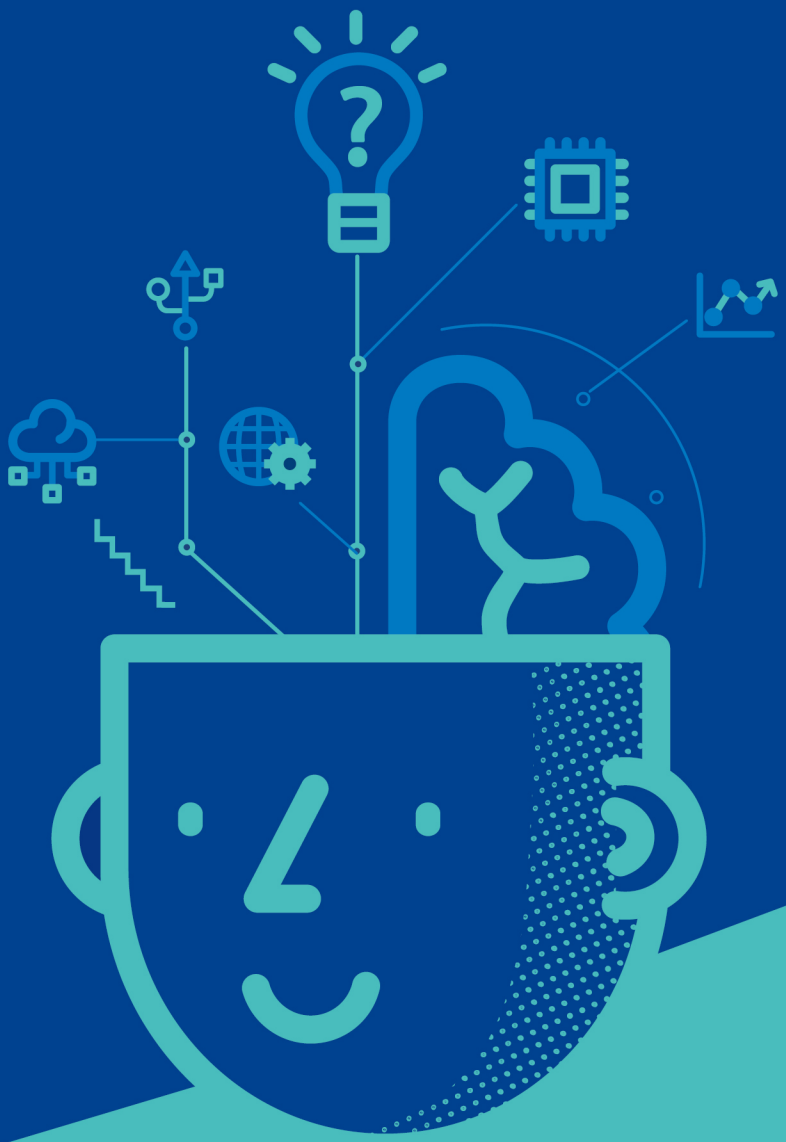
서울형 인공지능 윤리교육자료

초등학교용



부록

인공지능 윤리 교육 총론

부록

인공지능 윤리 교육 총론

1 인공지능 윤리 교육 총론

인공지능 윤리교육의 총체적 방향은 국가 수준 인공지능 윤리 기준에 서 제시된 10대 핵심요건을 교육과정에 적용할 수 있도록 개발된 내용 체계 및 성취기준을 활용하여 초·중·고 서울형 인공지능 윤리 교육의 기초 자료로 제시

<표 1> 인공지능 윤리 교육의 내용 체계

영역 (10대 핵심요건)	일반화된 지식 (해석된 원칙)	내용		기능
		구성 요소	주요 주제	
인권보장	인공지능은 인간의 기본권을 침해하지 않는 범위 내에서 사용되어야 한다는 점에 대한 이해를 바탕으로, 인공지능이 인권존중을 실현하고 인권 침해를 방지하는 데에 활용될 수 있는 방안에 대해 고민한다.	인간의 기본권 존중 기본권 침해 방지	• 인공지능 기술 사용이 인간의 기본권을 침해할 소지가 있는가? • 인공지능의 개발과 활용은 인권존중을 실현하는 데에 도움을 줄 수 있는가? • 인공지능은 인권의 침해를 방지하는 데에 사용될 수 있는가?	[인공지능 기술을 통한 인권 존중] • 인권보장 실현 및 침해 방지 방안 탐구하기 • 인권 침해 가능성 이해를 통해 인권 감수성 기르기
프라이버시 보호	인공지능 기술로 인해 프라이버시가 침해될 수 있다는 사실을 인지하고, 이를 방지할 수 있는 방안에 대해 탐구한다.	프라이버시 침해 이해 개인정보의 합리적 사용	• 인공지능의 개발과 활용 과정에서 프라이버시 침해가 발생할 수 있는가? • 인공지능 시스템은 개인정보의 오남용을 어떻게 방지할 수 있는가? • 인공지능 기술이 사용될 경우 프라이버시 보호 체계나 개인정보 분류 및 활용의 방법을 어떻게 정해야 하는가?	[인공지능 기술 활용시 프라이버시 보호] • 프라이버시 침해 가능성 이해하기 • 개인정보의 합리적 사용 방안 탐구하기

