

Horizon Europe



Skills and needs data and analysis

D7.5 - Policy Brief 1

Document Control

Information:

Settings	Value
Deliverable No.	D7.5
Document Title:	Policy Brief 1: Skills and needs data and analysis
Author(s):	Raquel Amaro & Ana Leite Policy Brief Leaflet - Mindshaker
Reviewer(s):	Angels Catena; Regina Duarte
Sensitivity:	Public
Date:	23/07/2024

Document Location: The latest version of this controlled document is stored in OneDrive-
fcs.h.unl.pt/iRead4Skills/Project/Work Packages/WP7/Policy Briefs

TABLE OF CONTENTS

IREAD4SKILLS PROJECT SUMMARY 3

1. INTRODUCTION 5

2. EVIDENCE AND ANALYSIS 7

 2.1 DATA ON READING NEEDS AND SKILLS 7

 2.2 COMPLEXITY FRAMEWORK 10

 2.3 DATA SETS: CORPORA AND LEXICONS 11

3. POLICY IMPLICATIONS AND RECOMMENDATIONS 12

4. PROJECT IDENTITY 14

5. REFERENCES 15

6. POLICY BRIEF LEAFLET VERSION 16

iRead4Skills Project Summary

Reading skills are fundamental to acquire technical and scientific knowledge, either in formal education and training contexts, as Adult Learning (AL) and Vocational Educational Training (VET), or in practical and empirical working contexts, as when companies provide written specific information and/or instructions to their workers. People with low literacy skills are less able to acquire and sustain transversal and durable skills needed to stay apace with the changing job market and to lead meaningful and complete lives.

Responding to the need to create the conditions for the successful development of skills matched to the functional needs of adult population, and contributing directly to many of the issues established in the *European Skills Agenda for Sustainable Competitiveness, Social Fairness and Resilience* (EU 2021), the project **iRead4Skills - Intelligent Reading Improvement System for Fundamental and Transversal Skills Development** has as main objectives:

1. To assess and reduce reading skills gaps in low literacy adults for the development of fundamental and transversal skills by:

i) establishing and monitoring the relation between improvement in literacy skills and improvement in other fundamental skills (such as numeracy or ICT) and transversal skills (such as technical skills, communication skills, teamworking skills, foreign language skills, customer handling skills, problem solving skills, learning skills and planning and organization skills) as well as motivation, citizenship, and well-being.

ii) providing an intelligent reading system that evaluates texts complexity and suggests appropriate readings according to the user literacy level. The system would also be used to assist trainers and text producers to create or adapt texts with the appropriate level of complexity for their target readers.

2. To support the adoption and diffusion of innovation in the way low literacy adults acquire and sustain transversal and durable skills, matching the always evolving needs of the job market, by developing and putting to work an open access system, available for both trainers and trainees, that allows for advanced digitally enabled training and skills development.

Partners

No	Participant organisation name	Country
1	NOVA University Lisbon	Portugal
2	Catholic University of Louvain	Belgium
3	Luxembourg Institute of Socio-Economic Research	Luxembourg
4	Ministry of Education	Portugal
5	Universitat Autònoma de Barcelona	Spain
6	Mindshaker Serviços Informáticos, Lda	Portugal
7	INESC ID - Engenharia de Sistemas e Computadores: I&D em Lisboa	Portugal
8	University of Santiago de Compostela	Spain

Advisory Board

Anne Mességué, National Coordinator of ANLCL, in France

Cécilia Märki, Head of Basic Skills Department from SVEB FSEA, in Switzerland

Luís Rothes, Coordinator of the PIAAC, OCDE, in Portugal

Zoltán Várkonyi, Secretary General of the European Basic Skills Network, in Hungary

Ethics Board

Dagmar Monett Díaz, Prof. in AI and Software Engineering, Head of Computer Science Dept. Berlin School of Economics and Law, in Germany

João Queirós, Subcoordinator of PIAAC, OCDE, in Portugal

Ulises Cortés, Prof. in AI, scientific coordinator of High-Performance Artificial Intelligence group, Universitat Politècnica de Catalunya/Barcelona Supercomputing Center, in Spain

This document includes an Appendix with a leaflet version.

I. Introduction

Reading skills are essential to acquire technical and scientific knowledge. This is especially relevant in the context of formal education and training contexts, as Adult Learning (AL) and Vocational Educational Training (VET), and in practical and empirical work contexts, such as when companies provide information and/or specific written instructions to their workers. However, promoting and motivating reading habits and skills in adults is quite challenging due to the lack of dedicated and/or adequate reading materials.

Supporting the adoption and diffusion of innovation in adult training, the iRead4Skills project aims to promote the development of reading skills through an innovative intelligent system that evaluates texts complexity and suggests reading materials adequate to the user reading level, which can also be used by trainers in the creation or adaptation of texts with the appropriate level of complexity for their individual students.

Literacy: oral reading fluency; reading comprehension; spelling; writing.

Reading Skills: reading fluency; reading comprehension.

This means contributing:

- to **upskilling and reskilling of people**,
- to developing their **flexibility and adapting abilities**, including staying abreast with new skills related to technology, but also
- to **new educational strategies** related to formal education, responding to individual and social needs, and boosting adults' motivation and participation, and
- to **innovative and inclusive education systems**, using digital technologies to provide quality training.

