



# Research on the Mechanism of Intelligent Education Enabled by Natural Language Processing Technology

Yueran Zhang<sup>(✉)</sup>

School of Marxism, Beijing Jiaotong University, Beijing 100044, China  
20119003@bjtu.edu.cn

**Abstract.** The development of computer technology and AI computing power has created conditions for the deep integration of natural language processing technology and education. Natural language processing technology is capable of text retrieval and extraction, text classification, intelligent question and answer, text generation and machine translation, which can help improve the synergy and sustainability of intelligent development in education. This paper proposes the “NLP + Education” enabling model based on a double-circle mechanism, taking intelligent education stakeholders, technology enabling interfaces, technology enabling platforms, technology entry and infrastructures as enabling indicators, through the technical implementation path of education intelligence and the user access path of intelligent education stakeholder, a bidirectional cycle of internal and external linkage will be generated, this will help to create an innovative mechanism for transforming education based on natural language processing technology. Additionally, this paper also provides some references and suggestions for the challenges that natural language processing technology may face, such as manpower shortage, technical barriers, resource constraints and so on.

**Keywords:** NLP Technology · Intelligent Education · Circulation Mechanism · Deep Learning

## 1 Introduction

As an endogenous force in the new round of technological revolution and industrial transformation, artificial intelligence (AI) is a powerful force in upgrading global production methods. Solving the problem of unequal distribution of educational resources is a global challenge. The “Artificial Intelligence + Education” (AIED) enabling system can effectively save human capital and time cost, it can generate personal education models to promote the intelligent development of education. As a core component of artificial intelligence, natural language processing technology (NLP) is a framework for human-machine interaction and communication based on the analysis, simulation, conversion and generation of human natural language, it can help users build intelligent platforms for information searching, content recommending, question identifying and

text structuring. Though human-machine communication and cooperation, we can produce individual solutions to meet the changing needs of content and industry, in order to form “intelligent education” during the whole education process. It plays a significant role in promoting the reconstruction of education ecology, enriching education materials, upgrading education tools, improving education services, strengthening precision education and so on. Natural language processing technology has become an important technical supports for the intelligent transformation and integrated development of education.

At present, global development of education intelligence has been upgraded to a two-way human-machine facilitation stage. At the end of 2022, ChatGPT has generated global attention. As a chatbot program based on natural language processing technology, demonstrating higher intelligence and linguistic capabilities in continuous dialogue, contextual semantic generation, literary composition, code writing and revision, which marks a new stage in intelligent teaching and precision teaching based on big data analysis, and indicates a new breakthrough in intelligent education.

The educational enabling of natural language processing technology has also been widely discussed in academic circle. Intelligent products based on NLP technology help improve responsiveness and usability, reducing human dependency in today’s automated world [1]. Thanveer et al. believe that NLP technology generates text data models through extensive language reading to achieve user feedback on services and products [2]. At the same time, technology discrimination create risks of deep falsification, human rights ethics, and academic misconduct such as plagiarism, data falsification, and experimental falsification, which also create challenges for the learning work of educators and scholars around the world [3]. With the theme of the application of natural language processing technology in intelligent education, searching for the terms “Natural Language Processing”, “NLP” and “Education”, about 3043 papers on related topics can be filtered, spanning the period from 2000 to 2022. See Figs. 1, 2 and 3 for details.

