**1.1 Clinical Problems and Demonstrators**

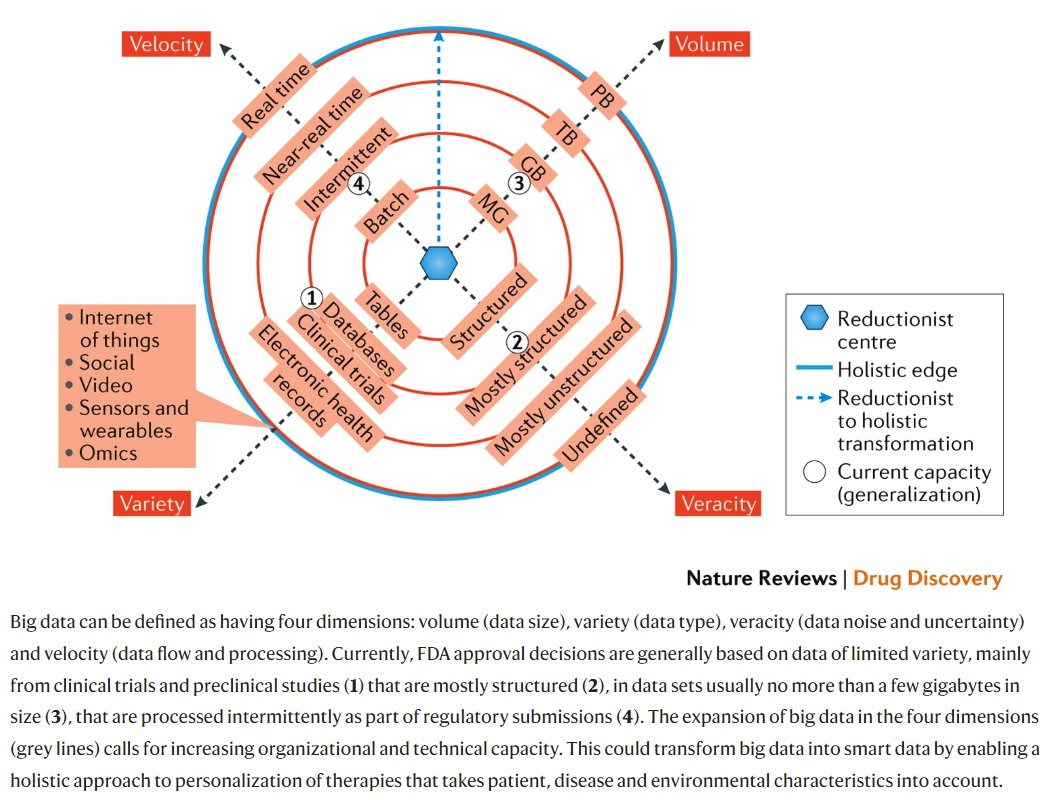
Dr Eoin Mckinney

**Outline**

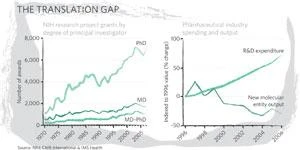
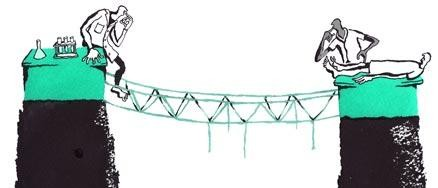
* Definitions
* The problem with translation
* Example of success?
* Progress, challenges, and opportunities
  + Deployment
  + Data access
  + Beyond supervised assistance
  + To interpret or not to care?
* Summary
* Q&A

**Definition of AI and ML**

* AI: capability of a machine to imitate intelligent human behaviour
* ML: the field of study that gives computers the ability to learn without explicitly being programmed

****

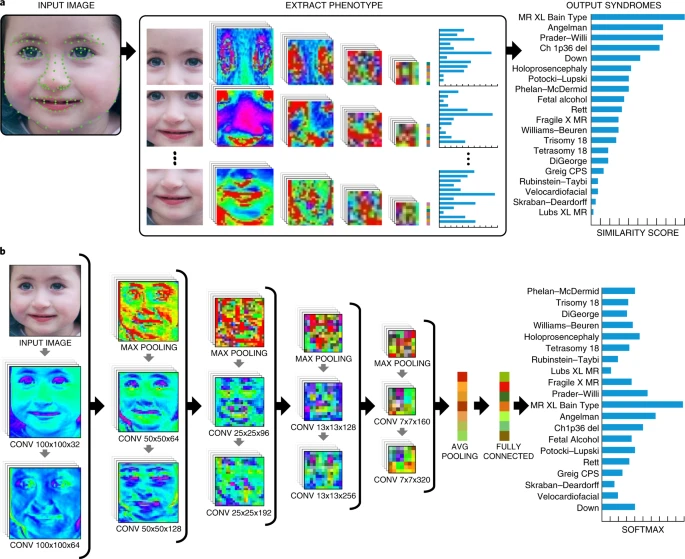
* [**https://www.nature.com/articles/nrd.2017.26**](https://www.nature.com/articles/nrd.2017.26)
* contexts in terms of application has changed drastically and will change more
* Moving outwards from the circle
  + data moving towards real-time information
  + New methods required
* Why are we still poised? 
* How can we change to realisation?
  + The AI chasm (Nature 453;12 2008) — Crossing the valley of death
    - <https://www.nature.com/articles/453840a>



* + Getting towards impact is difficult
    - Presentation will cover key aspects
    - e.g. not translatable in 5 to 10 years. no impact made following publications

