

Annex 22: 인공 지능(Artificial Intelligence)

Reasons for changes: Not applicable (new annex).

변경 이유: 해당 없음(신규 부록)

Document map

목차

1. 적용 범위(Scope)
 2. 원칙(Principles)
 3. 예정 용도(Intended Use)
 4. 허용 기준(Acceptance Criteria)
 5. 시험 데이터(Test Data)
 6. 시험 데이터 독립성(Test Data Independency)
 7. 시험 실시(Test Execution)
 8. 설명 가능성(Explainability)
 9. 신뢰도(Confidence)
 10. 운영(Operation)
- 용어 정의(Glossary)

1. 적용 범위(Scope)

This annex applies to all types of computerised systems used in the manufacturing of medicinal products and active substances, where Artificial Intelligence models are used in critical applications with direct impact on patient safety, product quality or data integrity, e.g. to predict or classify data. The document provides additional guidance to Annex 11 for computerised systems in which AI models are embedded.

활성 성분과 의약품 제조에 사용되며 환자 안전, 제품 품질 또는 데이터 완전성에 직접적으로 영향을 미치는 중요 용도(예, 데이터 예측 또는 분류)로 인공 지능 모델이 사용되는 모든 종류의 컴퓨터화 시스템에 이 부록이 적용된다. AI 모델이 내장된 컴퓨터화 시스템과 관련하여 부록 11을 보완하는 추가 가이드라인을 이 문서에서 제시한다.

The document applies to machine learning (AI/ML) models which have obtained their functionality through training with data, rather than being explicitly programmed. Models may consist of several individual models, each automating specific process steps in GMP.

명백하게 프로그래밍된 것이 아니라 데이터로 학습하여 기능을 확보한 기계 학습(AI/ML) 모델에 이 문서가 적용된다. 각기 특정 GMP 프로세스 단계를 자동화하는 여러 모델로 모델이 구성될 수 있다.

The document applies to static models, i.e. models that do not adapt their performance during use by incorporating new data. The use of dynamic models which continuously and automatically learn and adapt performance during use, is not covered by this document, and should not be used in critical GMP applications.

사용 도중에 새로운 데이터를 통합하여 성능을 조정하지 않는 모델인 정적 모델에 이 문서를 적용한다. 사용 도중에 지속적이고 자동적으로 학습하고 성능을 조정하는 동적 모델은 이 문서의 대상이 아니고, 이러한 모델을 중요 GMP 용도로 사용해서는 안 된다.

The document applies to models with a deterministic output which, when given identical inputs, provide identical outputs. Models with a probabilistic output which, when given identical inputs, might not provide identical outputs are not covered by this document and should not be used in critical GMP applications.

동일한 입력 시에 동일하게 출력되는 결정론적 출력 모델에 이 문서가 적용된다. 동일한 입력 시에 동일하게 출력되지 않을 수 있는 확률적 출력 모델은 이 문서의 대상이 아니며, 이러한 모델을 중요 GMP 용도로 사용해서는 안 된다.

Following the above, the document does not apply to Generative AI and Large Language Models (LLM), and such models should not be used in critical GMP applications. If used in non-critical GMP applications, which do not have direct impact on patient safety, product quality or data integrity, personnel with adequate qualification and training should always be responsible for ensuring that the outputs from such models are suitable for the intended use, i.e. a human-in-the-loop (HITL) and the principles described in this document may be considered where applicable.

이에 따라 이 문서는 생성형 AI와 LLM에 적용되지 않으며, 이와 같은 모델을 중요 GMP 용도로 사용해서는 안 된다. 환자 안전, 제품 품질 또는 데이터 완전성에 직접적인 영향이 없는 비중요 GMP 용도로 사용한다면, 적절한 자격을 갖추고 교육 훈련을 거친 작업자가 상기 모델의 출력 결과가 예정 용도에 적합함을 보증해야 한다. 해당되는 경우에 "사람 개입 방식"(HITL)과 이 문서에 기술된 원칙을 고려할 수 있다.

2. 원칙(Principles)

- 2.1. Personnel. In order to adequately understand the intended use and the associated risks of the application of an AI model in a GMP environment, there should be close cooperation between all relevant parties during algorithm selection, and model training, validation, testing and operation. This includes but may not be limited to process subject matter experts (SMEs), QA, data scientists, IT, and consultants. All personnel should have adequate qualifications, defined responsibilities and appropriate level of access.