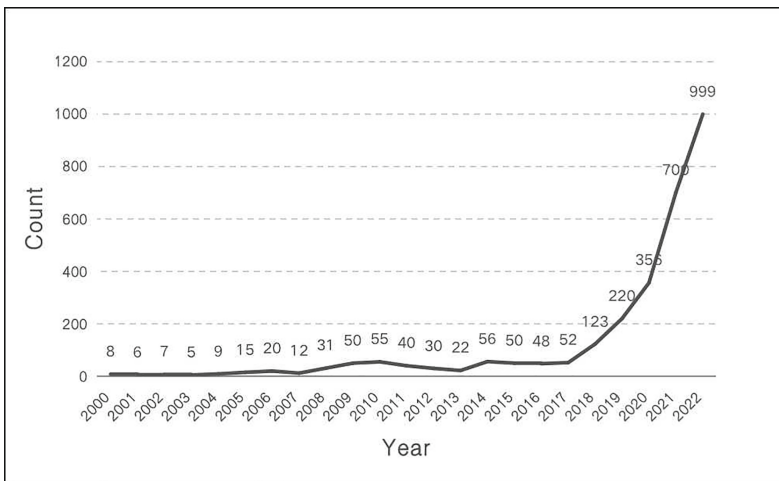
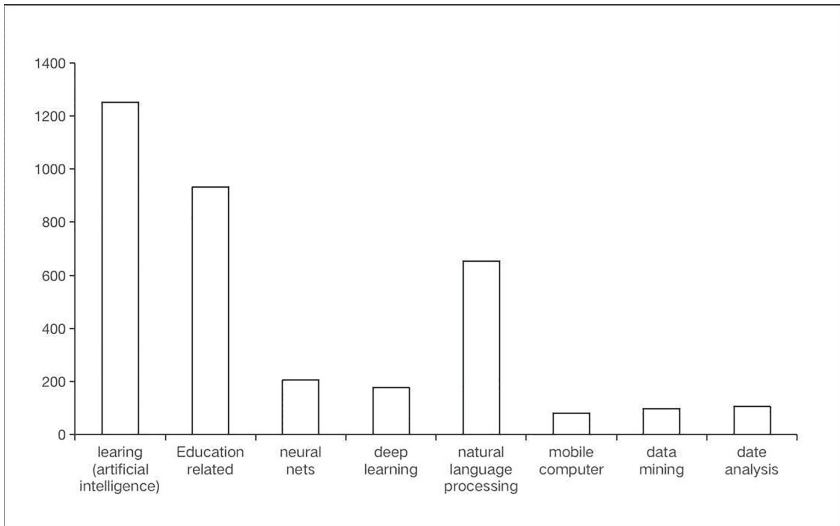
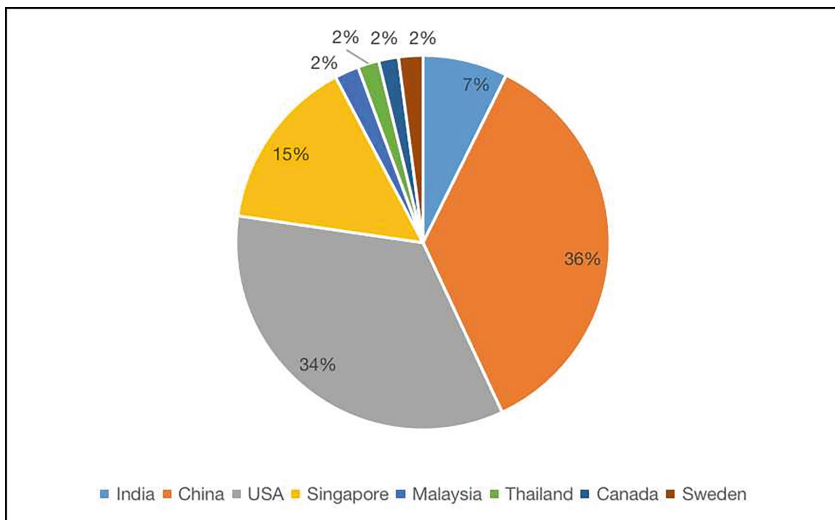


Fig. 1. Year trend of literature volume



**Fig. 2.** Hot spot distribution statistics



**Fig. 3.** Global literature distribution

According to the year trend of literature volume, the research on this topic was relatively poor and the number of literature was small in the first 8 years, from 2008 to 2017, the research status of this topic showed fluctuation; since 2018, the number of literature on this topic climbed sharply and reached a peak in 2022. From the perspective of the hot topics involved, this topic covers a wide range, mainly focusing on “artificial intelligence + education” (1251 articles), education-related fields (931 articles) and natural language technology assistance (654 articles). In terms of the regional distribution

of the literature, based on statistical analysis of literature affiliation and conference locations, China (36%) and the United States (34%) have more research participation and are the two main countries studying this topic globally. Meanwhile, Southeast Asia has also contributed to the exploration of this topic, showing an active trend.

Through literature analysis, academic research mainly focuses on technology iteration, application integration and risk avoidance, but there is a lack of research on system mechanism. This paper proposes the “NLP + Education” model based on a double-circle mechanism, which building a sustainable mechanism for natural language processing technology enabling intelligent education, it deeply expands the efficiency and dimensions of the education intelligence enabling system.