**An example given: Identifying facial phenotypes of genetic disorders using deep learning**

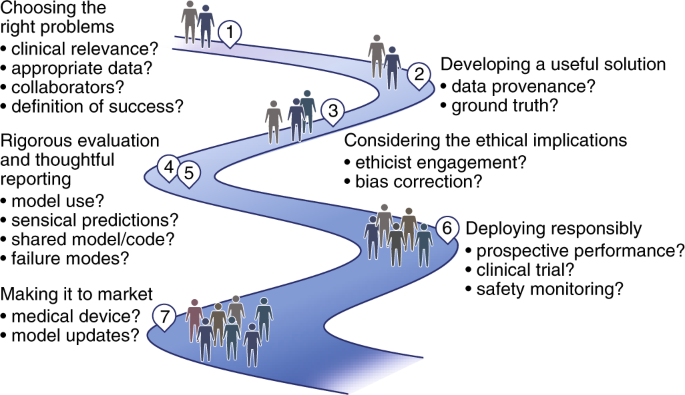
* <https://www.nature.com/articles/s41591-018-0279-0>



* DeepGestalt
  + Potentially adds considerable value to phenotypic evaluation in clinical genetics, genetic testing, research and precision medicine

**Road to impact**

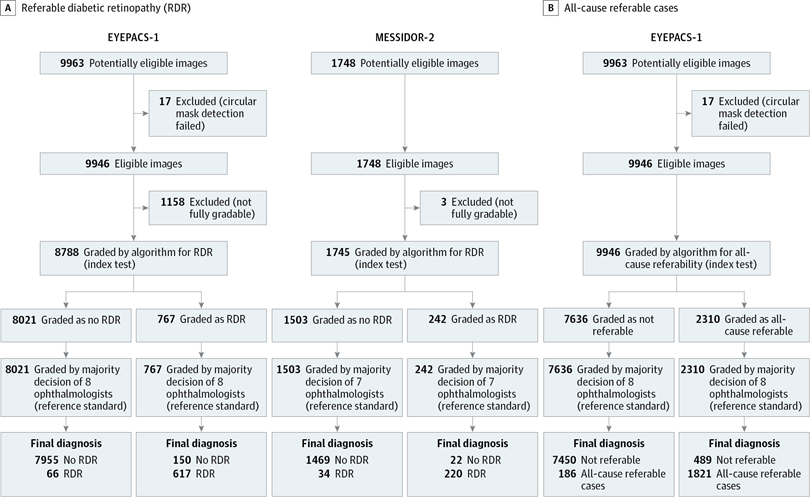
* <https://www.nature.com/articles/s41591-019-0548-6>

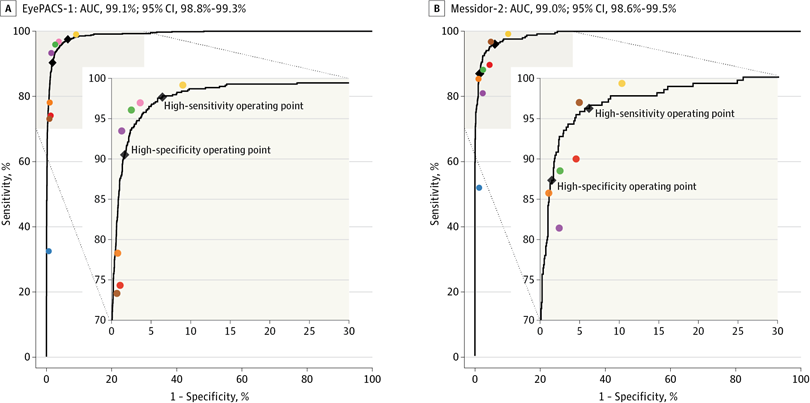


* 2. Useful model solution should result in
  + important impacts
* 4. Model use:
  + Critical consideration
* 6. Deploying responsibly:
  + Ensuring when something is deployed, it remains useful
* 7. Making it to market:
  + Up to date model

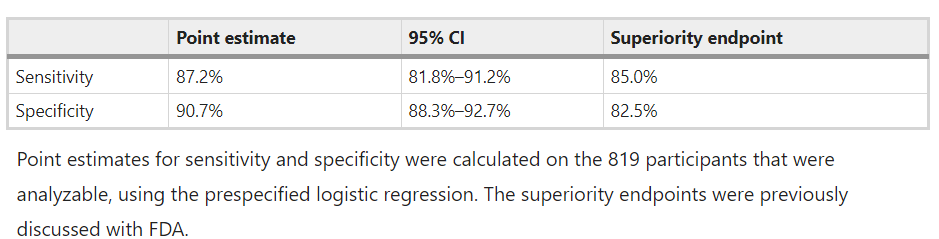
**Making it to market and staying up to date – Retinal photographs example:**

* <https://jamanetwork.com/journals/jama/fullarticle/2588763>

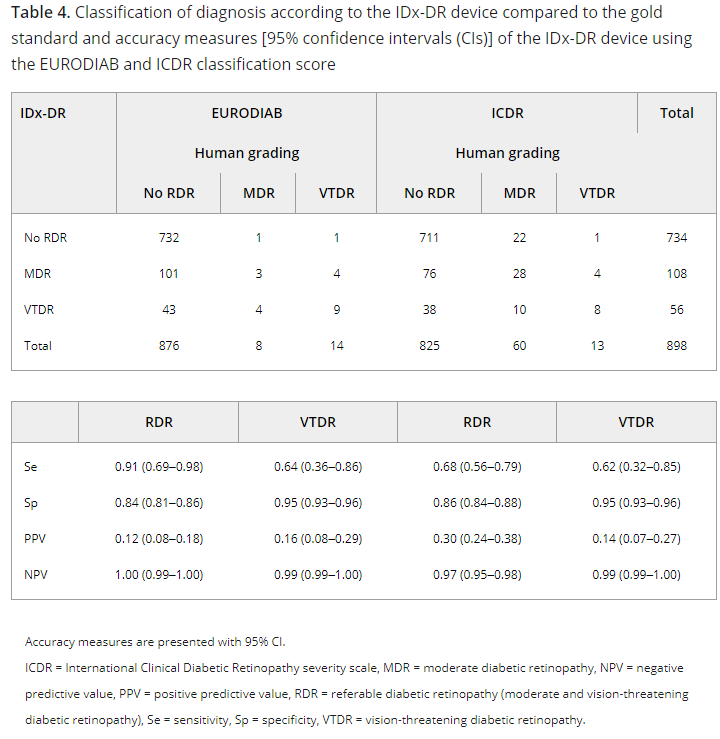
****

****

* CNN developed to train and diagnose diabetic retinopathy
* high sensitivity and specificity when compared to **expert opinions**
  + replacing medical individual with algorithms
  + still require camera and photo-taking
  + well defined goals and well annotated dataset (ground truth)
* Years after: pivotal trial on the model
  + **Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices**
  + <https://www.nature.com/articles/s41746-018-0040-6>



