Risk
modelling.research

DS projects

Skoltech, April 2021

Sberbank

Risk Modelling Function in Numbers

Data scientists (DS)

Over **200** DS in Risk Function

Model development: **120** DS

Model validation: **80** DS

Models

Approx. **500** models (inventory)

300+ annual model flow

(new / revised models)

What should be taken into account:
the difference between data driven projects and other IT projects

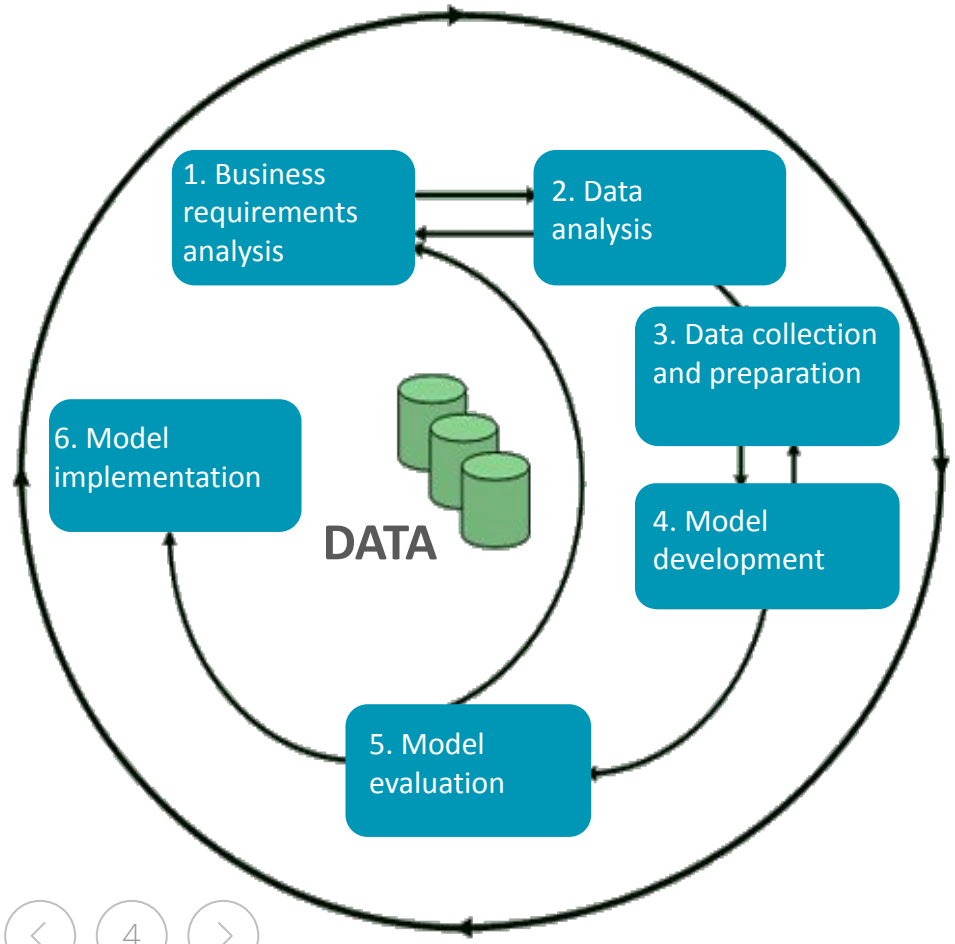
Data is required

Uncertainty of results

Hypothesis are tested

Model lifecycle and DS product

CRISP-DM

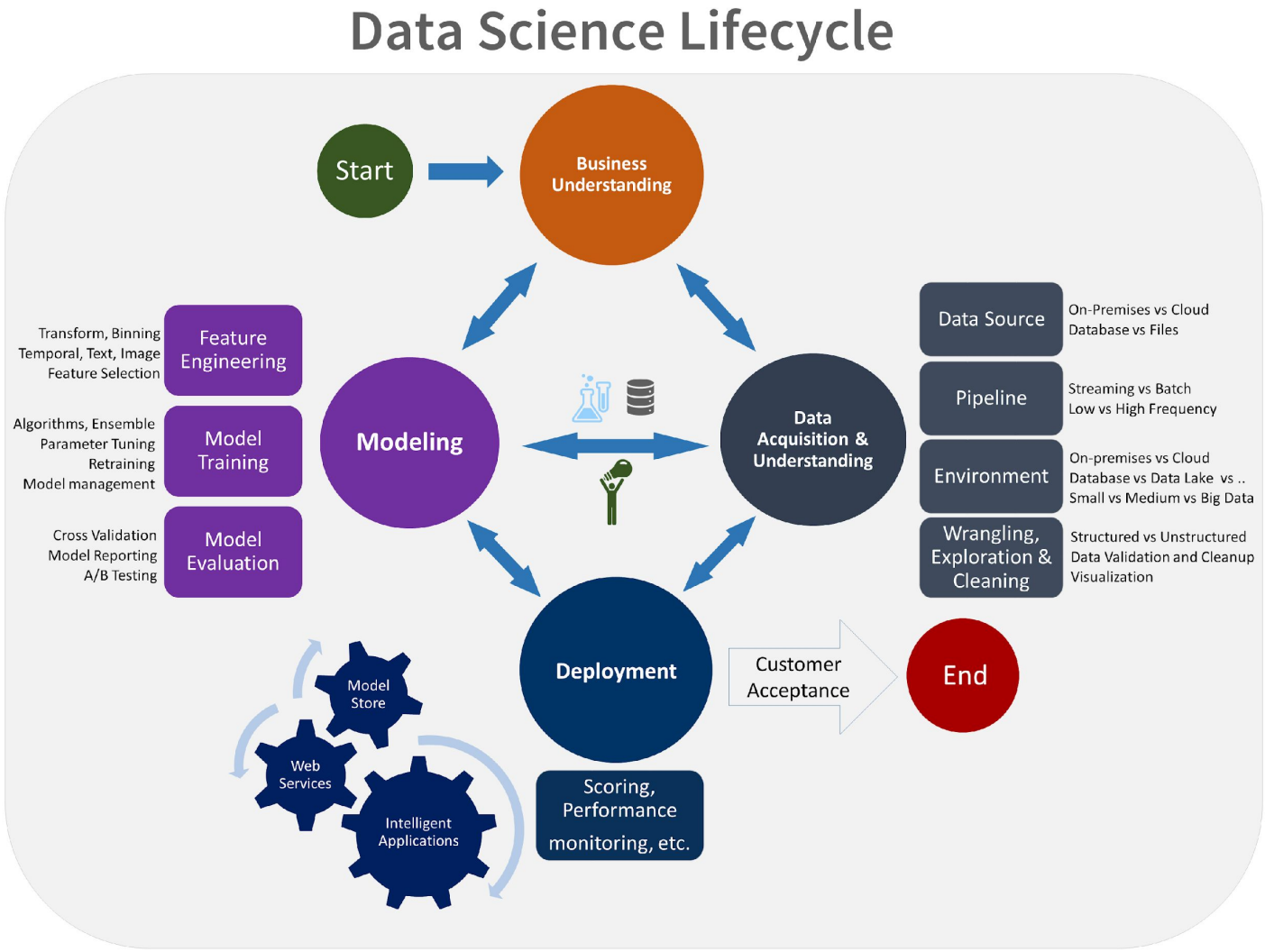


The process of units interaction in modelling is developed with account for the best international standards

CRISP-DM - Cross Industry Standard Process for Data Mining (1999), used in 40+% cases

	Main process steps	Output
Research	▪ Business requirements analysis ▪ Task definition ▪ Data analysis	▪ Proof of concept ▪ Detailed requirements for development
Development	▪ Data preparation ▪ Feature engineering ▪ Modelling ▪ Validation ▪ Interpretation of results ▪ Model implementation ▪ Model usage and monitoring	▪ Development report

One more Microsoft standard: Team Data Science Process (TDSP)