## **2 The Application of NLP Technology in the Development of Educational Intelligence**

Natural language processing technology is mainly engaged in the natural language communication between man and machine. Natural language is the general language used to maintain people’s daily communication and life, including English, Chinese, French, Russian, etc. In education, natural language communication is the basis of “teaching” and “learning” between users. Therefore, natural language processing technology is a fundamental technical tool for the development of intelligent education. The deep integration of natural language processing technology and education has led to the widespread use of intelligent tools in teaching and learning. The breakthroughs in crucial technologies can drive innovation in natural language processing technology enabling mechanisms [4].

### **2.1 The Development of Intelligent Education Enabled by NLP Technology**

The development of intelligent education enabled by natural language processing technology can be roughly divided into four stages: the embryonic stage before 2000, the integration stage (machine learning) from 2000 to 2008, the breakthrough stage (deep learning) from 2008 to 2018 and the intensive stage (large-scale language models) from 2019 to the present. Before 2000, as a subsystem of artificial intelligence, the application of natural language processing technology in education was not mature. The birth of intelligent computer aided instruction (ICAI) technology and other models laid the foundation for the deep enabling of natural language technology. From 2000 to 2008, natural language technology was deeply integrated with education, machine learning became the mainstream direction in this stage, and great progress was made in traditional issues such as semantic understanding, analysis and judgment. The period from 2008 to 2018 was a breakthrough stage in natural language technology enabling education, in terms of technology, the development of intelligent education has shifted from machine learning to deep learning. The development of neural networks has greatly promoted the realization of this transition, word embedding and lexical representation have brought a fundamental breakthrough to natural language processing technology enabling ability, resulting in the birth of GPT language model and BERT pre-training language model, which are able to generate analysis results and intelligently output

new content through the input of text, images, music and data [4], it provides more specific technical support for the intelligent development of education. Recent years, countries all over the world have actively invested in and participated in the deployment of intelligent education, for instance, China has promoted the systematic implementation of the construction of intelligent education project since 2017. The year 2019 to now is a period of rapid development, the transformer based pre-training model makes the boundary between large-scale language model and natural language further narrow, and improves the ability of AI technology represented by natural language processing technology to understand the world. From the GPT-2 and GPT-3 language models to ChatGPT, a large number of models have developed in a spurt, technology iteration has improved the fitting degree of realistic education, and natural language processing technology has ushered in the golden age of enabling education, as shown in Fig. 4.

The rapid development of natural language processing technology can promote the development of educational tools, the process of innovation and optimization of intelligent tools in educational activities is the process of natural language processing enabling intelligent education [4]. With the effects of environmental conditions, market factors,

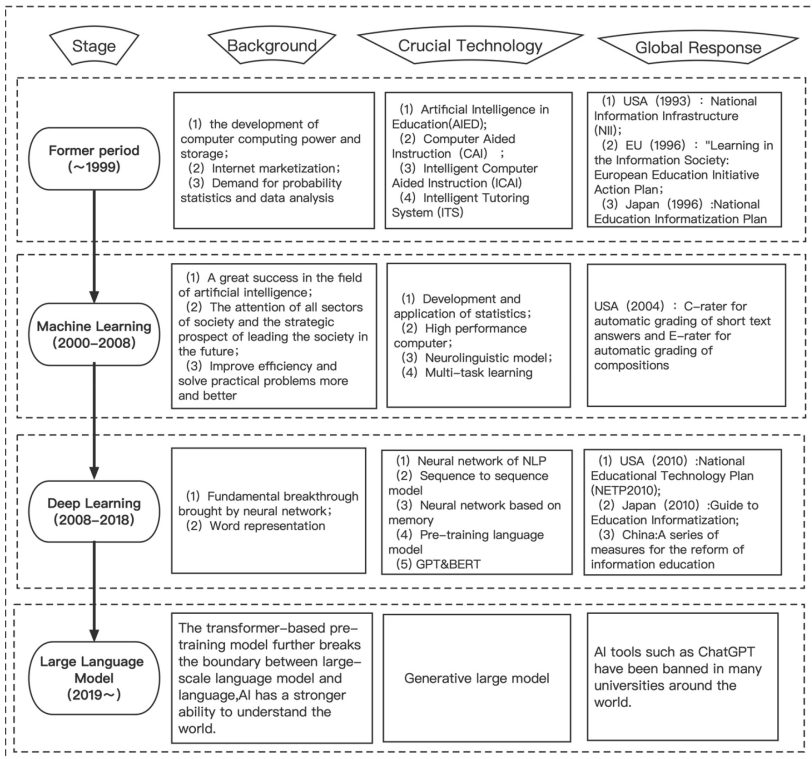


Fig. 4. The development of intelligent education enabled by natural language processing technology