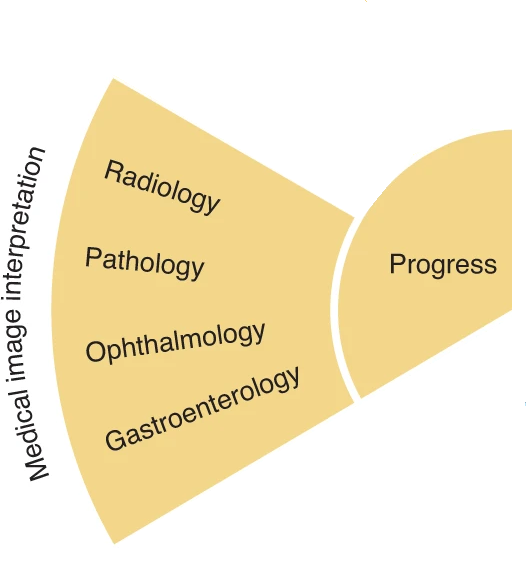
* + Requirements of FDA to define the superiority end-point
    - not algorithm better than human: does not have to be perfect
    - drop-outs of samples from data (10%): unreadable
      * not included in clinical trials, point estimates cleared superiority when added later on
    - pivotal trial: testing whether algorithm better than
      * does not consider economical impact
      * uncertain if the algorithm can make healthcare system effective
    - Further studies conducted years after: <https://onlinelibrary.wiley.com/doi/10.1111/aos.13613>



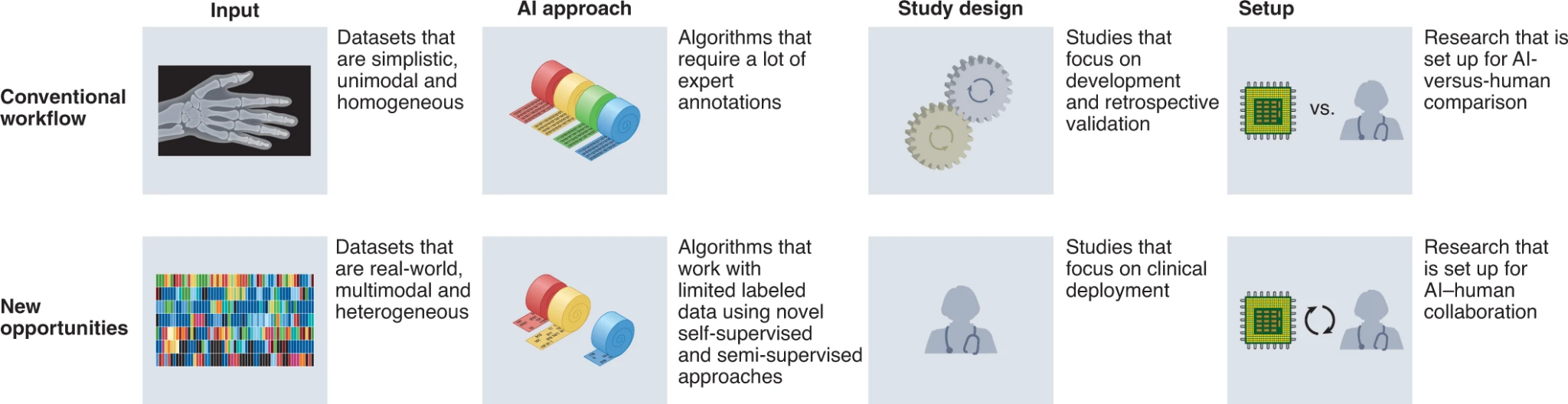
* + - Risk missing cases due to severe end-points,
      * sensitivity can be more important than specificity
      * drop of up to 40%, images not up to quality -- cost efficiency
  + In the end, Algorithm passed
    - unanswered questions:
      * Will it be used in hospitals?
      * Is it cost saving?
  + Some of these are due to reproducibility:
    - Independent replication shows low reproducibility
      * **Reproduction study using public data of: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs**
        + <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0217541>
      * **Evaluation of Artificial Intelligence–Based Grading of Diabetic Retinopathy in Primary Care**
        + <https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2703944>
    - Reproducibility: reproducing the results shared in the initial publication
      * Challenging
      * commercial sensitivity: no code sharing or dataset sharing
      * random seeding: may not be reproducibility
      * Data release changes
      * code version updates
      * inherent methods (stochastic descent)
    - Independent replication:
      * Substantially impacts from new data quality
      * Much more lower than published results
* Publication showing 35 highly-cited biomarker publications:
  + Comparison of Effect Sizes Associated With Biomarkers Reported in Highly Cited Individual Articles and in Subsequent Meta-analyses
  + <https://jamanetwork.com/journals/jama/article-abstract/900417>
  + Once reached to clinical deployment
    - performance is expected to reduce significantly

**Problems tackled**

* AI in health and medicine (<https://www.nature.com/articles/s41591-021-01614-0>)

