. Thus, in order to engage children, **RS** should propose materials that are not only relevant but at the right developmental level. This means that it should be understandable and provide some challenge to children as means to expand their current level of reading and support learning. The right level of challenge for each child is different and changes dynamically according to cognitive development, personal interest and experience, and the amount and type of guidance provided by teachers and parents. Besides, the various stakeholders pose different requirements on the **RS** to engage with. Delivering inspiring material for preparing classes would attract teachers, while providing support to children in topics they and their parents find hard (e.g. foreign languages, math and science) would entice parents. Therefore, as the experience of interacting with a **RS** has to be conducive to discovery and learning (as engagement is not always aligned with learning [23]) it is very complex to assess its performance too (see 3.7).

3.5 The need for explanations

For Nunes and Jannach [25] a “key requirement for the success and adoption of [recommender] systems is that users must trust system choices or even fully automated decisions." Explanations of the generated recommendations can help fulfill this requirement [33], yet, it is complex problem on its own [18], as it involves leveraging diverse data sources and perspectives of what will make them useful for the users. When it comes to educational **RS** for children, the transparency afforded by explanations paired with the recommendations becomes a must. At the same time, to be of use, explanations should appeal to the different stakeholders involved. For example, from a child’s perspective, preliminary studies show that kids must know at least the sources of the suggestions they are offered (e.g., peers or teachers) it they are to be of value; explanations could also support the exploration of the presented suggestions[35]. From an educator’s perspective (e.g., a parent or teacher) explanations became useful for understanding the rationale behind recommendations (e.g., commercial bias), the alignment to the educational context (e.g., prioritize learning), and the suitability to children’s needs (e.g., readability or appropriateness levels), to name a few.

3.6 The importance of ethics

The undesired and unpredictable behaviour of algorithms in **RS** is currently an ethical and social concern. These concerns apply even more to children than they do to adults. Algorithms in educational **RS** affect a child’s education. Unreliable, unreadable or irrelevant recommended information will directly harm the child [19]. The use of children’s profile and other data must also be dealt with very correctly on ethical grounds. Children need special protection when collecting and processing their data because they may be less aware of the risks involved. For young children, the **RS** needs to get consent from whoever holds parental responsibility for the child. This is not easy as they are not using the **RS**, and even if the parents are asked, it is complex to make the right consent decisions.

Finally, it is essential that children’s rights [2] are secured in the design process of the **RS** including their interactive needs, data safety, disclosure, and privacy. For example the involvement of children’s perspective in the design of the interface is of paramount importance (see Sections 3.1, 3.4). If adults are the major voices driving **RS** that target them, then equally children should play an active role in the design of tools for them to use.

3.7 The challenges with assessment

Lastly, attention must be paid to determining the degree to which educational **RS** for children are "good". As stated by Huibers et al. [16] "It seems to be that children are more interested in content that is interesting, amusing or informative, and not just results that are precise and relevant". For this reason, simply relying on top-N or predictive metrics will not be sufficient, and, in some cases, it will not be even possible [10, 14]. For example, it has been reported that children do not take advantage of the full-spectrum of the Likert scale for rating purposes, as such, RMSE is not applicable, as the penalization for incorrect predictions would be negligent if ratings are in their majority 4s and 5s [14]. Furthermore, traditional evaluation measures characterize algorithmic performance (see Sections 2 and 3.3), but overlook requirements from other stakeholders and the context itself in the portrayal of system performance that truly reflects if a given recommender algorithm lives up to the complex requirements of the proposed task. In fact, an educational **RS** for children cannot be evaluated only in terms of immediate gratification, measurable with fun, but by the satisfaction of children having somehow suffered and acquired a new competence. And for that it requires expensive longitudinal studies and interdisciplinary teams of evaluators. On its own, evaluation is a hard problem for **RS** research in general, when combined with the other layers discussed, it only helps shed a brighter light on the complexity of building meaningful and useful **RS** for children in an educational setting.

4 NEXT STEPS

We introduced seven layers of complexity that can be considered when looking at educational **RS** for children. We also discussed why each layer contributes to making this particular scenario complex. While both industry and academia have attempted to address some of the open problems presented [3, 17, 29, 30], none of proposed solutions simultaneously tackles all the layers outlined. Next steps in this area include exploring layers proposed (considering there might be others yet to emerge) and identifying the necessary course of action to approach in the quest for an ideal **RS** for children in an educational setting. We argue that this will naturally require multidisciplinary collaboration environments that accommodate academia vs. industry perspectives, as well as researchers on diverse, yet necessary and complementary, areas of study, including child development, psychology, education, edtech, literacy development, human-computer interaction, information retrieval, and graphic design.