The goals of this project focus on reading skills and are attuned with creating the conditions for the successful development of skills matched to needs, responding directly to many of the issues established in the European Skills Agenda for Sustainable Competitiveness, Social Fairness and Resilience (EU 2021) and contributing to several European Commission's policy priorities, such as *A Europe fit for the digital age*, by empowering people and educational systems with innovative and inclusive technologies, *An economy that works for people*, by contributing to the European Skills Agenda, or *A new push for European democracy*, by providing people with skills that allow them to be more informed, to deal with misinformation and to be more active and participating citizens.

“Literacy and numeracy skills form a Foundation for developing higher-order cognitive skills such as analytic reasoning and are essential for accessing and understanding specific domains of knowledge. In addition, they are relevant across a range of life contexts, from education and work to home, social life, and interaction with public authorities.” (OECD 2019:36)

By combining an interdisciplinarity team covering different fields of expertise such as ICT, Linguistics, Economics and Education, the iRead4Skills project will provide groundbreaking research and innovation, going beyond the current state of the art in several key areas:

Assessing the response to specific needs from the AL and VET communities involved and devising ways to evaluate results to inform proposals for employers and policy makers.

The training difficulties of adult learners can range from basic reading skills to motivational factors. The development of the iRead4Skills system goes beyond existing studies and tools due to its relational and cooperative nature. Based on input from end-users and covering their real needs, the project will provide a new way to assess the response of individuals, skill development institutions, and employers to issues on reading skills and to gather data to inform policy makers on the conditions for the successful development of skills. At this stage of the project, we have reached out and secured the involvement and cooperation of both trainees and trainers in AL and VET and compile some interesting data on reading skills needs and preferences.

Developing new complexity measures and base data sets

iRead4Skills solution goes beyond existing systems by developing and testing a language-designed system based on the direct input from AL and VET trainers and learners. This means that actual end-users will provide the base classification of texts according to their sensitivity to complexity, and that new complexity levels and features are to be researched from these new and valuable data sets. This also means new complexity measures based not on proficiency levels of a second language, but on issues of complexity for native speakers, which can be drastically different. At this stage of the project, we have already defined the complexity framework relevant for our target audience and purposes, and the classification and annotation of texts by the end-users in Portugal, France and Belgium and Spain.

Testing and developing Natural Language Processing systems directly with end-users

iRead4Skills innovates the current paradigm for developing machine-learning systems, devised, and tested only on data sets. Current systems, produced in academia, are not developed nor tested in real case scenarios, where end-users evaluate their usability and utility. In the iRead4Skills project, we will go the extra mile, and test the system directly with specific end-users (i.e., low literacy adults). This will allow us to devise an intelligent tool, attuned to the user, but also to gather solid evidence to assess its impact.

Promoting the use of Art and Culture as a means to develop fundamental and transversal skills

By including reading materials from a wide range of domains (literature, history/social studies, science, etc.), iRead4Skills provides a foundation of knowledge that will help adult learners to be better readers in all domains. In addition to professional and scientific training, the project will contribute to raise interest in humanistic, scientific, and technological knowledge of adult population, to bring adult readers closer to literature, culture, and general information; and to improve communication and educational practices for all.

All in all, project aims at **all people with lower literacy skills interested in reading**, and at **people/entities/organizations promoting reading**. AL and VET centres in Belgium, France, Portugal, and Spain, including trainers and trainees can be identified as the iRead4Skills direct beneficiaries.

The project primary stakeholders for exploitation purposes are national and local entities interested in the promotion of reading habits/culture dissemination and consumption, such as

- i) governmental ministries, general directorates of education, etc., to enforce more effective teaching/learning methods and improve the job market and society, but, also,
- ii) culture actors, such as libraries, publishers, and content producers - to augment their audiences/customers.

To reach these audiences, the iRead4Skills project has already established contacts and cooperation protocols with AL and VET centres and professionals in the relevant geopolitical intervention areas, namely, Belgium, France, Portugal, and Spain, having also reached AL centres in Argentina. We have also secured cooperation bridges with relevant stakeholders dedicated to skills, adult training, and reading and literacy promotion, such as ANQEP - *National Agency for Qualification and Professional Training*, Portugal, ESNB - *European Basic Skills Network*, SVEB - *Swiss Federation for Adult Learning*, Switzerland, ANLCL - *Agence nationale de lutte contre l'illettrisme*, France, SPRP - *Savoirs pour Réussir*, France, *Ille & Vilaine*, France, *Xunta de Galicia - Consellería de Cultura, Educación, Formación Profesional e Universidades*, Spain, *Fundació Gentis*, Spain, *Federación de Asociaciones de Educación de personas Adultas*, Spain, FECEI - *Federación Española de Centros de Enseñanza de Idiomas*, Spain.

Besides the publication of objective results concerning the relation between the improvement of reading skills and other fundamental and transversal skills, the project also expects the production of new fundamental

knowledge on complexity analysis, within the different key areas coverage by the project (Linguistics, Education, Natural Language Processing, Socioeconomics), usable in many other contexts.

The major and more tangible result is, however, the development of the **iRead4Skills system, a new ICT system, open access web-based system**, able to provide:

- i) **complexity analysis** of texts,
- ii) **reading suggestions** according to their complexity level,
- iii) **assistance in the simplification** of texts.