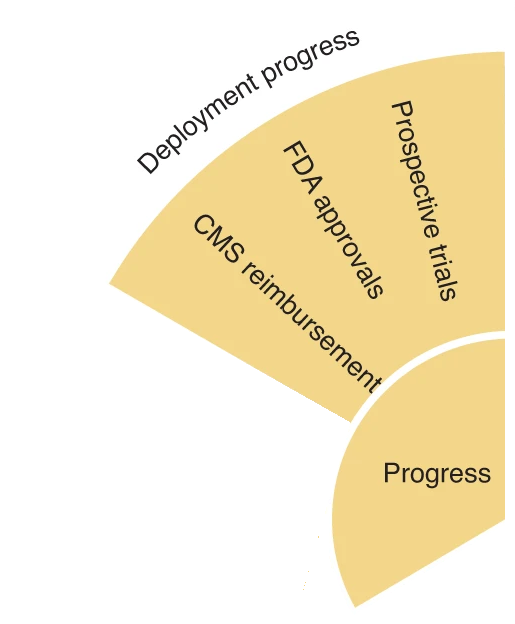


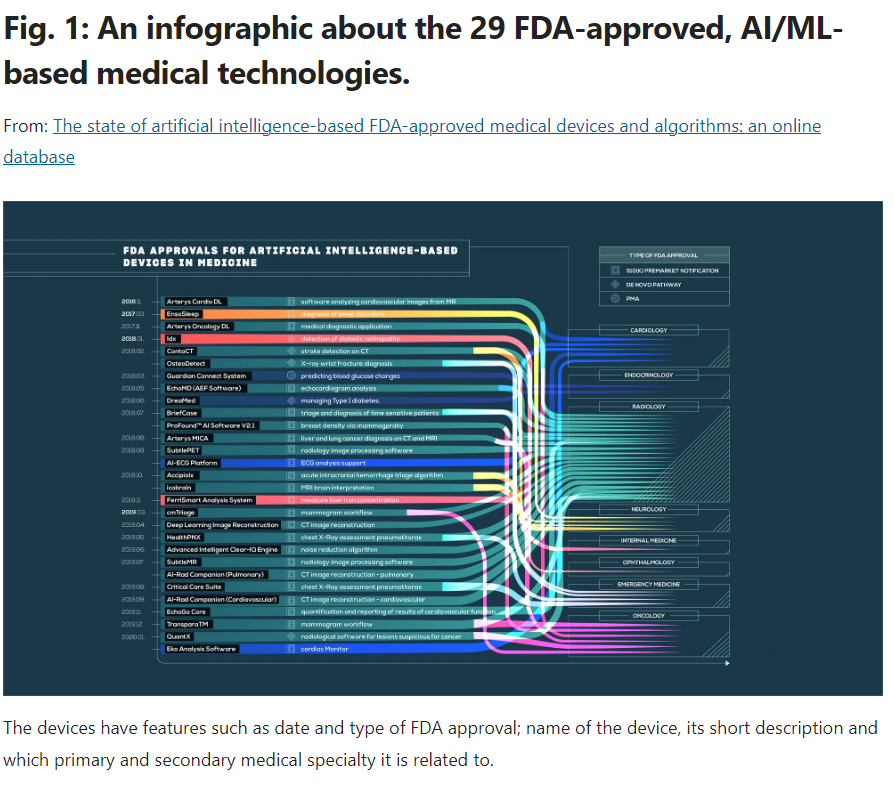
* Increasing success and usage in DL and ML methods, especially in medical imaging interpretation



* **Clinical impact and quality of randomized controlled trials involving interventions evaluating artificial intelligence prediction tools: a systematic review**
  + - <https://www.nature.com/articles/s41746-021-00524-2>

**Deployment**

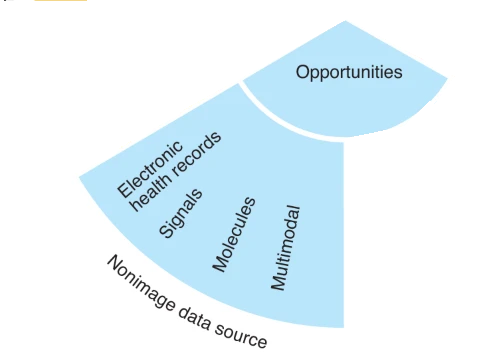
* <https://www.linkedin.com/pulse/fda-proposed-regulatory-framework-modifications-samd-adafinoaiei>
* 
* Orphan device assignment for FDA approval
* assisted device = low risk
* Comparison with human performance
* Stringent data quality selection