작업자. GMP 환경에서 사용하는 AI 모델의 예정 용도와 관련 리스크를 적절하게 이해하기 위하여, 알고리즘 선정과 모델 훈련, 밸리데이션, 시험, 운영 시에 모든 관련자가 긴밀하게 협력해야 한다. 프로세스 SME, QA, 데이터 과학자, IT, 컨설턴트를 포함하되 이에 국한되지 않는다. 모든 작업자가 적절한 자격을 갖추고 작업자의 책임을 규정하며 접근 수준을 적절하게 설정한다.

- 2.2. Documentation. Documentation for activities described in this section should be available and reviewed by the regulated user irrespective of whether a model is trained, validated and tested in-house or whether it is provided by a supplier or service provider.

문서 관리. 모델 훈련, 밸리데이션, 시험을 누가 하건(자체적으로 또는 공급업체나 서비스 제공업체), 이 섹션에 기술된 활동에 관한 문서를 규제 대상 사용자가 확보하고 검토한다.

- 2.3. Quality Risk Management Activities described in this document should be implemented based on the risk to patient safety, product quality and data integrity.
환자 안전, 제품 품질, 데이터 완전성 관련 리스크를 토대로, 이 문서에 기술된 품질 리스크 관리 활동을 추진한다.

3. 예정 용도(Intended Use)

- 3.1. Intended use. The intended use of a model and the specific tasks it is designed to assist or automate should be described in detail based on an in-depth knowledge of the process the model is integrated in. This should include a comprehensive characterisation of the data the model is intended to use as input and all common and rare variations; i.e. the input sample space. Any limitations and possible erroneous and biased inputs should be identified. A process subject matter expert (SME) should be responsible for the adequacy of the description, and it should be documented and approved before the start of acceptance testing.

예정 용도. 모델이 통합된 프로세스에 대한 심층적인 지식을 바탕으로, 모델의 예정 용도와 모델이 지원하거나 자동화하도록 설계된 특정 태스크를 자세히 기술한다. 모델이 입력 요소로 사용하는 데이터, 그리고 모든 일반 편차와 희귀 편차(즉, 입력 샘플 스페이스)의 포괄적인 특성 평가 결과를 포함하여 기술한다. 한계와 입력 요소의 오류/편향 가능성도 파악한다. 프로세스 SME가 기술된 정보의 적절성을 책임지고, 승인 시험을 시작하기 전에 문서화하여 승인한다.

- 3.2. Subgroups. Where applicable, the input sample space should be divided into subgroups based on relevant characteristics. Subgroups may be defined by characteristics like the decision output (e.g. 'accept' or 'reject'), process specific baseline characteristics (e.g. geographical site or equipment), specific characteristics in material or product, and characteristics specific to the task being automated (e.g. types and severity of defects).

세부 그룹. 해당되는 경우에는 관련 특성을 바탕으로 입력 샘플 스페이스를 세부 그룹으로 나눈다. 의사 결정 출력 결과(예, "승인" 또는 "기각"), 프로세스 특이적 베이스라인 특성(예, 지리적 장소나 설비), 물품 또는 제품의 특정 특성, 자동화 대상 태스크 관련 특성(예, 결함 종류와 심각성) 등 특성별로 세부 그룹을 정할 수 있다.

- 3.3. Human-in-the-loop. Where a model is used to give an input to a decision made by a

human operator (human-in-the-loop), and where the effort to test such model has been diminished, the description of the intended use should include the responsibility of the operator. In this case, the training and consistent performance of the operator should be monitored like any other manual process.

HITL. 사람 작업자가 내리는 결정에 필요한 입력 요소를 모델이 제공하고(사람 개입 방식), 이와 같은 모델의 시험 활동이 축소된 경우, 작업자의 책임을 포함하여 예정 용도를 기술한다. 이 경우에 다른 수동 프로세스와 마찬가지로 작업자 교육 훈련과 실행 일관성을 모니터링한다.

4. 허용 기준(Acceptance Criteria)

- 4.1. Test metrics. Suitable, case dependent test metrics, should be defined to measure the performance of the model according to the intended use. As an example, suitable test metrics for a model used to classify products (e.g. 'accept' or 'reject') may include, but may not be limited to, a confusion matrix, sensitivity, specificity, accuracy, precision and/or F1 score.