It is important to underline the project's singularity that emerges from this section. On the one hand, it benefits from an interdisciplinarity team covering different fields of expertise such as ICT, Linguistics, Economics and Education. On the other hand, it is based on input from end-users (i.e., low literacy adults) which meets their real needs and at the same time makes possible an innovative approach for developing and testing machine-learning systems. Finally, unlike in other resources, iRead4Skills is based on issues of complexity for native speakers and not on proficiency levels of a second language.

The successful achievement of iRead4Skills project contributes in a significant scale to enable policymakers to better understand, measure and reduce skills gaps, resulting in policy recommendations towards learner centred flexibility in AL and VET contents and training materials, and towards a generalized use of these type of systems in cooperation with the culture and science actors. Using and disseminating information on the complexity (or its lack thereof) of existing and new literature and other reading materials in libraries, book covers, newspapers, etc., would immediately open a whole new world of information for a still excluded part of the population.

2. Evidence and Analysis

This Policy Brief provides a brief overview of the work developed so far, concerning data on reading needs and skills, and on the relation between reading skills gaps and socio-economic, employment and personal conditions, the advancement of the complexity framework directed to Adult Population with low literacy skills, and the collection of datasets, as well as the lessons learned in this cooperative process.

Two specific surveys were designed and disseminated to collect the data. The first survey focuses on reading skills and adult learning and was carried out among trainers and trainees. It reported skills in the categories Home, Transport, Shopping and Healthcare and revealed that respondents like to read about current issues of society, health, and wellness. The second survey focuses on how reading difficulties impact access to employment and individual well-being. Even though there is a high number of missing observations for certain questions, it is important to underline that this second survey allows to observe the impact that reading training has had on people's literacy skills, their other skills and professional situation.

2.1 Data on reading needs and skills

To collect the relevant data in a bottom-up approach, we have designed and disseminated two specific surveys:

I. Reading Skills survey, a qualitative and a quantitative survey on people who are following AL and VET courses or that act as trainers, designed with the participation of the in-the-field partners and others with knowledge on reading skills and adult learning, to assure that the survey covers the specific information required for the subsequent work packages (i.e., text genres, domains, cultural and scientific expectations,

etc.). The survey collected information from adults, trainers/teachers and trainees/students in AL and VET centres in Portugal, Belgium, France, and Spain.

Considering that the target group include peoples with low reading skills, the survey included an oral version of all the contents, so people could provide informed consent concerning the goals of the survey, the data collected, and the data protection measures in place and could listen to the questions and select the appropriate answers. Equivalent versions of the survey in French, Portuguese, and Spanish were devised.

A summary of the most relevant results is presented below:

Overall, **more women responded** to the survey. Most of the respondents stated that they **read to learn new things** and, after that, to learn about the news. In general, respondents like to read about **current issues of society**, and **health and wellness**. Considering the reading skills and needs in the categories *Home*, *Transport*, *Shopping* and *Healthcare*, the respondents reported that:

- (i) **Home**: they **do not read** or they **read and understand everything** in home-related texts (**adverts in newspapers, magazines or on the internet, information or adverts on posters, lease agreement and condominium regulations**).
- (ii) **Transport**: they generally **read and overall comprehend timetables, destinations and stops on posters or billboards**, but do **not read or fully understand information about strikes and disruptions** in transport information.
- (iii) **Shopping**: they **read and understand** (a little or a lot) the information presented in **products packaging and labelling** when shopping, and **product and restaurant prices on posters, flyers, leaflets or catalogues**.
- (iv) **Health**: they **read partly or fully understood** and texts related to healthcare issues (**medical prescriptions, medicine leaflets and packaging, information about health services**, etc.).

Full information on the survey, as well as the three full surveys openly accessible here: <https://zenodo.org/records/10179536>.

II. Overall skills and gaps survey, aiming at studying how reading difficulties impact access to employment and individual well-being in adults who have undergone reading education in AL and VET centres, including aspects such as personal confidence, acquisition of other skills (mathematics, computer literacy, soft skills, etc). These impacts were examined from both a subjective perspective, through individuals' perceptions of the consequences of their reading difficulties on various aspects of their lives, and from an objective perspective, through changes in personal circumstances enabled by participating in reading education (human capital, employment status, etc.).

The same survey was available in three languages, Spanish, French, and Portuguese. Respondents were given the option to answer the survey in multiple sittings. Participation in the survey was voluntary and respondents are informed of the survey's objective and their rights regarding data protection.

The questionnaire consisted of the following six modules:

- **sociodemographic characteristics**: to characterize the respondent and, ultimately, to study whether the consequences of reading difficulties vary based on sociodemographic characteristics. Respondents are specifically asked to indicate their gender, year, and country of birth, as well as their occupational status.
- **education and reading learning**: to identify if the respondent's reading difficulties stem from their educational background. Understanding the origin of reading difficulties is crucial since they can influence the impact these difficulties have on other areas of life. In this module, respondents are specifically asked

about the country where they received their education, whether they faced academic challenges, and in which languages they learned to read.

- **reading education:** characterize the reading education received by the respondents from training centres, to understand why the respondents underwent such training and whether they completed it or not. These questions are essential because the way the learning was conducted determines the benefits the respondent gained from it.
- **reading proficiency:** to identify whether the training received enabled the respondents to improve their reading abilities. Respondents are asked to assess their reading skills both before and after undergoing the training. Reading proficiency is evaluated in the context of everyday life and, if applicable, in the context of professional life.
- **other skills and health:** to characterize the respondents' other skills (mathematics, computer literacy ...) and how their reading difficulties may have affected these skills. It also asks respondents about their health status and any learning disorders that may have caused difficulties in acquiring skills.
- **workplace skills:** to study the mismatch between the skills required in the workplace and the skills possessed by the respondent. It also allows for an examination of whether improved reading proficiency has enabled respondents to enhance their skills in other aspects of their professional life (technical skills, planning, communication, etc.).