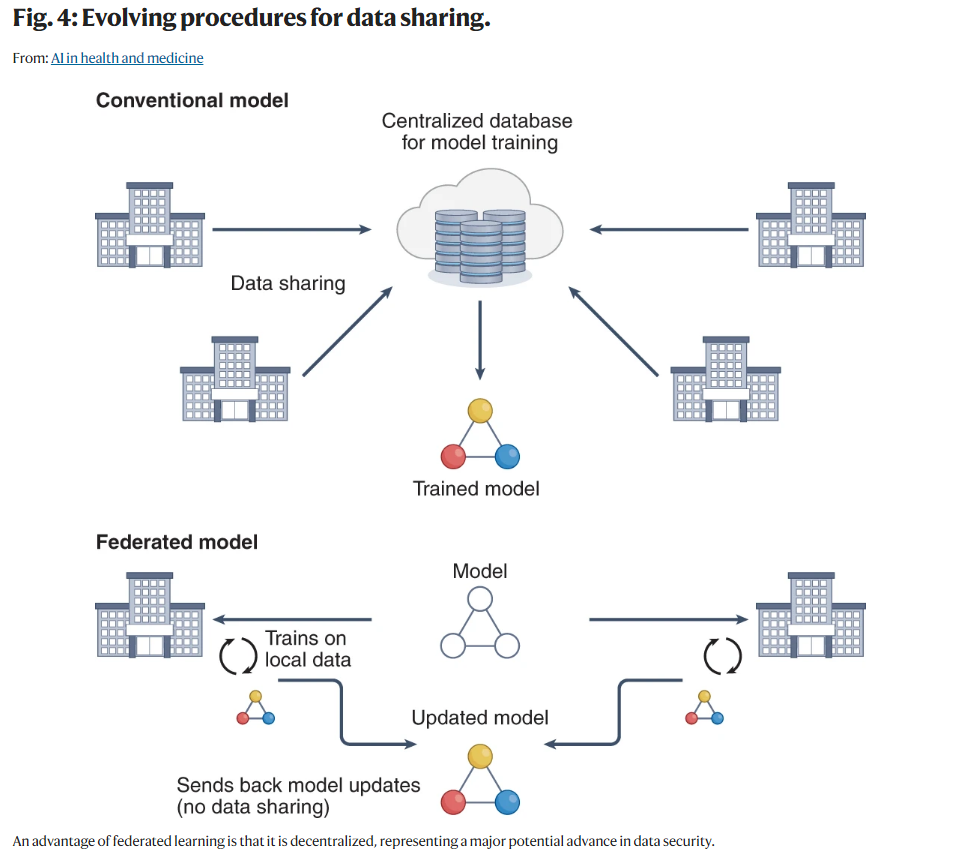


<https://www.nature.com/articles/s41746-020-00324-0>

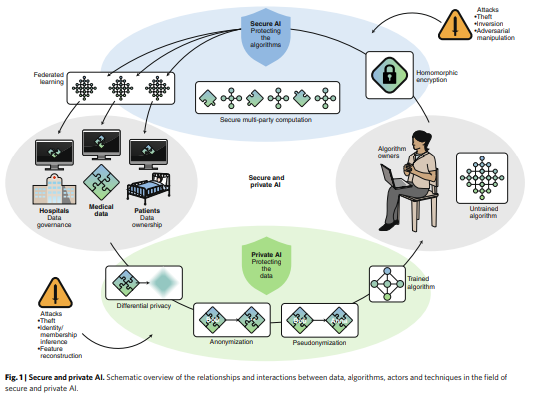
* + Much progress based on images, followed by cardiology (ECG)
    - approval is not enough, further demonstration of impact beyond approval

**Data access**

****

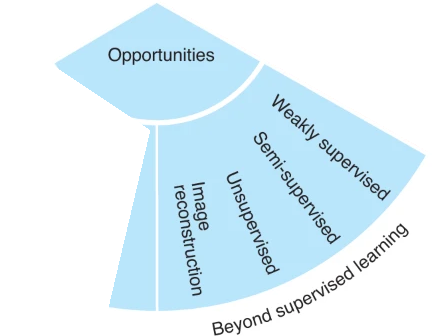


* + Historically, local learning and independently validated
    - replaced by central learning mechanism (data and parameters housed centrally)
    - challenges in:
      * collaboration
      * security
      * data ownership
    - Replaced by federated model

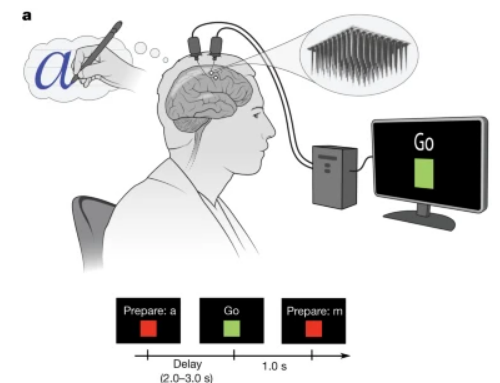


* + Further advances: swarm learning (block chain access)
    - <https://mediatum.ub.tum.de/doc/1602022/1602022.pdf>
    - <https://www.nature.com/articles/s41586-021-03583-3>

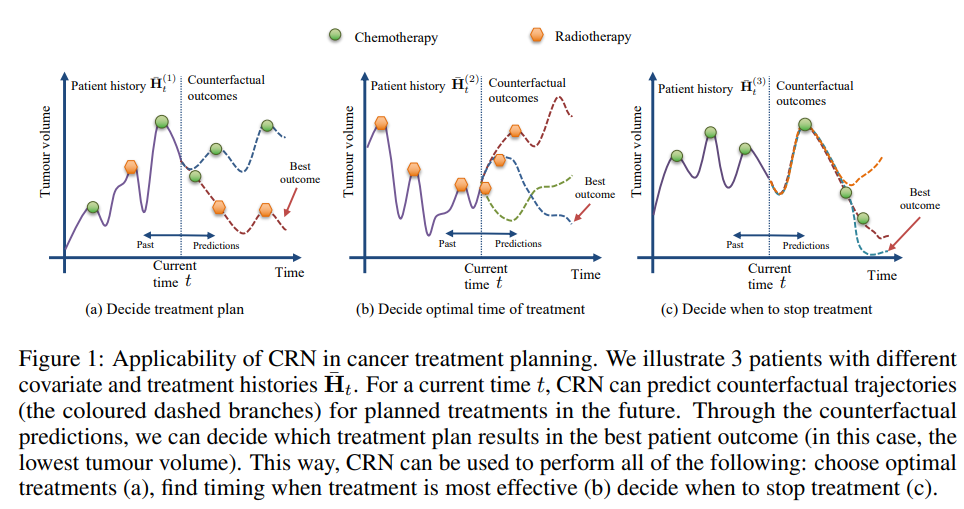
**Beyond supervised assistance**

****

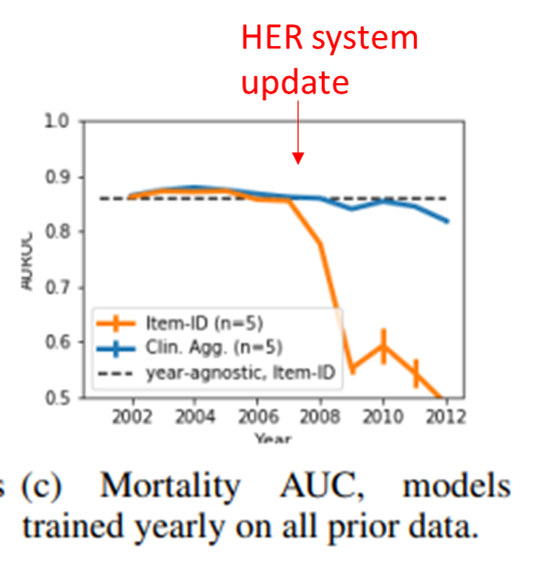
* **Example: High-performance brain-to-text communication via handwriting**
  + <https://www.nature.com/articles/s41586-021-03506-2>

a, To assess the neural representation of attempted handwriting, participant T5 attempted to handwrite each character one at a time, following the instructions given on a computer screen (bottom panels depict what is shown on the screen, following the timeline). Credit: drawing of the human silhouette created by E. Woodrum. 