ACKNOWLEDGMENTS

Work partially supported by National Science Foundation, awards 1565937 and 1763649.

REFERENCES

[1] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. 2017. Recommender systems as multistakeholder environments. In *Procedings of the 25th Conference on User Modeling, Adaptation and Personalization*. ACM, 347–348.

[2] D4CR Association. 2019. Designing for Children's Rights - Home Page. http://designingforchildrensrights.org/

[3] Alejandro Baldominos and David Quintana. 2019. Data-Driven Interaction Review of an Ed-Tech Application. *Sensors* 19, 8 (2019), 1910.

[4] Robert Bodily and Katrien Verbert. 2017. Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies* 10, 4 (2017), 405–418.

[5] Toine Bogers, Marijn Koolen, Bamshad Mobasher, Alan Said, and Casper Petersen. 2018. 2nd workshop on recommendation in complex scenarios (complexrec 2018). In *Procedings of the 12th ACM Conference on Recommender Systems*. ACM, 510–511.

[6] Robin Burke and Himan Abdollahpouri. 2016. Educational recommendation with multiple stakeholders. In *2016 IEEE/WIC/ACM International Conference on Web Intelligence Workshops*. IEEE, 62–63.

[7] Robin D Burke, Himan Abdollahpouri, Bamshad Mobasher, and Trinadh Gupta. 2016. Towards Multi-Stakeholder Utility Evaluation of Recommender Systems. In *UMAP (Extended Proceedings)*.

[8] Yashar Deldjoo, Cristina Frà, Massimo Valla, Antonio Paladini, Davide Anghileri, Mustafa Anil Tuncil, Franca Garzotta, Paolo Cremonesi, et al. 2017. Enhancing children's experience with recommendation systems. In *Workshop on Children and Recommender Systems (KidRec'17)- RecSys 2017*.

[9] Allison Druin, Elizabeth Foss, Hilary Hutchinson, Evan Golub, and Leshell Hatley. 2010. Children's Roles Using Keyword Search Interfaces at Home. In *Procedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 413–422.

[10] Michael Ekstrand. 2017. Challenges in Evaluating Recommendations for Children. In International Workshop on Children & Recommender Systems. Available at: shorturl.at/osFV9.

[11] Michael D Ekstrand, Ion Madrazo Azpiazu, and Katherine Landau Wright. 2018. Retrieving and Recommending for the Classroom. *ComplexRec 2018* 6 (2018), 14.

[12] Rosemany Evans. 2017. Four Dimensions of Education: 21st Century Teaching & Learning. https://www.utschools.ca/blog/content/four-dimensions-education-21st-century-teaching-learning

[13] Andrea Garavaglia, Valentina Garzia, and Livia Petti. 2013. The integration of computers into the classroom as school equipment: a primary school case study. *Procedia-Social and Behavioral Sciences* 83 (2013), 323–327.

[14] Michael Green, Oghenemaro Anuyah, and Devan Karsann. 2019. Evaluating prediction-based recommenders for kids. In 3^{rd} KidRec Workshop co-located with ACM IDC 2019. Available at: shorturl.at/hjnBI.

[15] Takeshi Horiuchi, Meagan Rothschild, Ray Barrera, and Surya Gururajan. 2018. Designing a Personally Meaningful ABCmouse.com: Challenges and questions in an EdTech recommendation system. In 2^{nd} KidRec Workshop co-located with ACM IDC 2018. Available at: shorturl.at/duvP6.

[16] Theo Huibers, Natalia Kucirkova, Emiliana Murgia, Jerry Alan Fails, Monica Landoni, and Maria Soledad Pera. 2019. 3rd KidRec Workshop: What does good look like?. In *Procedings of the 18th ACM International Conference on Interaction Design and Children*. ACM, 681–688.

[17] Theo Huibers and Thijs Westerveld. 2019. Relevance and utility in an educational search environment. In 3^{rd} KidRec Workshop co-located with ACM IDC 2019. Available at: shorturl.at/nAKNQ.

[18] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2019. Personalized explanations for hybrid recommender systems. In *Procedings of the 24th International Conference on Intelligent User Interfaces*. ACM, 379–390.