More information on the survey structure is openly accessible here: <https://zenodo.org/records/10181135>.

The surveys were widely disseminated with the support of all partners and peer institutions that work directly with AL and VET centres and announced in the project website (<https://iread4skills.com/activities/#surveys>; <https://iread4skills.com/fr/activites/>; <https://iread4skills.com/es/actividades/>; <https://iread4skills.com/ptpt/atividades/>), as well as in the project social media (Facebook, LinkedIn, X, YouTube). The dissemination of the surveys was also supported by dedicated media (see <https://www.youtube.com/watch?v=3e-qWMvG5rM>; <https://www.youtube.com/watch?v=dy94cBCxhoc>; and https://www.youtube.com/watch?v=Xz5c3_hRC2c).

Together with a literature review, the results from the second survey were further used to study the link between literacy skills, on the one hand, and other skills and working life, on the other hand. This study allows to observe the impact that reading training has had on people's literacy skills, their other skills and professional situation. The survey of individuals who have undergone training to enhance their reading skills revealed the challenges of conducting research on this group. Apart from Portugal, the response rate was extremely low. There is a high number of missing observations for certain questions due to respondents choosing not to answer or being unable to provide an opinion. This limitation restricts the possibility of utilizing certain questions and making comparisons based on the socio-demographic characteristics of the respondents. It is important to note that the results cannot be generalised to individuals who have taken reading courses in the countries studied due to the nonrepresentative sample. Nevertheless, the results enable us to draw conclusions and pinpoint some of the challenges faced by these individuals, as well identify new research directions, as briefly reported below:

- i) **Individuals who have received reading training tend to downplay their difficulties** in this area. In fact, when asked about their reading level prior to the training, most respondents did not mention any difficulties. For some, the training appeared to have made them aware of their shortcomings, as they reported facing more difficulties at the time of the survey than before the training. It seems that respondents adapt their reading needs based on their skill level. To verify this hypothesis, **a comparison of the reading needs of individuals undergoing training with those who have no difficulties in this area would be interesting.**

- ii) **Employees who encounter reading difficulties at work are often hesitant to seek assistance.** In fact, most of them do not ask for help when faced with a challenge.
- iii) **Respondents had mixed views on whether their reading difficulties caused other types of difficulties.** Half of the respondents reported no difficulties in the seven areas studied, including mathematics, computing, employment opportunities, career advancement, daily activities, personal confidence, and relationships with others. However, 15% felt that their reading difficulties negatively impacted these areas.
- iv) **Respondents aged under 40 reported encountering fewer difficulties than their older counterparts.** This may be due to unawareness of their difficulties in some areas of life.
- v) **Employed respondents agreed that their improved reading skills had a positive impact on other professional abilities,** such as teamwork.

The full literature review, and well as the results from the second survey are openly accessible in: <https://zenodo.org/records/10459030>.

2.2 Complexity framework

Considering the overall goals of the iRead4Skills project, the methodology for establishing the framework for the complexity levels addressed in the project considered three major steps:

- i) Compiling and evaluating existing tools and frameworks relevant for text readability and complexity assessment.
- ii) Determining a base proposal for three text complexity levels relevant for our target audience (adult native speakers with low reading skills) and for our goals (to promote the development of reading skills through an innovative intelligent system that evaluates texts complexity and suggests reading materials adequate to the user reading level, which can also be used by trainers in the creation or adaptation of texts with the appropriate level of complexity for their individual students).
- iii) Validating and making the necessary adjustments to the base proposal to inform a thorough and objective characterization of the text complexity levels addressed by the project.

In general terms, this meant gathering information to juxtapose the descriptors from PIAAC, ALTE and CEFR¹ proficiency levels into three levels that would reflect different reading complexity/difficulties for Spanish, French and Portuguese texts, basing the literacy skills on the literacy model of the first three levels from PIAAC (ranging from A1 to B1 CEFR levels). These are broadly characterised both internationally and cross-linguistically with norm-referenced descriptors that provide guidelines and bring reliability and comparability to teaching environments and assessment across different societal realities. The resulting set of descriptors were then presented for independent classification to trainers from AL and VET centres in Belgium, France, Portugal, and Spain. The results of this task were then discussed in focus groups to understand and adjust, whenever possible, conflicting judgements. **The result** is a framework for text analysis and classification, consisting of language-dependent tables with the characterization of the **complexity levels relevant for adult population with low literacy skills**.

¹ PIAAC - *Program for the International Assessment of Adult Competencies* (OECD 2013), ALTE - *Association of Language Testers in Europe* (ALTE 2020) and CEFR - *Common European Framework of Reference* (CEFR 2020).

For the iRead4Skills project, we decided on three text complexity levels – **Very Easy** (corresponding to A1), **Easy** (corresponding to A2), **Plain** (corresponding to B1).