시험 지표. 케이스 의존적 시험 지표를 적합하게 규정하고, 예정 용도를 바탕으로 모델 성능을 평가한다. 예를 들어 제품을 분류(예, "승인" 또는 "기각")하는데 사용되는 모델의 적합한 시험 지표로 혼동 행렬, 민감성, 특이성, 정확성, 정밀성 및/또는 F1 점수가 있으나 이에 국한되지 않는다.

- 4.2. Acceptance criteria. Acceptance criteria for the defined test metrics should be established by which the performance of the model should be considered acceptable for the intended use. The acceptance criteria may differ for specific subgroups within the intended use. A process subject matter expert (SME) should be responsible for the definition of the acceptance criteria, which should be documented and approved before the start of acceptance testing.

허용 기준. 모델의 성능이 예정 용도에 적합하다고 판단되는 허용 기준을 지정 시험 지표에 대해 설정한다. 예정 용도의 특정 세부 그룹에 따라 허용 기준이 다를 수 있다. 프로세스 SME가 허용 기준을 규정하고, 승인 시험을 시작하기 전에 허용 기준을 문서화하고 승인한다.

- 4.3. No decrease. The acceptance criteria of a model, should be at least as high as the performance of the process it replaces. This implies, that the performance should be known for the process which is to be replaced by a model (see Annex 11 2.7).

성능 기준 저하 금지. 모델이 대체하는 기존 프로세스의 성능 이상으로 모델의 허용 기준을 설정한다. 그러므로 모델이 대체하고자 하는 프로세스의 성능을 알고 있어야 한다(부록 11의 2.7 참조).

5. 시험 데이터(Test Data)

- 5.1. Selection. Test data should be representative of and expand the full sample space of the intended use. It should be stratified, include all subgroups, and reflect the limitations, complexity and all common and rare variations within the intended use of the model. The criteria and rationale for selection of test data should be documented.
선정. 예정 용도의 전체 샘플 스페이스를 대표하고 확장한 시험 데이터를 선정한다. 모든 세부 그룹을 포함하고 예정 용도의 모든 일반/희귀 편차와 복잡성, 한계를 반영하며 계층화하여 선정한다. 시험 데이터 선정 기준과 근거를 문서화한다.
- 5.2. Sufficient in size. The test dataset, and any of its subgroups, should be sufficient in size to calculate the test metrics with adequate statistical confidence.
충분한 규모. 시험 데이터 세트와 세부 그룹은 적절한 통계적 신뢰도로 시험 지표를 계산하기에 충분한 규모여야 한다.
- 5.3. Labelling. The labelling of test data should be verified following a process that ensures a very high degree of correctness. This may include independent verification by multiple experts, validated equipment or laboratory tests.
라벨링. 매우 높은 정확도를 보증하는 프로세스에 따라 시험 데이터의 라벨링을 확인한다. 여러 전문가의 독립적인 확인, 밸리데이션된 설비 사용 또는 시험실 시험으로 확인할 수 있다.
- 5.4. Pre-processing. Any pre-processing of the test data, e.g. transformation, normalisation, or standardisation, should be pre-specified and a rationale should be provided, that it represents intended use conditions.
전처리. 시험 데이터 전처리(예, 변환, 정규화, 표준화) 방법을 미리 규정하고 근거를 제시한다. 예정 사용 조건을 반영하여 정한다.
- 5.5. Exclusion. Any cleaning or exclusion of test data should be documented and fully justified.
배제. 시험 데이터 정제 또는 배제를 문서화하고 타당성을 충분하게 제시한다.

- 5.6. Data generation. Generation of test data or labels, e.g. by means of generative AI, is not recommended and any use hereof should be fully justified.

데이터 생산. 시험 데이터 또는 라벨 생산(예, 생성형 AI를 사용해 생산)은 권장되지 않으며, 이런 데이터를 사용한다면 타당성을 충분하게 제시한다.

6. 시험 데이터 독립성(Test Data Independency)

- 6.1. Independence. Effective measures consisting of technical and/or procedural controls should be implemented to ensure the independency of test data, i.e. that data which will be used to test a model, is not used during development, training or validation of the model. This may be by capturing test data only after completion of training and validation, or by splitting test data from a complete pool of data before training has started.

독립성. 기술적 및/또는 절차적 관리로 구성된 효과적인 대책을 구비하여 시험 데이터의 독립성을 확보한다. 즉, 모델을 시험하는데 사용되는 데이터를 모델 개발, 훈련 또는 밸리데이션 단계에서 사용하지 않는다. 훈련과 밸리데이션을 완료한 다음에 시험 데이터를 확보하거나 훈련을 시작하기 전에 전체 데이터 풀에서 시험 데이터를 분할하는 방식으로 독립성을 확보할 수 있다.