[19] Natalia Kucirkova, Jerry Alan Fails, Maria Soledad Pera, and Theo Huibers. 2018. *Children and Search/Recommendations algorithms: What Adults Need to Know*. DigilitEY.

[20] Monica Landoni, Davide Matteri, Emiliana Murgia, Theo Huibers, and Maria Soledad Pera. 2019. Sonny, Cerca! Evaluating the Impact of Using a Vocal Assistant to Search at School. In *International Conference of the Cross-Language Evaluation Forum for European Languages*. Springer, To appear.

[21] Monica Landoni, Emiliana Murgia, Theo Huibers, and Maria Soledad Pera. 2019. My Name is Sonny, How May I help You Searching for Information?. In *KidRec '19: Workshop in International and Interdisciplinary Perspectives on Children & Recommender and Information Retrieval Systems, Co-located with ACM IDC, June 15, 2019, Boise, ID*. 6 pages.

[22] Monica Landoni and Elisa Rubegni. 2017. Aesthetic Relevance When Selecting Multimedia Stories. In International Workshop on Children & Recommender Systems. Available at: shorturl.at/uyBIR.

[23] Derek Lomas, Kishan Patel, Jodi L. Forlizzi, and Kenneth R. Koedinger. 2013. Optimizing Challenge in an Educational Game Using Large-scale Design Experiments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 89–98. https://doi.org/10.1145/2470654.2470668 event-place: Paris, France.

[24] Nikos Manouselis, Hendrik Drachsler, Riina Vuorikari, Hans Hummel, and Rob Koper. 2011. Recommender systems in technology enhanced learning. In *Recommender systems handbook*. Springer, 387–415.

[25] Ingrid Nunes and Dietmar Jannach. 2017. A systematic review and taxonomy of explanations in decision support and recommender systems. *User Modeling and User-Adapted Interaction* 27, 3-5 (2017), 393–444.

[26] Maria Soledad Pera, Jerry Alan Fails, Mirko Gelsomini, and Franca Garzotto. 2018. Building community: Report on kidrec workshop on children and recommender systems at RecSys 2017. In *ACM SIGIR Forum*, Vol. 52. ACM, 153–161.

[27] Maria Soledad Pera, Emiliana Murgia, Monica Landoni, and Theo Huibers. 2019. With a little help from myfriends: Use of recommendations at school. In *Proceedings of ACM RecSys 2019 Late-breaking Results*. 5 pages.

[28] Maria Soledad Pera and Yiu-Kai Ng. 2014. Automating readers' advisory to make book recommendations for k-12 readers. In *Procedings of the 8th ACM Conference on Recommender systems*. ACM, 9–16.

[29] Meagan Rothschild, Takeshi Horiuchi, and Marie Maxey. 2019. Evaluating "Just Right" in EdTech recommendation. In 3^{rd} KidRec Workshop co-located with ACM IDC 2019. Available at: shorturl.at/aGLRZ.

[30] Almudena Ruiz-Iniesta, Luis Melgar, Alejandro Baldominos, and David Quintana. 2018. Improving Children's Experience on a Mobile EdTech Platform through a Recommender System. *Mobile Information Systems* 2018 (2018).

[31] Markus Schedl and Christine Bauer. 2017. Online Music Listening Culture of Kids and Adolescents: Listening Analysis and Music Recommendation Tailored to the Young. In International Workshop on Children & Recommender Systems. Available at: shorturl.at/emP17.

[32] Avi Segal, Ziv Katzir, Ya'akov (kobi Gal, Guy Shani, and Bracha Shapira. 2014. EduRank: A Collaborative Filtering Approach to Personalization in E-learning. In *Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014)*. London, UK, 68–75.

[33] Nava Tintarev and Judith Masthoff. 2011. Designing and evaluating explanations for recommender systems. In *Recommender systems handbook*. Springer, 479–510.

[34] Sergio Duarte Torres, Djoerd Hiemstra, Ingmar Weber, and Pavel Serdyukov. 2014. Query recommendation in the information domain of children. *Journal of the Association for Information Science and Technology* 65, 7 (2014), 1368–1384.

[35] Katrien Verbert, Denis Parra, Peter Brusilovsky, and Erik Duval. 2013. Visualizing recommendations to support exploration, transparency and controllability. In *Procedings of the 2013 international conference on Intelligent user interfaces*. ACM, 351–362.