* + rich longitudinal NN
    - RNN to extract and learn what letter has been written
    - time-series information with clear ground truth
* Learning from an entire health record
  + Complexity of different information increases greatly
    - Application of NN made better
      * e.g. death, length of stay
      * better supporting the healthcare system
  + Individual treatment effect inference
    - **ESTIMATING COUNTERFACTUAL TREATMENT OUTCOMES OVER TIME THROUGH ADVERSARIALLY BALANCED REPRESENTATIONS**
      * <https://arxiv.org/pdf/2002.04083.pdf>

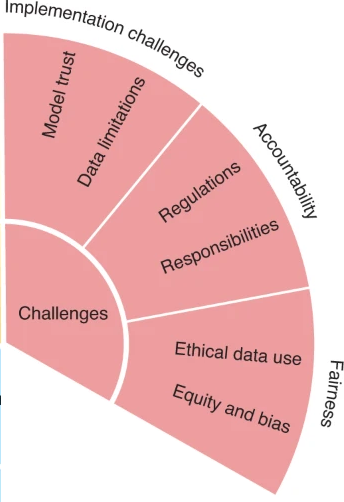


* + - Decide treatment plan, optimal time of treatment, when to stop treatment
    - Another example:
      * When treatment should be stopped,
* Rethinking clinical prediction: considering year of care and feature aggregation in ML
  + **Why machine learning must consider year of care and feature aggregation**
    - <https://arxiv.org/pdf/1811.12583.pdf>

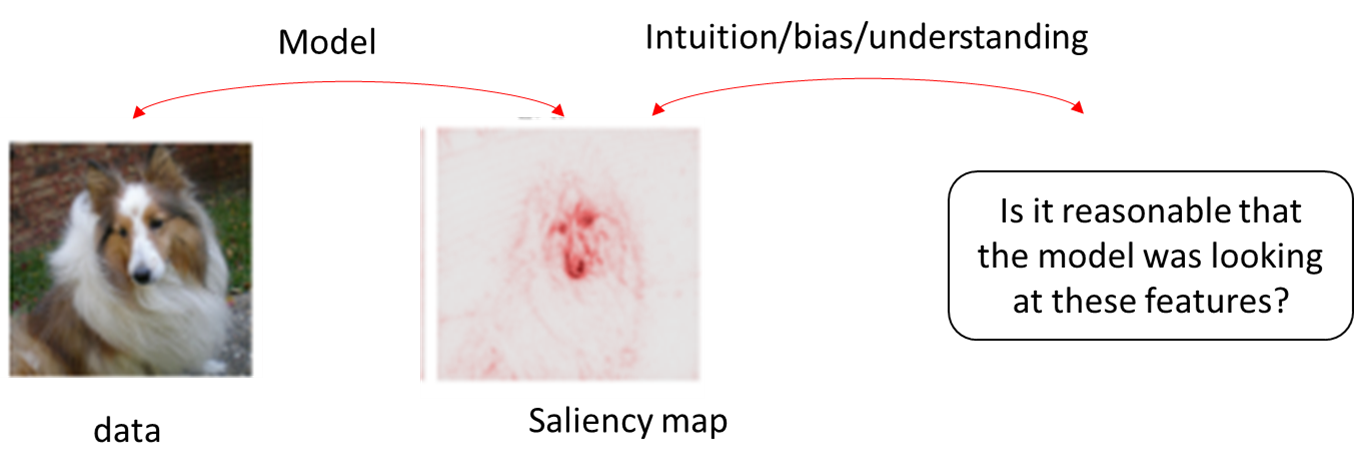
****

* + Dynamic system: changes with time
    - data changes dependent on the year and where collected
    - if learning is off-line at initial stage, the changes of healthcare system will lead to cliff-edge
      * constant learning
      * particular challenge at deployment