- 6.2. Data split. If test data is split from a complete pool of data before training of the model, it is essential that employees involved in the development and training of the model have never had access to the test data. The test data should be protected by access control and audit trail functionality logging accesses and changes to these. There should be no copies of test data outside this repository.

데이터 분할. 모델 훈련에 앞서 전체 데이터 풀에서 시험 데이터를 분할한다면, 모델 개발과 훈련에 관여하는 작업자가 시험 데이터 접근하지 못하게 하는 것이 중요하다. 접근 관리와 감사 추적 기능(접근 및 변경 내역 기록)으로 시험 데이터를 보호한다. 이 저장소 이외에는 시험 데이터 사본이 없어야 한다.

- 6.3. Identification. It should be recorded which data has been used for testing, when and how many times.

확인. 어떤 데이터가 언제, 몇 번 시험에 사용되었는지 기록한다.

- 6.4. Physical objects. When test data originates from physical objects, it should be ensured,

that the objects used for the final test of the model have not previously been used to train or validate the model, unless features are independent.

물리적 대상. 물리적 대상에서 시험 데이터가 유래한 경우, 최종 모델 시험에 사용되는 물리적 대상은 기능이 독립적이지 않다면 모델 훈련이나 밸리데이션에 사용된 것이 아니어야 한다.

- 6.5. Staff independency. Effective procedural and/or technical controls should be implemented to prevent staff members who have had access to test data from being involved in training and validation of the same model. In organisations where it is impossible to maintain this independency, a staff member who might have had access to test data for a model, should only have access to training and validation of the same model when working together (in pair) with a colleague who has not had this access (4-eyes principle).

작업자 독립성. 효과적인 절차적 및/또는 기술적 관리 대책을 구비하여, 시험 데이터에 접근한 작업자가 동일 모델의 훈련과 밸리데이션에 관여하지 못하게 한다. 이러한 독립성을 유지할 수 없는 조직인 경우, 모델 시험 데이터에 접근했을 가능성이 있는 작업자는, 이와 같은 접근을 한 적이 없는 다른 작업자와 함께(짝을 지어) 동일 모델의 훈련과 밸리데이션에 관여한다(4개의 눈 원칙).

7. 시험 실시(Test Execution)

- 7.1. Fit for intended use. The test should ensure that a model is fit for intended use and is 'generalising well', i.e. that the model has a satisfactory performance with new data from the intended use. This includes detecting possible over- or underfitting of the model to the training data.

예정 용도 적합성. 모델이 예정 용도에 적합하고 "우수하게 일반화"(즉, 새로운 데이터로도 모델이 예정 용도에 맞는 적합한 성능을 나타냄)함을 시험을 통해 확인한다. 훈련 데이터에 과다 또는 과소 피팅될 가능성도 시험을 통해 검출한다.

- 7.2. Test plan. Before the test is initiated, a test plan should be prepared and approved. It should contain a summary of the intended use, the pre-defined metrics and acceptance criteria, a reference to the test data, a test script including a description of all steps necessary to conduct the test, and a description of how to calculate the test metrics. A process subject matter expert (SME) should be involved in developing the plan.

시험 계획. 시험을 시작하기 전에 시험 계획서를 작성하고 승인한다. 예정 용도, 사전 설정 지표와 허용 기준, 시험 데이터 정보, 시험 실시에 필요한 모든 단계를 포함한 시험 스크립트, 시험 지표 계산 방법을 요약하여 작성한다. 프로세스 SME가 계획서 작성에 관여한다.

- 7.3. Deviation. Any deviation from the test plan, failure to meet acceptance criteria, or omission to use all test data should be documented, investigated, and fully justified.
일탈. 시험 계획 일탈, 허용 기준 부적합, 모든 시험 데이터 사용 누락을 기록하고 조사하고 충분하게 타당성을 제시한다.

- 7.4. Test documentation. All test documentation should be retained along with the description of the intended use, the characterisation of test data, the actual test data, and where relevant, physical test objects. In addition, documentation for access control to test data and related audit trail records, should be retained similarly to other GMP documentation.

시험 문서. 예정 용도, 시험 데이터의 특성, 실제 시험 데이터, 그리고 관련성이 있을 때는 물리적 시험 대상에 대한 정보를 기술한 자료와 함께, 모든 시험 문서를 보관한다. 이외에도 시험 데이터 접근 관리 문서와 관련 감사 추적 기록도 다른 GMP 문서와 마찬가지로 방식으로 보관한다.