technological synergies, industry needs and the continuous improvement of crucial technologies, the contribution of natural language processing technology to education is gradually becoming more scalable, general and concise, reinforcing the level and depth of its empowerment.

## **2.2 The Concrete Pattern of the Intelligent Development of NLP Enabling Education**

The development of computing power has promoted the advancement of artificial intelligence technologies such as text recognition, human-machine interaction Q&A, etc., it achieves breakthroughs and innovations in education scenarios, teacher assistance, education management, education feedback and other applications, which bring opportunities to education industry.

Educational informations acquisition based on text retrieval and extraction technology: extracting crucial information from the input text and generating a brief summary of the text. The text summary is aimed at converting text into a short summary containing key information. Text summaries can be divided into single-document summaries and multi-document summaries according to the type of input. In education, text searching can quickly find key information while effectively extracting the main content of a text, extremely reducing the reading time of educators and student users.

Educational resources integration based on text classification technology: text classification is a fundamental function of natural language processing technology in educational enabling applications. It refers to the process of mapping text containing information to a pre-given category of topics by means of artificial intelligence models, the purpose of categorising and integrating text information is achieved by tagging different labels. The sources of text information including but not limited to internet databases, SNS databases, knowledge databases, teaching databases. For intelligent education, the text classification function of natural language processing technology, mainly used for evaluating lesson plans, classifying teaching materials and evaluating students' homework, etc. It can help teachers and students make initial judgement about article style and direction, and improve the efficiency of reading and decision-making.

Educational tools upgrade based on intelligent question and answer technology: intelligent question and answer is the process of organising disordered corpus information into a structured format. Building knowledge classification models can more accurately understand natural language forms of user questions. Therefore, the users' questions need to be contextualised in a certain chronological order, and the correct matching answer is generated by retrieving a heterogeneous corpus or knowledge database of questions and answers. Today, with the continuous development of intelligent question and answer technology, it can accurately understand students' questions and provide precise answers, reducing embarrassing situations when the teacher is not present and making a great contribution to the development of intelligence.

Educational competencies progress based on text generation technology: generating high quality articles, conversations and summaries can reduce the amount of manual text writing. Natural language generation systems must go through the steps of planning and merging content, in order to produce natural texts. With the gradual development of big

model technology, ChatGPT is widely used in education, a large amount of high quality text can be generated to provide effective reference for teachers and students.

Educational contents enrichment based on machine translation technology: machine translation is the process of automatically translating text from one language to another without human involvement. Machine translation now goes beyond word-for-word translation, it can convey the full meaning of the original language text with the target language, and analyse all text elements to identify how words interact with each other. In education, this technology is commonly used in translating texts and articles, providing a great convenience for teachers and students.

### 3 The Technical Path of “NLP + Education” Model Based on Double-Circle Mechanism

To achieve intelligent education, we need to coordinate internal and external mechanisms to transform and upgrade education intelligently. With the model of “AI + Education”, we propose the “NLP + Education” enabling model based on a two-way cycle mechanism, as shown in Fig. 5.

The “NLP + Education” model consists of two main cycles, the internal cycle and the overall cycle. It involves five major enabling indicators: intelligent education stakeholder, technology enabling interface, enabling platform, technology entry and related infrastructure.