The grading system was influenced by the CEFR's frame of reference (Council of Europe, 2020), namely in its labels (A1, A2, B1) and the naming system chosen considering two criteria: versatility and transparency. Versatility, since CEFR is the most widely used frame of reference for languages in Europe. Future users of our complexity analysis system will be able to easily understand and use the grading system, as well as adapting it to other grading systems. Transparency, considering simplicity and other already existing initiatives (e.g., Plain language). Given the potential target users of the iRead4Skills system, we consider that transparent and familiar names, characterizing the texts and not the readers, will contribute to the use of the system.

The reason for postulating these three levels of text complexity is also linked to the target audience of the project. These three levels of complexity were postulated considering that i) they were the relevant levels for developing and promoting reading skills and habits (e.g., starting in A1 and not pre-A1 or A0, which are closer to the alphabetization process), and ii) they were the ones most problematic for trainers, since reading materials of these levels are not easy to find.

The levels of complexity defined for the project can be functionally described as follows:

Very easy: Texts that are fully or almost fully understood by everyone, including people with very low schooling (i.e., that did not finish the primary school (ca. 6th year)) and almost no reading experience. It roughly corresponds to CEFR A1 level.

Easy: Texts that are fully or almost fully understood by people with low schooling (i.e., that completed the primary school but do not have more than the 9th year) and have poor reading experience. It roughly corresponds to CEFR A2 level.

Plain: Texts that are understood the first time they are read by people that completed the 9th year and have a functional-to-average reading experience. It roughly corresponds to CEFR B1 level.

The complete description of the iRead4Skills Complexity Levels and framework can be openly accessed here: <https://zenodo.org/records/10459090>.

2.3 Data sets: corpora and lexicons

- iRead4Skills Dataset 1: corpora by level of complexity for FR, PT and SP

The compilation of this first dataset of texts was based on the complexity levels established as relevant for the project and on the expected needs of learners and trainers, resulting from the surveys and close cooperation with AL and VET centres. This data set is a collection of written texts of several genres and levels of complexity, in txt format, and is composed of three sub corpora: French, Portuguese and Spanish. Each of the sub corpora considers different complexity levels and covers texts from the following communication domains, and respective genres and subtypes:

personal communication; institutional/professional communication; social media; commercial communication/dissemination; non-fiction book; fiction book; didactic book; academic/school; political communication/dissemination; legal documents; religious texts/dissemination.

This dataset provides the basis for the training and test sets for the complexity analysis systems for the three languages of the project: French, Portuguese, and Spanish. The dataset will be further enhanced, validated, and annotated by end-users, originating forthcoming versions and a second, derived, dataset.

Full access to the dataset is available here: <https://zenodo.org/records/10055909>.

- Lexicons per complexity level: baselines

As described in the literature from its beginning to date, and as considered in the descriptors of complexity in use for the iRead4Skills purposes, the words used in the texts are an essential factor of complexity. An automatic complexity analysis requires, thus, information on the lexicon used and/or expected at each level. However, complexity analysis targeting adult native speakers may require lexical resources that somewhat differ from existent resources related to lexical complexity directed to second language learners, as the passive knowledge and the needs from native speakers are expected to be different.

To define the relevant lexicons, the iRead4Skills project follows a mixed approach, as manually built resources are richer in terms of some specific information but tend to be shorter and more-time consuming to build, whereas data-driven and automatic inferred resources are quicker to get but require representative and balanced corpora as well as other structured source resources.

iRead4Skills benefits from a mixed approach considering the different available resources for each language and taking advantage from the expertise gathered in the iRead4Skills consortium, the compilation of the lexicons to be used in the project combines:

- corpus-driven methods (e.g., for Portuguese, to extract graded lexicons per CEFR levels and to extract core/fundamental general language vocabularies),
- automatic inference (e.g., for French using graded lexicons for CEFR levels and distributional data from iRead4Skills validated corpus),
- expert grading (e.g., for Spanish to define core vocabularies usable for complexity analysis and for automatic inference tools; for Portuguese, to validate and further adapt CEFR graded lexicons to native speakers' adult case).

The idea is to allow for the use of different data sources, of different analysis and processing systems, and well as of different pipelines to reach the best outcome. Graded corpora considered as input include the corpora specifically built for the iRead4Skills project (see D3.7 Data set 1: corpora by level of complexity FR, PT and SP). The resulting lexicons will also be subject to the target-audience indirect validation, as adult learners will participate in a classification task for validating the corpora compiled, that contemplates the option of annotating words and expressions perceived as complex.

Both the lexicons and the second, annotated dataset are currently being built. Further details on the Baselines for complexity lexicons definition established for the iRead4Skills project can be openly accessed here: <https://zenodo.org/records/10069793>.

3. Policy Implications and Recommendations

Following the iRead4Skills outlined paths towards achieving solid results, with the cooperation of direct beneficiaries and stakeholders, the work developed so far has also required engaging directly with people in AL and VET formal and informal training contexts. To reach these audiences, the iRead4Skills project has established contacts and cooperation with AL and VET centres, as well as individual professionals in social and education associations, located within the relevant geopolitical intervention areas, namely, Belgium, France, Portugal, and Spain. The policy implications and recommendations listed in this section reflect both the insights and remarks derived from the research activities developed, and the preliminary conclusions collected from

and informed by contact with end-users. Our goal is to contribute to European initiatives and interventions aimed at promoting reading skills and habits in the adult population, facilitating access to information, scientific and technical knowledge, and culture.