8. 설명 가능성(Explainability)

- 8.1. Feature attribution. During testing of models used in critical GMP applications, systems should capture and record the features in the test data that have contributed to a particular classification or decision (e.g. rejection). Where applicable, techniques like feature attribution (e.g. SHAP values or LIME) or visual tools like heat maps should be used to highlight key factors contributing to the outcome.

특성 기여도. 중요 GMP 용도로 사용되는 모델의 시험 시에, 특정 분류 또는 결정(예, 기각)에 기여한 시험 데이터의 특성을 파악하고 기록한다. 해당되는 경우에는 히트맵 같은 시각적 도구나 특성 기여도 기법(예, SHAP 값 또는 LIME)을 사용해 결과에 기여한 주요 요소를 파악한다.

- 8.2. Feature justification. In order to ensure that a model is making decisions based on relevant and appropriate features and based on risk, a review of these features should be part of the process for approval of test results.

특성 타당성. 적절한 관련 특성과 리스크를 토대로 모델이 의사 결정을 내리도록 보증하기 위하여, 시험 결과 승인 시에 특성을 검토한다.

9. 신뢰도(Confidence)

- 9.1. Confidence score. When testing a model used to predict or classify data, the system should, where applicable, log the confidence score of the model for each prediction or classification outcome.

신뢰도 점수. 데이터 분류나 예측에 사용되는 모델을 시험할 때, 해당되는 경우에 예측 또는 분류 결과별로 모델의 신뢰도 점수를 기록한다.

- 9.2. Threshold. Models used to predict or classify data should have an appropriate threshold setting to ensure predictions or classifications are made only when suitable. If the confidence score is very low, it should be considered whether the model should flag the outcome as 'undecided', rather than making potentially unreliable predictions or classifications.

기준값. 데이터 분류나 예측에 사용되는 모델에 기준값을 적절하게 설정하여, 적합할 때만 예측 또는 분류를 하도록 보증한다. 신뢰도 점수가 매우 낮으면, 모델이 신뢰할 수 없는 예측 또는 분류를 하기보다는 "미결정"으로 결과를 표시하는지 검토한다.

10. 운영(Operation)

- 10.1. Change control. A tested model, the system it is implemented in, and the whole process it is automating or assisting should be put under change control before it is deployed in operation. Any change to the model itself, the system, or the process in which it is used, including any change to physical objects the model is using as input, should be documented and evaluated to determine if the model needs to be retested. Any decision not to conduct such retest should be fully justified.

변경 관리. 시험 완료된 모델, 모델이 구축된 시스템, 모델이 자동화하거나 지원하는 전체 프로세스를 배치하여 운영하기 전에 변경 관리 대상에 포함시킨다. 모델 자체, 시스템 또는 해당 프로세스의 변경(모델이 입력 요소로 사용하는 물리적 대상의 변경 포함)을 문서화하고 평가하여, 모델을 다시 시험할 필요가 있는지 판단한다. 다시 시험하지 않기로 결정하는 경우에는 타당성을 충분하게 제시한다.

- 10.2. Configuration control. A tested model should be put under configuration control

before being deployed in operation, and effective measures should be used to detect any unauthorised change.

구성 관리. 시험 완료된 모델을 배치하여 운영하기 전에 구성 관리 대상에 포함시키고, 무허가 변경을 감지하는 효과적인 대책을 구비한다.

- 10.3. System performance monitoring. The performance of a model as defined by its metrics should be regularly monitored to detect any changes in the computerised system (e.g. deterioration or change of a lighting condition).

시스템 성능 모니터링. 지표를 토대로 모델 성능을 정기적으로 모니터링하여, 컴퓨터화 시스템의 변화(예, 조명 조건 저하 또는 변경)를 감지한다.

- 10.4. Input sample space monitoring. It should be regularly monitored whether the input data are still within the model sample space and intended use. Metrics should be defined for monitoring any drift in the input data.

입력 샘플 스페이스 모니터링. 입력 데이터가 계속해서 모델 샘플 스페이스와 예정 용도 이내인지 정기적으로 모니터링한다. 입력 데이터의 드리프트를 모니터링하기 위한 지표를 규정한다.