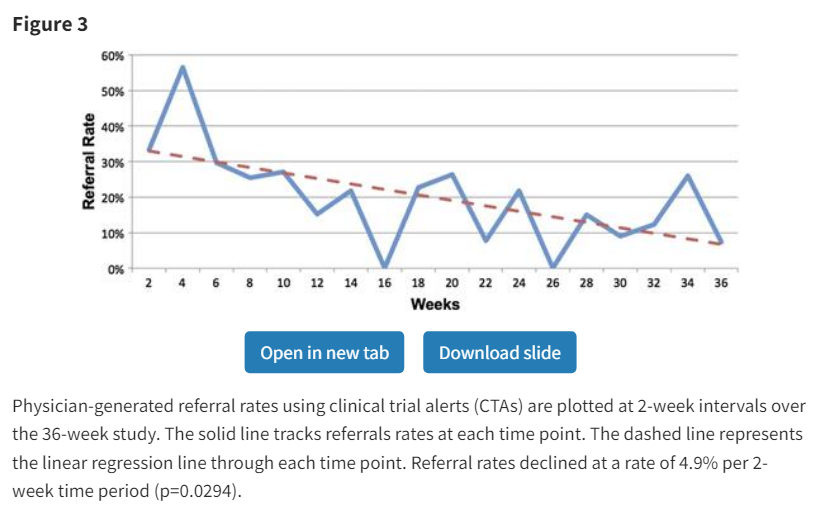
**Interpretability**

****

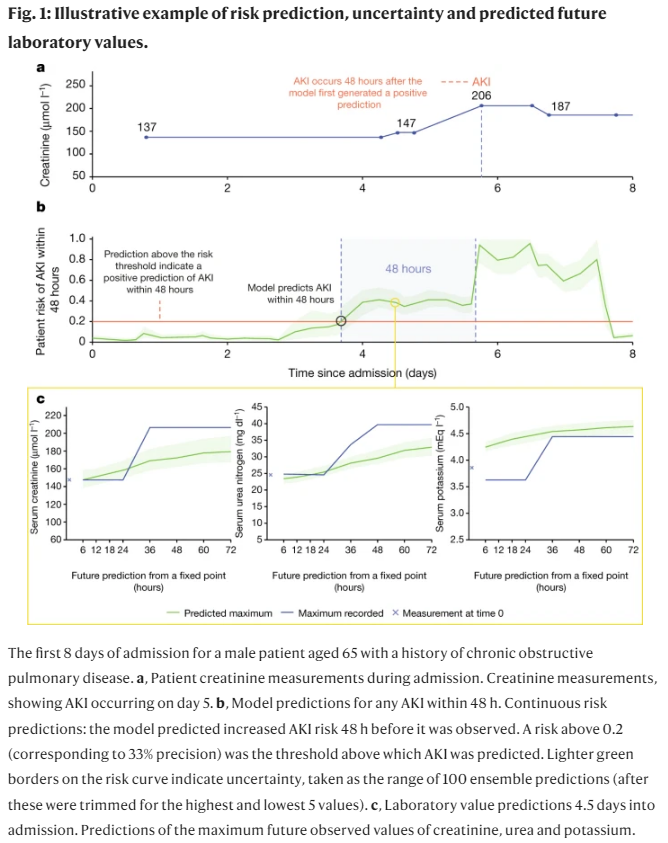
* Concepts of interpretability and why is it important
* Interpretability:
  + step 1: insights into decision-making process
  + step2 : mechanistic understanding of whether the process was sensible
    - e.g. saliency map
    - much harder to define and look at
    - potentially biased: based on your own intuition
    - many are not interpretable
  + e.g. feature of a dog (dog or cat?)



* What do clinicians want?
  + Blackbox models usually outperform transparent models
    - arguably should develop more blackbox?
  + Feature importance
    - Clinicians wants to understand the individual variances
  + Instance level explanations
    - i.e. what other instances should be considered from the model if two predictions are similar?
  + Temporal explanations?
    - how changes in longitudinal is factored in?
  + Transparent design
    - Not always available
  + Avoidance of alarm fatigue
    - as demonstrated previously
    - why the prediction is made thus understanding what risks the alarm fatigue



* + - **Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study**
    - <https://doi.org/10.1136/amiajnl-2011-000743>
* Example of alarm fatigue:

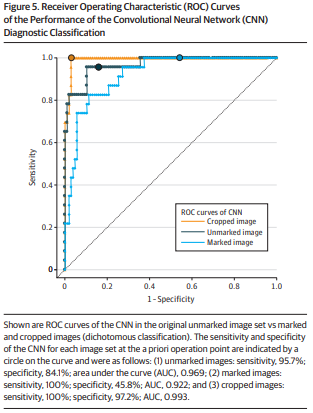
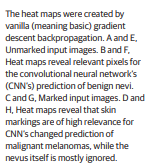
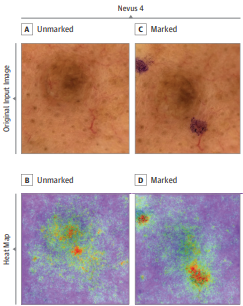


* + **A clinically applicable approach to continuous prediction of future acute kidney injury**
    - <https://www.nature.com/articles/s41586-019-1390-1>
  + model can predict kidney injury with 2 day lead time
  + ratio to true to false positive was 2:1
    - early stage of clinical performance, showing alarm fatigue
* A need to understand?
  + Can highly complex models be explained? do they need to be?
  + Legal perspective:
    - EU GDPR Article 13-15
    - must be interpretable: "meaningful information"
  + Risk of apparently reassuring saliency maps
    - saliency maps:
      * or feature weights
      * human needs to interpret the actual feature and quantifying it
  + Explaining explanations: quantifying interpretation?
  + risk of bias/ model debugging
* **Intuitive interpretation**