▪ Access to information - information in plain language

There is a generalized lack of information available in accessible written formats. This is clearly evidenced by the corpora collected and by the efforts spent to find such texts. Specifically, there is a significant shortage of texts with low complexity in functional areas, such as pricing information, bills, and services conditions, which are essential for people to make important decisions regarding issues like acquiring services, renting a house, shopping, etc. This indicates that, despite several plain language initiatives being active globally, institutions and companies are still not actively implementing them.

- ⇒ **Directives regarding the information that must be provided to the consumers and/or general population could also mention the need to provide this information in plain language.**

The developed complexity descriptors, along with the resources produced, such as the basic/fundamental lexicon expected at each level - extracted from authentic texts and validated by end-users - are crucial resources to help in the production of such texts.

▪ Motivation and adequation of AL and VET training

The materials specifically designed for language teaching are not adequate for the adult population with low literacy skills, as these materials do not cover: i) the topics and communication situations in which people are truly interested (e.g., healthcare, medical appointments); ii) the text genres and subtypes that people will actually need to read (e.g., medical leaflets and prescriptions, lease contracts).

- ⇒ **To facilitate access to texts and reading for adult learners in a sustained and informed manner, adult training should include and promote the use of authentic texts and consider real situations, including informal training contexts (such as leisure, household activities, etc.).**

The survey results inform about the perceived trends but can also be used to collect updated information, on different occasions, including local and specific classes.

Considering the 2030 targets set in the Commission's European Pillar of Social Rights Action Plan at the Porto Social Commitment by which at least 60% of all adults of Europe should participate in training every year.

The devised complexity descriptors, along with the resources produced, such as the basic/fundamental lexicon expected in each level - extracted from authentic texts and validated by end-users - are key resources to help in the production of such texts.

According to the latest European statistics on lifelong learning activities, in Europe, 43,7% of the adults aged 25-64 (ca. 125,5 million) participated in some form of education or training in 2016, 80% of which were job-related.

These numbers are even higher in most of the countries in Europe where the working languages of the iRead4Skills project are spoken: Belgium 45,2%; France 51,3% (61,5 million); Luxemburg 48,1% (70 million); Portugal 46,1%; Spain 43,4%; Switzerland 69,1%. This means that the iRead4Skills project can directly impact the lives of more than 50 million trainees, besides all involved trainers and teachers.

▪ Cooperative and collaborative approaches in research and knowledge development

It was particularly difficult to recruit and sustain stable and motivated cooperation from the end-users, especially trainers. Trainers expressed feeling overwhelmed by constantly being asked to participate in and provide information for different studies (e.g., PIAAC), with no real and direct effect on their work and without the proper recognition of their efforts.

- ⇒ **Provide attractive compensation for the segments of population asked to contribute to research and innovation. This could include allowing teachers/trainers to allocate some of their working time to participating in research and innovation projects or providing adequate monetary compensation for the time dedicated to the projects, without requiring the trainers' affiliation institution (if any) to be formally a beneficiary of the project.**

Recognizing that end-users/target beneficiaries and professionals in the field are essential parts of a cooperative, bottom-up approach to developing knowledge and innovation will enhance the expected impact of investments in research and innovation, as well as society's overall trust in and relationship with science.

4. Project Identity

Coordinator:	Raquel Amaro, PhD NOVA University Lisbon Lisbon, Portugal raquelamaro@fcsh.unl.pt
Funding programme:	HORIZON Europe
Duration:	March 2023 - February 2026 (36 months)
Website:	https://iread4skills.com/
Social Media:	https://www.facebook.com/iread4skills.com https://twitter.com/iRead4Skills https://www.linkedin.com/company/iread4skills/ https://www.youtube.com/@iRead4Skills

5. References

ALTE (2002). The ALTE CAN DO PROJECT(1992–2002). https://www.alte.org/can_do/alte_cando.pdf

Council of Europe (CE) (2020). Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume. Council of Europe Publishing, Strasbourg. <https://www.coe.int/langcefr>

European Union (EU) (2021). *European Skills Agenda for Sustainable Competitiveness, Social Fairness and Resilience*. European Commission. <https://ec.europa.eu/migrant-integration/sites/default/files/2020-07/SkillsAgenda.pdf>

OECD (2013). Technical report of the survey of adult skills (PIAAC). OECD Publishing. Paris.

OECD (2019). Skills Matter: Additional Results from the Survey of Adult Skills. OECD Publishing, Paris, <https://doi.org/10.1787/605ec8b3-en>

6. Policy Brief leaflet version

Policy Brief 1

Skills and needs data and analysis

The iRead4Skills project, funded by the European Commission, aims to reduce literacy gaps in low-literacy adults by promoting both fundamental and transversal skills.

Our objective is to **assess and improve reading skills**, thus enhancing other essential skills such as numeracy, ICT, communication, and problem-solving. By providing an intelligent reading system that suggests texts appropriate to users' literacy levels, we support the adoption and diffusion of innovative **adult training methods**.

Reading skills are crucial for acquiring technical and scientific knowledge, particularly in formal education, vocational training, and practical work contexts. However, promoting reading habits among adults is challenging due to the lack of dedicated and appropriate reading materials. Our project addresses this by developing a system that evaluates **text complexity** and suggests reading materials aligned with the user's literacy level. This system also aids trainers in **creating or adapting texts** with the right level of complexity for their students.

