

# KEDI 미국원자력SMR 지수 산출방법 (Methodology Book)

Oct 2024

## 1. 지수 소개

- A. 본 문서는 KEDI 미국원자력SMR 지수(KEDI U.S. Nuclear Power SMR Index)를 산출하고 관리하는 방법론에 대한 이해를 돕기 위하여 작성되었습니다.
- B. KEDI 미국원자력SMR 지수는 미국 증권거래소에 상장된 종목들 중 Large Language Model (LLM) 기술을 공시보고서와 뉴스 자료에 적용하여 원자력과 SMR 산업에 관련된 키워드와 유사도가 점수가 높은 종목을 선정합니다. 이후 유동시가총액 가중 방식으로 비중을 구성한 지수입니다.
- C. 키워드를 기반으로 한 LLM적용은 기업의 재무실적과 직접적인 연관성이 높지 않을 수 있는 한계가 존재합니다. 따라서 분기별 실적 설명자료를 토대로 KEDI 지수위원회의 검토를 거쳐 연관성이 낮은 종목은 제외할 수 있으며, 이 경우 차 순위 종목이 최종 포트폴리오로 선정됩니다.
- D. KEDI 미국원자력SMR 지수는 가격지수(Price Return) 입니다

| 구분     | 기준일        | 기준지수    | 발표일         | 산출시간         | 리밸런싱주기 |
|--------|------------|---------|-------------|--------------|--------|
| 일반전략지수 | 2019.01.04 | 1,000pt | 2024.11. 01 | 8:30AM (KST) | 분기별    |

## 2. 지수 구성 방법

### A. 기초 투자 유니버스 선정

KEDI 미국원자력SMR 지수의 유니버스를 구성하기 위하여, 다음과 같은 요건에 부합하는 기초 투자 유니버스를 선정합니다.

- 미국 증권거래소(뉴욕증권거래소, 나스닥 등)에 상장된 보통주
- 정기변경일 기준 유동 시가총액 \$100mn 이상인 기업
- 정기변경일 기준 3개월 일 평균 거래대금이 \$1mn 이상인 기업
- KEDI-Akros Industry Classification Standard (KAICS) 분류 기준 다음 분류에 해당하는 기업
  - **Primary Level 5**
    - Nuclear Electric Power Generation
    - Other Metal Ore Mining
    - Industrial Valve Manufacturing
    - Fossil Fuel Electric Power Generation
  - **Secondary Level 5**
    - All Other Industrial Machinery Manufacturing

\* KEDI Akros Industry Classification Standard (KAICS)는 아크로스테크놀로지스에서 개발한 Akros Industry Classification Standard (AICS)를 기반으로, KEDI 지수위원회의 검토를 거쳐 조정된 분류 체계입니다. AICS는 미국 NAICS(North American Industry Classification System)의 산업분류 체계를 기반으로 기업 개요와 매출구성을 대규모 언어모델(LLM)을 활용하여 종목별로 구분한 업종 분류입니다

## B. 키워드 선정

키워드는 U.S. Department of Energy에서 발간된 Nuclear Energy: Supply Chain Deep Dive Assessment(2022.02.24)를 기반으로 원자력과 SMR 산업의 밸류체인과 차세대 기술을 반영하는 키워드를 선정하였습니다.

- Uranium Mining and Refining
- Uranium Enrichment
- Reactor Construction, Equipment, and Materials
- Nuclear Power Generation
- Small Modular Reactors (SMR)

키워드는 산업의 성장 및 변화 혹은 국가기관 및 연구소에서 인정하는 출처 자료의 변경에 따라 KEDI 지수위원회의 검토를 거쳐 수정될 수 있습니다.

## C. 데이터 수집

모든 데이터는 정기변경일을 기준으로 국가기관 자료, 사업보고서, 검색 및 뉴스 자료를 수집합니다. 특히 사업보고서의 경우 비중결정일을 기준으로 보통주에 대해서는 미국 증권거래위원회에 공시되어 있는 최신 날짜의 분기 보고서 (10-Q) 자료를 사용합니다. 단, 어떠한 사유로 분기 보고서 (10-Q) 자료가 존재하지 않는 경우, 연간 보고서 (10-K)로 대체하여 사용합니다.