Intelligent education stakeholder include public users, students, educators, educational administrators and related technical personnel. They are the main strength and element source of the “NLP + Education” enabling model. Technology enabling interface includes educational information retrieval and extraction ability, text classification ability, intelligent question-and-answer ability, educational text generation ability and machine translation ability. Enabling platform includes education operation management platform, enabling model development platform, natural language processing training platform, ability management platform, intelligent education model optimization platform, education data platform and management platform. It provides the possibility and necessity for the continuous flow and healthy circulation of intelligent education circulation mechanism. Technology entry includes educational natural language resource databases, data mining and cleaning, feature extraction, knowledge graphs, artificial intelligence algorithms. Related infrastructure is a material precondition for the achievement of the above indicators, including power supply, computing power supply, network capacity supply, scientist resource supply, policy and legal guarantee and financial support. It connects the internal and external mechanisms of the “NLP + Education” enabling model, also creates a stable environment for the value transformation of technology enabling and the sustainable development of intelligent education.

The internal mechanism of “NLP + Education” enabling model consists of four enabling indicators, namely intelligent education stakeholder, technology enabling interface, enabling platform and technology entry. Two enabling paths can be generated to realize the internal circulation mechanism. One is the technology realization path based on intelligent education (Path ①, Technology access → technology enabling interface → technology enabling platform → education feedback collector → education natural

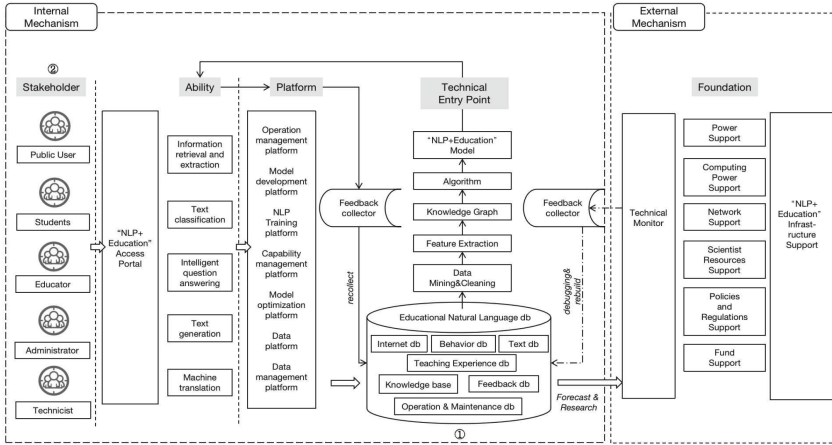


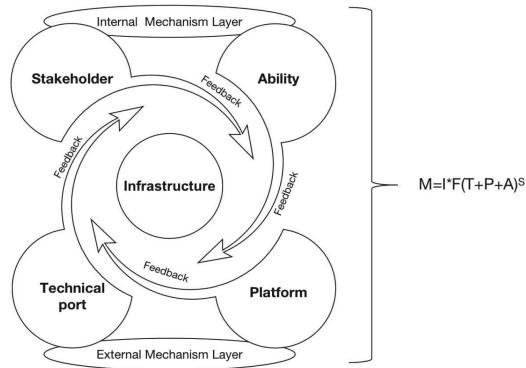
Fig. 5. Two-way cycle framework of “NLP + Education”

language resource library), building a complete closed-loop of corpus input, output and feedback update. The other is the user access path based on intelligent education stakeholder (path ②, intelligent education stakeholders → technology enabling interface → technology enabling platform → educational natural language resource library), which realizes a complete closed-loop of user input, capability acquisition, platform use and data feedback. The external mechanism of the “NLP + Education” model is primarily a technological infrastructure guarantee mechanism, the intelligent education levels can be predicted and evaluated by form a feedback collection through technology monitors. The internal and external mechanisms are parallelized through the debugging of the resource base to build a two-way circular system.

The bidirectional cycle framework of “NLP + Education” can realize the intelligent development and transformation of Education, which can precisely evaluate the accuracy and availability of the overall design. Based on above research, the enabling model and enabling formula of “NLP + Education” can be constructed, and the feature models are generated from different dimensions, as detailed in Fig. 6.

In the “NLP + Education” enabling formula  $M = I * F(T + P + A)S$ . M (Mechanism) represents general circulation mechanisms. I (Infrastructure) represents physical infrastructure requirements during the circle. F (Feedback) represents the feedback generated during the two-way cycle mechanism. T (Technology), P (Platform) and A (Ability) respectively represent the technology, platform and ability requirements in intelligent education sessions. S (Stakeholder) represents relevant stakeholders of promoting intelligent education. The enabling formula can help coordinate the implementation of various factors, promoting the sustainability and effectiveness of technology empowerment.