영역 (10대 핵심요건)	일반화된 지식 (해석된 원칙)	내용		기능
		구성 요소	주요 주제	
다양성 존중	인공지능 기술이 활용되는 모든 영역에서 다양성과 대표성이 반영되어야 하는 중요성을 인식하고, 접근성 보장 및 혜택의 공정한 분배 방안에 대해 탐구한다.	다양성 및 대표성 반영 접근성 보장 혜택의 공정한 분배	• 인공지능 기술은 사용자를 차별하지 않고 다양성과 대표성을 반영하는가? • 사회적 약자 및 취약 계층의 인공지능 기술 및 서비스에 대한 접근성이 보장되고 있는가? • 인공지능 기술로 인한 혜택은 어떤 방식으로 특정 집단이 아닌 모든 사람에게 공정하게 분배될 수 있는가?	[인공지능 기술을 통한 다양성 존중] • 다양성과 접근성 보장의 필요성 인식하기 • 인공지능 기술로 인한 혜택의 공정한 분배 방안 탐구하기
침해 금지	인공지능이 인간에게 피해를 입히는 목적으로 활용되지 않아야 함을 이해하고, 인공지능이 야기할 수 있는 위험과 부정적 결과에 대한 대처방안에 대해 탐구한다.	인간 보호 피해처리의 합리성 위험 최소화	• 인공지능이 인간에게 피해를 입히는 목적으로 활용되는 것을 어떻게 방지할 수 있는가? • 인공지능으로 인한 피해 발생시 이를 어떻게 해결할 수 있는가? • 인공지능 시스템이 야기할 수 있는 위험요소를 최소화 할 수 있는 방안은 무엇인가?	[인공지능으로 인한 피해 예방] • 피해 예방 및 피해 발생시 효율적이고 합리적인 대처 방안에 대해 탐구하기 • 인공지능 기술로 인한 위험 요소의 최소화 방안 탐구하기
공공성	인공지능은 공공성 증진과 인류의 공동 이익을 위해 활용되어야 함을 이해하고, 인공지능이 사회에 미치는 순기능을 극대화하고 역기능을 최소화하기 위한 방안에 대해 탐구한다.	사회의 공동이익 추구 공공성 보장	• 인공지능 기술은 사회의 공공성과 인류의 공동이익을 증진시키는가? • 인공지능 기술로 인한 공공성의 훼손이나 사회에 미치는 악영향을 어떻게 최소화할 수 있는가?	[인공지능 기술을 통한 공공성 증진] • 사회의 공공성 증진의 필요성 이해하기 • 공공성 훼손의 최소화 방안 탐구
연대성	인공지능 기술을 통해 다양한 집단 간의 통합과 연대성을 확보할 필요성을 이해하고, 인공지능의 개발 및 활용에 관해 국제적 협력 방안을 탐구하며, 미래세대를 포함한 전 세대를 충분히 배려하도록 노력한다.	통합과 연대 국제적 협력 전 세대 배려	• 인공지능 기술을 활용할 경우 다양한 집단 간의 통합과 연대를 이루는데 도움을 줄 수 있을까? • 인공지능의 개발 및 활용에 대한 국가간 합의 및 국제적 협력은 어떻게 이룰 수 있는가? • 인공지능 기술의 개발 및 활용이 미래세대를 포함한 전 세대를 충분히 배려하고 있는가?	[인공지능 기술을 통한 통합과 연대 확보] • 통합과 연대의 필요성 이해하기 • 국제적 협력 방안 탐구하기 • 전 세대 배려하기