## D. Large Language Model (이하 “LLM”) 기법을 통한 필터링

### i. 1차 Large Language Model (이하 “LLM”)을 이용한 필터링

기초 투자 유니버스에 해당하는 기업들에 대해서 국가기관 자료, 사업보고서, 검색 및 뉴스 자료를 기반으로, 모든 기업들이 각각의 키워드에 대해 얼마나 높은 유사도를 가지는지 LLM을 통하여 다음과 같이 측정합니다 (**절대적 기준의 유사도 점수**):

$$L(K_j, P_k) = \begin{cases} 1.0, & \text{기업정보와 키워드의 유사도가 "매우 높음"} \\ 0.75, & \text{기업정보와 키워드의 유사도가 "높음"} \\ 0.5, & \text{기업정보와 키워드의 유사도가 "중간"} \\ 0.25, & \text{기업정보와 키워드의 유사도가 "낮음"} \\ 0.0, & \text{기업정보와 키워드의 유사도가 "매우 낮음"} \end{cases}$$

|                  |                   |                                |
|------------------|-------------------|--------------------------------|
| L:               | LLM 기반 절대적 유사도 점수 | 개별 기업과 개별 키워드 간의 LLM 기반 유사도 점수 |
| K <sub>i</sub> : | 키워드               | 개별 키워드                         |
| P <sub>k</sub> : | 기업 정보             | 개별 기업에 대한 자료, 보고서, 검색 및 뉴스 정보  |

기초 투자 유니버스에 해당하는 모든 종목들에 대해서 1차 LLM 기법을 통하여 절대적 기준의 유사도 점수를 측정합니다. 이때, 절대적 기준의 유사도 점수가 0.5점 보다 낮은 점수로 측정된 기업들은 해당 기업 정보와 유사도가 “낮음” 혹은 “매우 낮음”으로 간주되어 투자 유니버스에서 제외됩니다.

## ii. 2차 Large Language Model 기법을 통한 스코어링

1차 LLM 스코어링(절대적 기준의 유사도 점수) 결과가 0.5점 이상인 기업들에 대해서 국가기관 자료, 사업보고서, 검색 및 뉴스 자료를 기반으로, 각각의 키워드에 대해 어떠한 기업이 다른 기업 대비 높은 유사도를 가지는지 LLM을 통하여 다음과 같이 측정합니다 (상대적 기준의 유사도 점수):

$$M(K_i, P_j, P_k) = \begin{cases} 1.0, & \text{기업 } j \text{와 기업 } k \text{ 를 비교했을 때 기업 } j \text{의 유사도가 기업 } k \text{에 비해 "높음"} \\ 0.0, & \text{기업 } j \text{와 기업 } k \text{ 를 비교했을 때 기업 } j \text{의 유사도가 기업 } k \text{에 비해 "낮음"} \end{cases}$$

|                  |                   |                            |
|------------------|-------------------|----------------------------|
| M:               | LLM 기반 상대적 유사도 점수 | 기업 대 기업 LLM 기반 키워드 유사도 점수  |
| K <sub>i</sub> : | 키워드               | 개별 키워드                     |
| P <sub>j</sub> : | 기업 정보             | 기업에 대한 자료, 보고서, 검색 및 뉴스 정보 |
| P <sub>k</sub> : | 기업 정보             | 기업에 대한 자료, 보고서, 검색 및 뉴스 정보 |

각 기업들이 첫 번째부터 마지막 순위에 이르기까지 모든 가능한 조합에 대한 상대적 비교 평가가 이루어집니다. 모든 비교를 통해 완성된 기업 우선순위에 대해 가장 낮은 점수의 기업은 0점, 가장 높은 점수의 기업은 1점으로 평가하며 상대 유사도 점수(이하 “LLM 상대 스코어”)로 합니다.

## E. 최종 편입 종목 선정

각 키워드 중 최소 하나에서 절대적 기준의 유사도 점수가 0.5점 이상인 기업을 최종 편입 종목으로 선정합니다. 만약 종목수가 10개 미만일 경우 유사도 점수 합계가 높은 차순위 종목 중 시가총액이 큰 기업을 우선 편입합니다.

키워드 기반의 LLM 기술로 선정된 기업이 원자력 및 SMR 산업과 재무적 연관성이 낮을 수 있습니다. KEDI 지수위원회는 최종 의사결정을 가진 기구로써, 분기별 실적 자료 등을 기반으로 연관성이 낮다고 판단되는 종목은 제외할 수 있으며, 이 경우 차 순위 종목이 최종 포트폴리오로 선정됩니다.