**Fig. 6.** Enabling model of “NLP + Education”

## 4 Possible Challenges of Future Natural Language Technology in the Development of Intelligent Education

Based on the above discussion, it can be concluded that natural language processing technology has become a key technological support for the continued development of future educational intelligence. The effective synergy between natural language processing technology and other AI technologies also brings new opportunities for intelligent education. At the same time, further sustainable empowerment also faces challenges in terms of technical staff, technical implementation and technical application.

### 4.1 Human Resource Challenges

Artificial Intelligence development engineers require high levels of computer skills as well as a good background in mathematics and statistics. In addition, at the algorithm implementation stage, development engineers need to have excellent code engineering skills, be familiar with programming languages such as Python, and be able to engineer the implementation of algorithmic models. Most importantly, development engineers need to be deeply involved in educational scenarios and have a grasp of business requirements. It takes a lot of time to train a technician and requires the technician to be competent in all aspects of business, algorithms and engineering at the same time.

### 4.2 Technical Implementation Challenges

At the current stage, natural language has entered the era of large models, which then place a very high demand on computing power. According to relevant data, the total computing power consumption of ChatGPT developed by OpenAI is about 3640PF-days (i.e. if it calculates one thousand trillion times per second, it needs to calculate 3640 days), which requires 7–8 data centers with an investment scale of 3 billion and 500P computing power to support the operation [5]. At the same time, ChatGPT does not release its source code, making it extremely difficult for other companies to implement, creating a large technical barrier that will have an impact on downstream industries such as education.

### 4.3 Resource Support Challenges

High computing power requirements dictate high resource consumption, the high consumption of arithmetic resources, storage resources, network resources and power resources is most representative. According to statistics, an average sized transformer network with 65 million parameters, trained on 8 GPUs for 12 h, consumes 27 kWh of electricity and emits 11.8 kg of CO<sub>2</sub>, a larger BERT model with 110 million parameters, trained on 64 GPUs for 80 h, consumes 1507 kWh of electricity and emits 652.3 kg of CO<sub>2</sub> [6]. The large-scale application of education will cause a huge challenge to the development of natural resources and the investment of funds for environmental protection.

## 5 Conclusions

The application of natural language processing technology in education is already a global trend, how to promote technology integration and mechanism innovation is an essential subject for the realization of intelligent education. Natural language processing technology matches the current mainstream trend of intelligent education development, which can help strengthen deep learning, reshape the education environment, improve teaching efficiency and promote educational equity. Furthermore, the enabling mechanism is affected by internal and external factors such as insufficient investments, technical barriers and resource constraints, it is our common goal to build a global technology community and promote the intelligent development of human educational civilization. From the perspective of human technological evolution, the pattern of technological development is always a mixture of lows and highs, progress and challenges. Although ChatGPT caused a global controversy, but its application and promotion is still the general trend. The key to solving the problem is how to take advantage of the technology while establishing and improving appropriate management mechanisms to deal with its risks.

## References

1. A. Mondal, M. Dey, D. Das, S. Nagpal and K. Garda, "Chatbot: An automated conversation system for the educational domain," 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), Pattaya, Thailand, 2018, pp. 1-5.
2. T. Shaik et al., "A Review of the Trends and Challenges in Adopting Natural Language Processing Methods for Education Feedback Analysis," in *IEEE Access*, vol. 10, pp. 56720-56739, 2022.
3. Zhang Xiaheng. The logical deconstruction, impact analysis and policy recommendations of ChatGPT[J]. *Journal of Xinjiang Normal University*(Edition of Philosophy and Social Sciences),2023,40(5):18-28.
4. Zhangbo,Do Bohai.Natural language processing technology empowers the development of educational intelligence --from the perspective of AI scientists[J].*Journal of East China Normal University (Educational Science Edition)*,2022,40(09):19-31.

5. Jin Yezi. The popularity of ChatGPT drives the demand for computing power. Can China's computing power scale be supported[N]. China Business Daily,2023-02-15(A06).
6. Chai Qimin, Li Moyu. Application scenario and effect assessment of digital intelligence technology under global carbon neutralization process[J].Yuejiang Academic Journal,2022,14(05):51-60+173.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