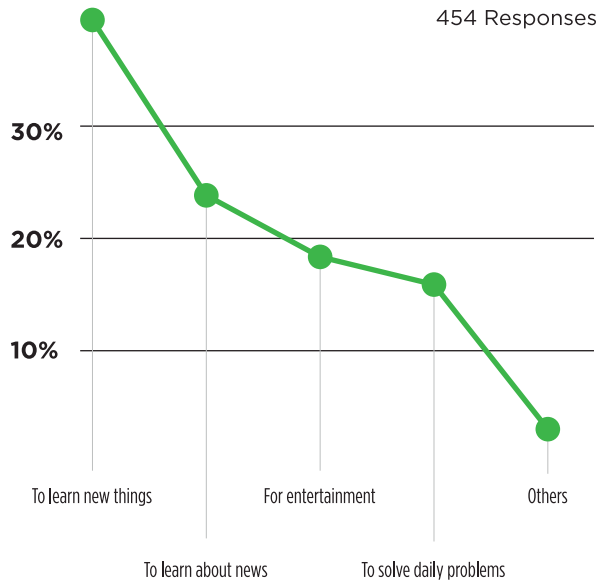
The major result of the iRead4Skills project is the development of an open-access, web-based ICT system that provides **complexity analysis of texts**, reading suggestions based on text complexity, and assistance in text simplification. This system benefits from an interdisciplinary team encompassing **ICT, Linguistics, Economics, and Education**, and is based on input from low-literacy adults, addressing their real needs. Unlike other resources, iRead4Skills focuses on text complexity for **native speakers** rather than second language proficiency levels. The project's success enables policymakers to better understand, measure, and reduce skills gaps, resulting in policy recommendations for learner-centered flexibility in Adult Learning (AL) and Vocational Educational Training (VET) materials. Additionally, the use of this system in cooperation with cultural and scientific actors, and the dissemination of information on text complexity, will open new avenues of information for previously excluded populations.



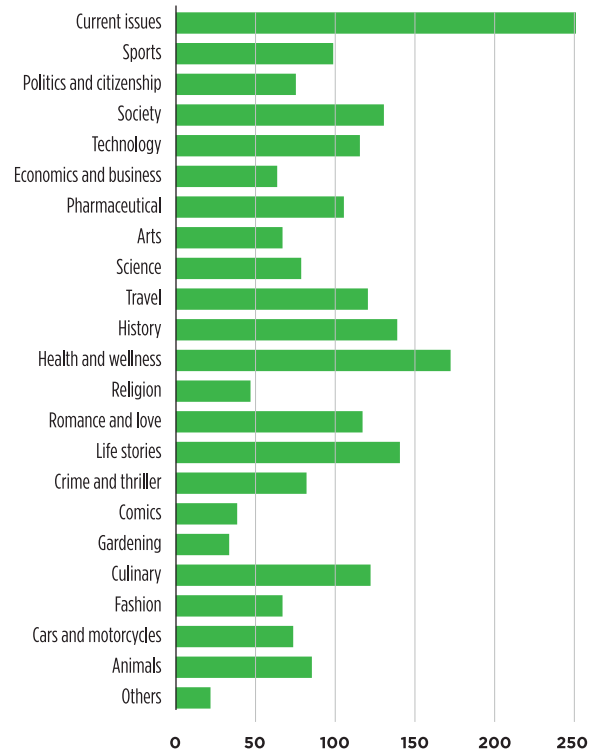
Evidence and Analysis

We conducted two surveys to gather data on reading needs and skills. The first survey targeted adults involved in Adult Learning (AL) and Vocational Educational Training (VET) in Portugal, Belgium, France, and Spain. The second survey explored how reading difficulties impact employment and individual well-being. Despite some challenges, such as low response rates outside Portugal, the surveys provided valuable insights.

WHAT DO LOW LITERACY PEOPLE READ FOR?



WHAT DO LOW LITERACY PEOPLE LIKE TO READ?



Survey Findings



Home-Related Texts:

Respondents generally understand home-related texts like adverts, lease agreements, and condominium regulations.



Transport Information:

Timetables and destinations are well understood, but information about strikes and disruptions is less clear.



Shopping Information:

Product packaging and labeling are mostly understood.

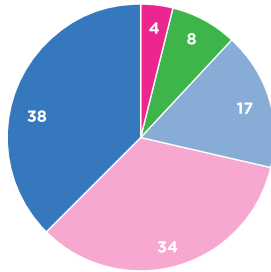


Healthcare Texts:

Medical prescriptions and information about health services are partially or fully understood.

Impact on Employment and Well-being

PRE-TRAINING PERCEPTION OF READING SKILLS IN RELATION TO EVERYDAY NEEDS (IN %)



Very low Low Medium Good Very good

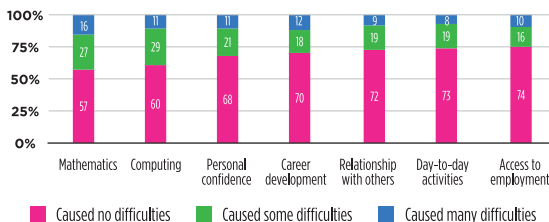
Field: Respondents who expressed themselves on the adequacy of their reading skills and their everyday needs prior to training (n=277).

MAIN REASONS WHY RESPONDENTS TOOK THE READING COURSE (IN %)



Field: Respondents who provided information on their motivation to follow a training course (n=311)

PERCEPTION OF THE CONSEQUENCES OF READING DIFFICULTIES IN SEVEN AREAS



Field: Participants who shared their views on the seven challenges that may arise due to their reading difficulties (n varies from 214 for Mathematics to 229 for Personal confidence or Relationships with others).