## F. 개별 종목의 편입 비중 산정 방식

각 기업별 Tier를 분류하여 아래표에 같이 비중을 부여합니다.

| 구분     | 기준                                                       | Tier 별 비중 | 개별종목 비중                                   | 비고                                                                     |
|--------|----------------------------------------------------------|-----------|-------------------------------------------|------------------------------------------------------------------------|
| Tier 1 | 키워드별 LLM 상대적 유사도 점수 1 위 종목                               | 50%       | 3 개월 평균<br>유동시가총액 가중비중<br>Cap 20%, Min 3% | 버퍼를: 정기변경 비중결정일에<br>기존 Tier1 종목이 Tier 1 기준을<br>충족하지 못하더라도 1 회 강등<br>유예 |
| Tier 2 | Tier1 에 선정된 종목을 제외하고<br>키워드별 LLM 상대적 유사도<br>점수가 가장 높은 종목 | 30%       | 3 개월 평균<br>유동시가총액 가중비중<br>Cap 10%         | -                                                                      |
| Tier 3 | 최종 편입 종목군에서 Tier1 과<br>Tier2 를 제외한 종목                    | 20%       | 3 개월 평균<br>유동시가총액 가중비중<br>Cap 5%          | -                                                                      |

동일 Tier 종목 선정에서 서로 다른 키워드에서 특정 종목이 중복 선정될 경우 하나의 키워드 종목으로 인정합니다.

### 3. 지수 산출 방법

#### A. 지수 값의 산출

KEDI 미국원자력SMR 지수는 다음 수식을 이용하여 유동 시가총액을 가중한 지수입니다.  
지수 산출 기준일은 2019년 01월 04일이며, 기준지수는 1,000으로 설정하였습니다.

$$I(t = 0) = 1000$$

$$I(t) = \frac{\sum_i P_{i,t} \times Q_{i,t}}{\sum_i P_{i,t-1} \times Q_{i,t-1}} \times I(t - 1)$$

|             |                   |
|-------------|-------------------|
| t           | 특정 시점 t           |
| I(t=0):     | 지수 I의 기준 값 (1000) |
| I(t):       | t 시점 지수 I의 산출 값   |
| I(t-1):     | t-1 시점 지수 I의 산출 값 |
| $P_{i,t}$ : | t 시점 종목 i의 종가     |
| $Q_{i,t}$ : | t 시점 종목 i의 보유량    |

#### 세부내용

위 수식을 이용하여 t 시점에 대한 지수 I의 값을 산출할 수 있습니다. 지수 I의 값에 대한 산출은 장 종료 시점(t-1)과 다음 영업일 개장 시점(t) 사이에 발생하는 종목의 시장가치 변화를 활용하여 이루어집니다. 따라서 해당 수식은 주식 분할 등 시장 중립적인 이벤트에 영향을 받지 않습니다.

## 4. 지수 관리 및 유지

### A. 정기변경

- i. 모든 영업일은 미국 주식시장 영업일을 기준으로 합니다.
- ii. 비중결정일 (Determination Date): 분기별 마지막 영업일 D를 기준으로 종목을 선정하며, 지수의 비중을 확정합니다.
- iii. 수행일 (Implementation Date): D+3 영업일 장 종료 시점에 정기 변경을 수행합니다.

### B. 수시변경

합병/분할, 상장폐지 등 특정 기업 이벤트 발생에 대한 수시변경은 일반적으로 「기업 이벤트 적용 방법론」을 따릅니다. 영업 손실 등 운용 상의 중대한 사유가 있을 경우 KEDI 지수 위원회의 검토를 거쳐 수시변경을 수행할 수 있습니다.

---

#### <책임의 한계와 법적고지>

이 문서를 포함한 KEDI(한국경제신문지수)의 지수 콘텐츠는 (주)한국경제신문과 서면동의 없이 복제, 전송 또는 배포할 수 없습니다. KEDI의 지수 콘텐츠는 정보 제공이 목적이며 산출 및 공표와 관련한 정확성 및 완전성을 보증하지 않습니다. (주)한국경제신문은 KEDI 지수를 기반으로 한 제 3자의 투자 상품에 대한 투자 의견을 표명하지 않으며, 지수 이용자와 제 3자간에 서비스를 매개하여 발생한 분쟁에 관여할 법적 의무가 없습니다. 또한 투자행위 등으로 발생한 손실 등에 대하여 책임을 부담하지 않습니다.