- 10.5. Human review. When a model is used to give an input to a decision made by a human operator (human-in-the-loop), and where the effort to test such model has been diminished, records should be kept from this process. Depending on the criticality of the process and the level of testing of the model, this may imply a consistent review and/or test of every output from the model, according to a procedure.

사람의 검토. 사람 작업자가 내리는 결정에 필요한 입력 요소를 모델이 제공하고(사람 개입 방식), 이와 같은 모델의 시험 활동이 축소된 경우, 이 프로세스의 기록을 구비한다. 프로세스의 중요도와 모델 시험 수준에 따라, 모델의 모든 출력을 절차에 의거하여 지속적으로 검토 및/또는 시험할 필요가 있음을 의미할 수 있다.

용어 정의(Glossary)

Artificial Intelligence – ‘AI system’ means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence

physical or virtual environments;

인공 지능 - 'AI 시스템'은 다양한 수준의 자율성을 갖고 작동하도록 설계되며, 구축 이후에 적응성을 나타낼 수 있으며, 명시적인 목적이나 암시적인 목적을 위해 시스템이 접수하는 입력 요소로부터 물리적 환경이나 가상 환경에 영향을 줄 수 있는 예측, 내용, 권고 또는 결정 등 출력을 생산하는 방법을 추론하는 기계 기반 시스템을 의미한다.

Deep Learning – Approach to creating rich hierarchical representations through the training of neural networks with many hidden layers

딥러닝 - 숨겨진 여러 레이어를 갖춘 신경망의 훈련을 통해 풍부한 계층적 표현을 생산하는 방법.

Feature – A pattern in data that can be reduced to a simpler higher-level representation

특성 - 더 단순한 상위 수준의 표현으로 전환될 수 있는 데이터의 패턴.

LIME – Local Interpretable Model-Agnostic Explanations; a technique that approximates any black box machine learning model with a local, interpretable model to explain each individual prediction.

LIME - 국소적 해석 가능한 모델 비종속 설명; 국소적 해석 가능한 모델로 블랙박스 기계 학습 모델을 근사하여 개별 예측 결과를 설명하는 기법.

Machine Learning – Machine learning refers to the computational process of optimising the parameters of a model from data, which is a mathematical construct generating an output based on input data. Machine learning approaches include, for instance, supervised, unsupervised and reinforcement learning, using a variety of methods including deep learning with neural networks.

기계 학습 - 기계 학습은 입력 데이터를 토대로 출력을 생산하는 수학적 구조인 모델의 파라미터를 데이터로 최적화하는 컴퓨팅 프로세스를 의미한다. 기계 학습 방식의 예로는 신경망 기반의 딥러닝을 포함해 다양한 방법을 사용하는 지도 학습, 비지도 학습, 강화 학습이 있다.

Model – Mathematical algorithms with parameters (weights) arranged in an architecture that allows learning of patterns (features) from training data

모델 - 훈련 데이터에서 패턴(특성)을 학습할 수 있는 아키텍처로 파라미터(가중치)를 구성한 수학적 알고리즘.

Overfitting – Learning details from training data that cannot be generalised to new data

오버피팅 - 새로운 데이터에 일반화할 수 없는 훈련 데이터의 세부 사항 학습.

SHAP – Shapley Additive Explanations; an explainable AI (XAI) framework that can provide model-agnostic local explainability for tabular, image, and text datasets

SHAP - 셰이플리 가산 설명; 표, 이미지, 텍스트 데이터 세트에 대하여 모델 독립적 국소적 설명을 제공할 수 있는 설명 가능한 AI(XAI) 프레임워크.

Static – Frozen model: A model where all parameters have been finally set, not allowing further adaption to new data.

정적 - 동결 모델; 모든 파라미터가 최종적으로 설정되어 새로운 데이터에 맞춰 추가로 조정되지 않는 모델.

Test dataset – The "hold-out" data that is used to estimate performance of the final ML model.

시험 데이터 세트 - 최종 ML 모델의 성능을 추정하는데 사용되는 "보류" 데이터.

Training dataset – The data used to train the ML model.

학습 데이터 세트 - ML 모델의 훈련에 사용되는 데이터.

Validation dataset (in AI) – The dataset used during model development, to inform on how to optimally train the model from training data. size smaller than the training set

밸리데이션 데이터 세트(AI) - 모델 개발 단계에서 사용되는 데이터 세트이며, 훈련 데이터로 모델을 최적으로 훈련하는 방법에 관한 정보를 제공하는 데이터 세트. 훈련 세트보다 규모가 작다.