영역 (10대 핵심요건)	일반화된 지식 (해석된 원칙)	내용		기능
		구성 요소	주요 주제	
데이터 관리	인공지능 기술을 활용하는 과정에서 수집되는 데이터는 그 목적에 부합하도록 활용해야 함을 이해하고, 데이터 수집과 활용의 전 과정에서 데이터 편향성을 최소화하는 방안을 탐구한다.	데이터 관리 편향성 최소화	• 인공지능 기술에 사용되는 데이터 관리를 어떻게 해야 하는가? • 데이터를 수집한 목적에 부합하도록 활용하게 하는 적절한 절차나 방안은 무엇인가? • 데이터 수집과 활용의 전 과정에서 데이터 편향성을 최소화하기 위한 노력에는 무엇이 있을까?	[적절한 데이터 관리와 편향성 최소화] • 데이터 적절한 관리의 필요성 이해하기 • 데이터의 편향성 최소화 방안 탐구하기
책임성	인공지능 시스템과 관련된 모든 과정에서 적절한 책임주체를 설정해야 함을 이해하고, 책임소재를 명확히 하기 위한 기록 및 추적의 필요성을 인식한다.	적절한 책임 주체 설정 기록 및 추적 활성화	• 인공지능 시스템의 작동, 사용, 오용 등에 대한 책임의 주체를 어떻게 정해야 할까? • 인공지능 시스템의 모든 작동 과정을 기록하고 추적할 수 있는 방안은 어떻게 마련할 수 있을까?	[적절한 책임주체 설정] • 책임주체 설정의 필요성 이해하기 • 기록 및 추적 방안의 중요성 인식하기
안전성	인공지능 시스템의 개발 및 활용 전 과정에서 안전을 보장하는 것이 필수적임을 이해하고, 안전성을 확보하는 다양한 방안을 탐구한다.	안전의 필요 인식 지속적인 안전성 확보 노력	• 인공지능 시스템의 안전성은 누가 어떻게 확보해야 할까? • 인공지능 시스템의 안전성은 어떻게 검증할 수 있을까? • 인공지능 시스템의 지속적인 모니터링, 위험 측정, 긴급상황 발생시 대처 기능의 마련 등은 안전성 확보를 위해 충분한가?	[안전성 확보 방안 강구] • 안전성 확보를 위한 주체별 방안 탐구하기 • 지속적인 안전성 확보 방안 탐구하기
투명성	사용자가 인공지능 시스템의 활용 및 발생할 수 있는 위험 등에 대해 제공받아야 하는 함의 중요성을 인식하고, 인공지능의 투명성과 설명 가능성을 확보하는 방안에 대해 탐구하고, 투명성의 확보가 궁극적으로 사회적 신뢰 형성에 기여함을 이해한다.	사전 고지 필요 인식 투명성의 합리적 이해 사회적 신뢰 형성	• 인공지능 시스템의 활용과 관련된 내용 중 어떤 것들을 사용자에게 고지하고 제공해야 할까? • 인공지능 시스템의 개발 및 활용의 과정에서 어떻게 작동 과정을 합리적이고 투명하게 확인할 수 있을까? • 적절한 투명성의 확보는 사회적 신뢰 형성에 어떤 영향을 미칠까?	[투명성 확보 방안 이해] • 사전 고지를 통한 투명성 확보의 방안 이해하기 • 투명성 확보가 사회적 신뢰 형성에 미치는 영향 인식하기

2 인공지능 윤리 교육의 성취기준

01 인권보장

인공지능은 인간의 기본권을 침해하지 않는 범위 내에서 사용되어야 한다는 점에 대한 이해를 바탕으로, 인공지능이 인권존중을 실현하고 인권 침해를 방지하는 데에 활용될 수 있는 방안에 대해 고민한다.

[인-01-01] 인공지능이 인간의 기본권(인권, 존엄성)을 침해할 수 있다는 사실을 이해한다.

[인-01-02] 사람을 돕는 인공지능이 인권을 보장하는데 사용될 수 있음을 이해한다.

[인-02-01] 인공지능이 인권을 존중하는 사례를 찾아 조사할 수 있다.

[인-02-02] 잘못된 인공지능 활용으로 인한 인권 침해 사례를 조사할 수 있다.

[인-02-03] 인공지능의 편향된 데이터와 알고리즘이 차별과 편견에 영향을 미친 사례를 통해 인권 감수성을 기른다.

[인-03-01] 인간의 기본권을 지키기 위하여 인공지능이 활용될 수 있는 방안에 대해 제시할 수 있다.