- Improved reading skills have a positive impact on other professional abilities, such as teamwork.
- Training increased awareness of literacy issues, though many respondents initially downplayed their difficulties.
- Employees who encounter reading difficulties at work are often hesitant to seek assistance.
- Respondents had mixed views on whether their reading difficulties caused other types of difficulties.
- Employed respondents agreed that their improved reading skills had a positive impact on other professional abilities.

COMPLEXITY FRAMEWORK

We developed a framework with three text complexity levels to match reading materials to users' literacy levels, based on input from AL and VET trainers and learners. The goal is to promote reading skills through an intelligent system that evaluates text complexity and suggests appropriate reading materials.

THE LEVELS OF COMPLEXITY FOR THE PROJECT ARE:



VERY EASY (CEFR A1)

Texts fully or almost fully understood by everyone, including those with very low schooling (up to 6th grade) and minimal reading experience.



EASY (CEFR A2)

Texts fully or almost fully understood by individuals with low schooling (completed primary school up to 9th grade) and poor reading experience.



PLAIN (CEFR B1)

Texts understood on the first read by individuals who completed up to the 9th grade and have functional-to-average reading experience.

Policy Implications and Recommendations

Our findings have several policy implications, leading to specific recommendations to improve literacy and training outcomes.

ACCESS TO INFORMATION

There is a significant shortage of low-complexity texts in essential areas like pricing information, bills, and service conditions. We recommend that directives regarding consumer information should mandate that such information be provided in **plain language**. The complexity descriptors and lexicons developed by the iRead4Skills project are crucial resources to assist in producing these texts.

MOTIVATION AND ADEQUACY OF AL AND VET

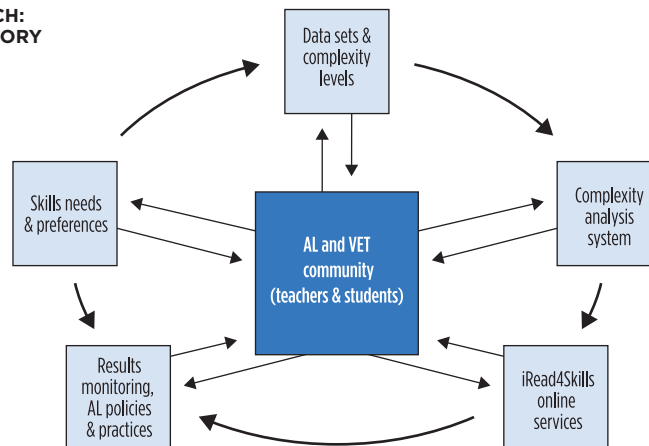
Materials designed for language teaching are often inadequate for low-literacy adults as they do not cover relevant topics and communication situations. To address this, adult training should include **authentic texts** that consider real-life situations. The survey results, along with the devised complexity descriptors and lexicons, can help produce these texts.

COOPERATIVE AND COLLABORATIVE RESEARCH

Recruiting and sustaining cooperation from end-users, especially trainers, is challenging. Trainers feel overwhelmed by constant requests to participate in studies without seeing tangible benefits or proper recognition. We recommend providing **attractive compensation** for their contributions. This could include allowing trainers to allocate some of their working time to participate in projects or providing monetary compensation without requiring their affiliation institution to be a formal project beneficiary.

Recognizing the essential role of end-users and professionals in a cooperative, bottom-up approach to developing knowledge and innovation will enhance the impact of investments in research and innovation, as well as improve society's trust in and relationship with science.

IREAD4SKILLS APPROACH: BOTTOM-UP PARTICIPATORY APPROACH



Project Impact and Future Steps

The iRead4Skills project has established cooperation with AL and VET centers and professionals in **Belgium, France, Portugal, and Spain**, engaging with individuals in both formal and informal training contexts. Targeting over 50 million trainees across Europe, the project aims to significantly impact literacy and related skills. By integrating expertise from ICT, Linguistics, Economics, and Education, and incorporating input from end-users, we are developing groundbreaking tools that surpass existing systems.

Our next steps involve **continuing to test** and refine the intelligent reading system with end-users, enhancing datasets and lexicons for complexity analysis, and further engaging with policymakers and stakeholders to implement our recommendations. The ultimate goal is to support European initiatives that **promote reading skills** and habits among adults, facilitating access to information, scientific and technical knowledge, and culture.

The successful implementation of iRead4Skills will enable policymakers to better understand, measure, and reduce skills gaps, leading to learner-centered flexibility in **AL and VET materials**. Additionally, widespread use of our system, in cooperation with cultural and scientific actors, will disseminate information on text complexity, opening new avenues of information for previously excluded populations.

For more information, please contact the project coordinator, Raquel Amaro, PhD, at NOVA University Lisbon. The project is funded by Horizon Europe and will run from March 2023 to February 2026. Visit our website at iRead4Skills and follow us on social media for updates.

Contact and Additional Information



Coordinator

Raquel Amaro, PhD - NOVA University Lisbon

Funding Programme

Horizon Europe

Duration

March 2023 - February 2026 (36 months)

Website

[iRead4Skills](https://iRead4Skills.com)

Social Media



Disclaimer: Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.