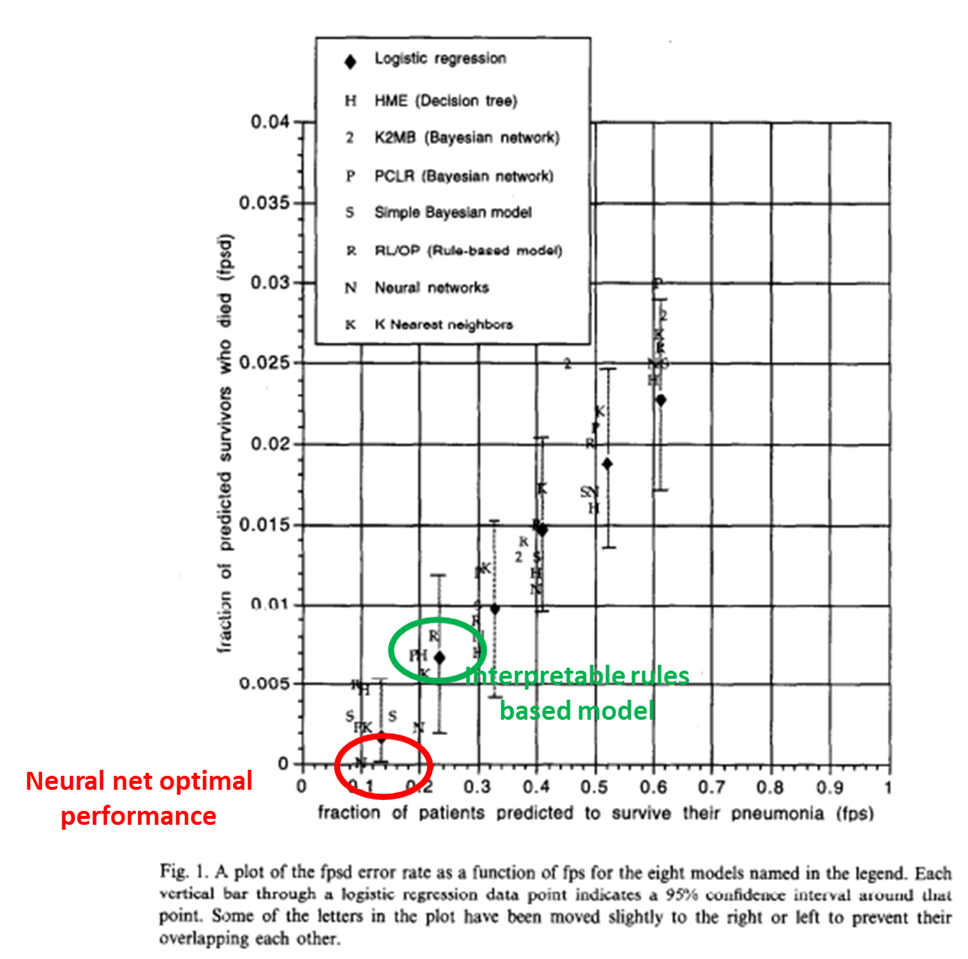


* + **Saliency Methods for Explaining Adversarial Attacks**
    - <https://www.semanticscholar.org/reader/d18a5a20e76d60206caee92b23db529329e34061>
  + should we be reassured if what the model is looking at “make sense”?
  + Saliency maps = make sense?
    - same features being looked at even reduced model accuracy
    - falsely reassured that the model is working as expected
  + “You could have many explanations for what a computer model is doing. Do you just pick the one you ‘want’ to be correct?” – Cynthia Rudin

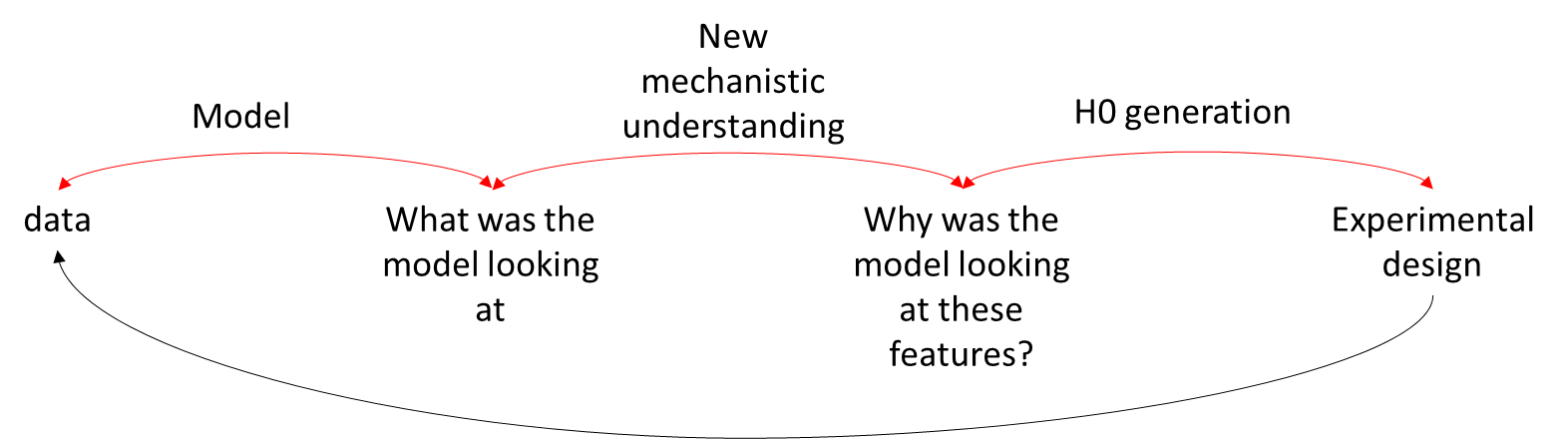
**Bias and debugging**



* **Association Between Surgical Skin Markings in Dermoscopic Images and Diagnostic Performance of a Deep Learning Convolutional Neural Network for Melanoma Recognition**
  + <https://jamanetwork.com/journals/jamadermatology/fullarticle/2740808>
* Model were markedly influenced by the skin marking on the marked lesions
  + falsely supporting diagnosis
* Other example risk due to model bias



* + **An evaluation of machine-learning methods for predicting pneumonia mortality**
    - <https://www.sciencedirect.com/science/article/pii/S0933365796003673>
    - the hospital data implied the rule “HasAsthma (x)⇒LowerRisk(x).”
    - data leak from asthmatic patients in the training dataset, not factored in
      * asthma was shown as low risk due to their chronic asthma treatment
  + **Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission**
    - <https://www.microsoft.com/en-us/research/wp-content/uploads/2017/06/KDD2015FinalDraftIntelligibleModels4HealthCare_igt143e-caruanaA.pdf>
    - Similarly
* **Clinical demonstrators to facilitate interpretation and debugging**
  + More information: [https://www.vanderschaar-lab.com/engagement-sessions/revolutionizing-healthcare/](https://ww.vanderschaar-lab.com/engagement-sessions/revolutionizing-healthcare/)
  + software will be showcased later in the week
    - clinical demonstrator allowing visualization of potential data leak for those unfamiliar to saliency maps
    - interpretable allowing debugging before clinical translation
* Interpretability in genomics
  + skipped
  + fundamental for interpretability, hypothesis generation rely on the results



* Summary
  + Definitions
  + Problem with translation
  + Example of successful models
  + Progress, remaining challenges, and opportunities
  + changes in infrastructure for deployment
  + Data accessibility
  + Beyond supervised assistance
  + interpretability very important for translation