[인-03-02] 인권을 존중하는 인공지능을 개발·활용하기 위한 노력(개인, 사회, 제도와 정책)에 대해 이해한다.

02 프라이버시 보호

인공지능 기술로 인해 프라이버시가 침해될 수 있다는 사실을 인지하고, 이를 방지할 수 있는 방안에 대해 탐구한다.

[프-01-01] 인공지능의 개발과 활용과정에서 사생활 침해가 발생할 수 있음을 이해한다.

[프-01-02] 인공지능 기술이 개인정보를 과도하게 활용할 수 있음을 이해한다.

[프-02-01] 인공지능의 개발과 활용과정에서 개인정보의 오남용이 발생할 수 있음을 이해한다.

[프-02-02] 인공지능 기술이 개인정보를 오남용 하였을 때 발생하는 문제 및 이에 대한 대처방안을 살펴본다.

[프-03-01] 인공지능 기술이 사생활 침해를 예방하고, 개인정보를 보호하기 위해 어떤 조치를 취할 수 있는지 살펴본다.

[프-03-02] 인공지능 시스템에서 개인정보가 유출되거나 사생활이 침해되었을 때 보호받을 수 있는 방법에 대해 이해한다.

[프-03-03] 인공지능 기술로 인해 침해될 수 있는 사생활과 개인정보를 보호하기 위해 개인 및 사회는 어떠한 노력을 기울일 수 있는지 살펴본다.

03 다양성 존중

인공지능 기술이 활용되는 모든 영역에서 다양성과 대표성이 반영되어야 하는 중요성을 인식하고, 접근성 보장 및 혜택의 공정한 분배 방안에 대해 탐구한다.

[다-01-01] 인공지능 기술이 사용자의 다양성(연령, 성별, 계층 등)을 온전하게 반영하지 못 할 수 있음을 이해한다.

[다-01-02] 인공지능 기술이 다양한 사용자에게 공정하게 적용될 수 있어야 함을 이해한다.

[다-02-01] 인공지능 기술이 사회적 약자 및 취약 계층의 접근성·사용성에 도움을 주는 사례를 찾아본다.

[다-02-02] 인공지능 기술이 사회 구성원에게 불이익을 주는 사례가 있는지 찾아보고 이를 보완하기 위한 노력에 대해 탐구한다.

[다-03-01] 인공지능 기술로 인한 혜택이 모든 사람들에게 공정하게 돌아가도록 하기 위해 어떠한 노력이 필요한지 이해한다.

[다-03-02] 인공지능 기술을 바탕으로 우리 사회의 다양성을 존중하기 위해 개발자와 사용자가 기울여야 하는 노력에 대해 토의한다.

04 침해금지

인공지능이 인간에게 피해를 입히는 목적으로 활용되지 않아야 함을 이해하고, 인공지능이 야기할 수 있는 위험과 부정적 결과에 대한 대처방안에 대해 탐구한다.

[침-01-01] 인공지능이 인간에게 피해를 입히는 목적으로 활용될 수 있음을 이해한다.

[침-01-02] 인공지능의 활용으로 자신이 다양한 피해를 입을 수 있음을 이해한다.

[침-02-01] 인공지능 시스템이 나에게 영향을 미칠 수 있는 위험요소를 인지할 수 있는 능력을 기르고, 이를 최소화할 수 있는 방안에 대해 탐구한다.

[침-02-02] 인공지능으로 인해 발생하는 다양한 피해 사례를 살펴보고 이를 해결하기 위한 방법에 대해 토의한다.

[침-03-01] 인공지능으로 인해 피해가 발생했을 경우, 피해보상과 재발 방지에 관해 개발자, 사용자, 그리고 정책 결정자가 기울일 수 있는 노력에 대해 이해한다.

[침-03-02] 인공지능 시스템이 가하는 위험 수준이 개인마다 서로 다를 때 이를 어떻게 처리하여야 하는지 논의한다.

05 공공성

인공지능은 공공성 증진과 인류의 공동이익을 위해 활용되어야 함을 이해하고, 인공지능이 사회에 미치는 순기능을 극대화하고 역기능을 최소화하기 위한 방안에 대해 탐구한다.

[공-01-01] 인공지능의 개발과 활용이 개인의 행복과 사회의 공공성 증진에 기여할 수 있음을 이해한다.

[공-01-02] 인공지능의 개발과 활용이 사회의 공공성을 훼손할 수 있음을 이해한다.

[공-02-01] 인공지능의 개발과 활용이 개인의 행복과 사회의 공공성을 증진한 사례와 그렇지 않은 사례에 대해 탐구한다.

[공-03-01] 인공지능이 공공성을 유지할 수 있도록 개발자, 사용자, 정책입안자가 기울이는 노력에 대해 이해한다.

[공-03-02] 인공지능 시스템이 사회에 악영향을 미칠 경우 이를 최소화하기 위해서 어떻게 하여야 하는지 토의한다.

[공-03-03] 인공지능 시스템이 공공성을 훼손할 때 취약한 사람은 누구이며 이를 방지하기 위해 어떤 노력을 하여야 하는지 토의한다.

06 연대성

인공지능 기술을 통해 다양한 집단 간의 통합과 연대성을 확보할 필요성을 이해하고, 인공지능의 개발 및 활용에 관해 국제적 협력 방안을 탐구하며, 미래세대를 포함한 전 세대를 충분히 배려하도록 노력한다.

[연-01-01] 인공지능이 사회의 통합과 발전에서 차지하는 역할의 범위와 중요성을 이해한다.

[연-01-02] 인공지능의 영향에 대한 사회적 제도적 차원의 정의와 원칙, 가이드라인을 조사한다.

[연-01-03] 인공지능 기술이 미래세대에게 이익이나 피해에 대해 알아보고 발생할 수 있는 문제를 해결하는 의지를 가진다.

[연-02-01] 인공지능 기술의 진보와 확산으로 상호작용하는 인간 사이의 필수적 관계를 강화 시키는 사례를 찾는다.

[연-02-02] 인공지능의 개발 및 활용에 대한 국가간 합의 및 국제적 협력의 방안을 제시할 수 있다.

[연-03-01] 인공지능 기술을 활용해 다양한 집단 및 세대 간의 통합과 연대를 증진하는 방안을 제시할 수 있다.

[연-03-02] 인공지능 기술을 기반으로 한 초연결 사회에서 편견과 차별을 없애고 연대성을 증진하는 방법에 대해 토의한다.

07 데이터 관리

인공지능 기술을 활용하는 과정에서 수집되는 데이터는 그 목적에 부합하도록 활용해야 함을 이해하고, 데이터 수집과 활용의 전 과정에서 데이터 편향성을 최소화하는 방안을 탐구한다.

[데-01-01] 인공지능 기술의 악용을 막기 위해 데이터를 관리하는 수준에 대해 토론을 통해 알아보고 데이터 관리의 의미를 이해한다.

[데-01-02] 데이터의 입수, 결합, 추출에서부터 폐기에 이르는 데이터의 생애주기를 이해하며, 데이터의 출처 및 인위적 세부 처리 과정에 대한 기록을 유지할 필요에 대해 이해한다.

[데-02-01] 데이터 관리의 방법과 중요성을 알고 데이터 관리가 필요한 사례를 찾아 제시한다.

[데-02-02] 데이터 편향성을 개선하기 위한 방안과 그 사례를 찾아 제시한다.

[데-03-01] 데이터를 수집해 목적에 부합하게 활용하는 절차나 방안을 토론을 통해 찾아 데이터 활용 방식을 제시한다.

[데-03-02] 편향되지 않은 데이터를 수집, 처리하는 인공지능을 개발하기 위한 방안을 토론을 통해 알아보고 이를 생활에 적용하고 실천한다.

08 책임성

인공지능 시스템과 관련된 모든 과정에서 적절한 책임주체를 설정해야 함을 이해하고, 책임소재를 명확히 하기 위한 기록 및 추적의 필요성을 인식한다.

[책-01-01] 인공지능 시스템의 작동, 사용, 오용 등에 대한 책임의 주체의 문제에 대해 이해하고 해결 방법에 대해 토의한다.

[책-01-02] 인공지능 시스템의 작동과정을 기록하고 추적하는 이유와 방법에 대해 이해한다.

[책-02-01] 인공지능 시스템을 개발, 배포, 사용할 때 시스템에 대한 책임의 문제 사례와 해결 방안을 제시 할 수 있다.

[책-02-02] 인공지능 운영자의 결정에 의해 영향을 받은 사례와 그 일을 해결하기 위한 방안을 제시 할 수 있다.

[책-03-01] 인공지능의 사용이나 오용에 대한 도덕적, 사회적 영향에 대한 책임의 주체를 정하고 그 이유를 설명 할 수 있다.

[책-03-02] 인공지능 시스템에서 의사결정 과정의 각 단계에 인간의 개입에 대한 것을 기록하는 필요성에 대해 토론하고 이해한다.

09 안전성

인공지능 시스템의 개발 및 활용 전 과정에서 안전을 보장하는 것이 필수적임을 이해하고, 안전성을 확보하는 다양한 방안을 탐구한다.

[안-01-01] 인공지능 시스템의 안전성을 확보하는 주체에 대해 조사하고 이해한다.

[안-01-02] 인공지능 시스템의 안정성을 검증하고 보장하기 위한 방안에 대해 찾고 토론한다.

[안-02-01] 인공지능 개발자가 인공지능 시스템 개발시 사용자의 안전에 대해 고려해야 하는 정도를 토론하여 제시한다.

[안-02-02] 인공지능이 수집할 데이터의 위험성과 안전에 대한 조치에 관한 사례를 찾고, 그 문제의 해결방안을 제시한다.

[안-02-03] 인공지능 운영자가 인공지능 시스템의 위험의 식별과 완화 조치를 위한 윤리 프레임 워크를 구성하고 구체적인 사례를 제시한다.

[안-03-01] 인공지능 운영자와 공급업체간의 성능 유지를 위한 지속적인 협력이 우리에게 끼치는 영향에 관해 알아보고 우리의 삶에 끼치는 영향에 대해 생각하고 발생할 수 있는 문제의 해결 방안을 제시한다.

[안-03-02] 인공지능 시스템에 의해 발생할 수 있는 문제를 해결하기 위한 '중지 버튼'의 필요성에 대해 토론을 통해 알아보고, 그 문제에 관한 구체적인 적용 방안을 제시한다.

10 투명성

사용자가 인공지능 시스템의 활용 및 발생할 수 있는 위험 등에 대해 제공받아야 함의 중요성을 인식하고, 인공지능의 투명성과 설명 가능성을 확보하는 방안에 대해 탐구하며, 투명성의 확보가 궁극적으로 사회적 신뢰 형성에 기여함을 이해한다.

[투-01-01] 인공지능 시스템과 관련하여 사용자에게 고지하고 제공해야 할 필요성을 이해한다.

[투-01-02] 사용자에게 인공지능 시스템을 사용하는 방법에 대한 교육자료나 면책 조항을 제시하는 것이 야기할 수 있는 문제에 대해 알아보고 이 문제를 해결하기 위한 방안을 제시한다.

[투-02-01] 인공지능과 관련된 중요한 결정, 특히 손실, 피해, 손상을 초해할 수 있는 결정에 대해 개발자가 추적 가능성을 고려하고 관련 내용 공개해야 할 필요성을 이해한다.

[투-02-02] 인공지능 운영자가 인공지능에 결정에 영향을 받은 사람들에게 그 결정 근거에 대해 접근할 수 있는 수단을 제공하는 방안과 이를 적용하기 위한 범위와 수준에 대해 토론하고 예시를 제시한다.

[투-03-01] 설명가능한 인공지능의 사용을 통해 사회적 신뢰 형성이 이루어질 수 있는 것을 이해하고 설명 가능한 인공지능을 보급하기 위해 노력하고 실천한다.

총 괄 조현석 | 서울특별시교육청 교육혁신과장

기 획 이현준 | 서울특별시교육청 장학관
박정희 | 서울특별시교육청 장학사

집 필 변순용 | 서울교육대학교 교수
개 발 김봉제 | 서울교육대학교 교수
김하민 | 월정초 교사
김태창 | 면남초 교사
김은수 | 인천가원초 교사
김진산 | 불광초 교사
박상아 | 진건초 교사
방은찬 | 의정부부용초 교사
이지형 | 서원초 교사
이혜령 | 군자초 교사

검 토 안성진 | 성균관대학교 교수
강윤지 | 교육연구정보원 교육연구사
김지혜 | 신서중 교사
김누리 | 세명초 교사
김재황 | 연가초 교사

발행일 2023년 2월

발행처 서울특별시교육청 교육혁신과(02-399-9059)

주 소 (03178)서울특별시 종로구 송월길 48

디자인 디투엔(02-2668-8880)

1. 이 자료집은 **서울특별시교육청 홈페이지** ▸ **교육청안내** ▸ **기구표** ▸ **교육혁신과** ▸ **부서업무방**에서 다운로드하실 수 있습니다.
2. 이 자료집은 교육적 목적으로 활용할 수 있으며, 기타 용도의 무단배포 및 복제를 금합니다.

서울특별시교육청