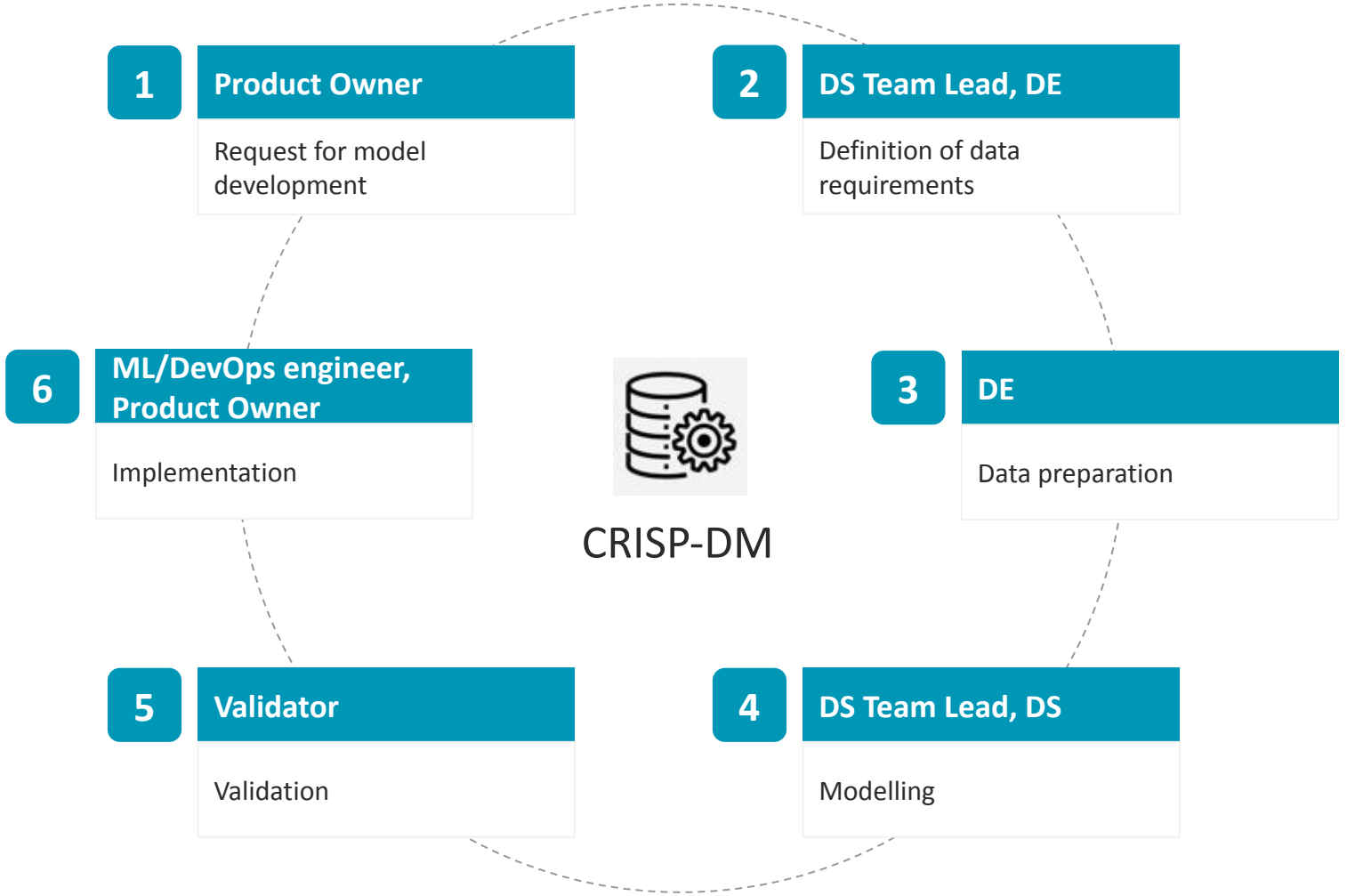


Role model according to modelling standards

CRISP-DM

MIT DS Model

Agile Methodology



10+ Roles

DS
Team Lead DS
Data Scientist
Team Lead DA
Data Analyst

DO
Data Engineer
DE Team lead

DevOps
Machine Learning Engineer
Model DevOps Engineer
IT Process Owner

Business
Product Owner / Process Leader
Business Owner of IT System
Owner of Corporate Datamart

Validation
Validator

Internal Audit
Model Auditor

Model is a product or part of a product

Project leader

Product owner / Model sponsor:
Owner of the process the model is developed for

Data stream

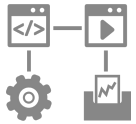
CDO

- Data Mart for development
- Data Mart for production

Model stream

CDS Risks

- Data analysis
- Algorithm selection
- Model development



IT stream

Product owner

- Model deployment in production environment

Monitoring

Product owner, CDO, Validation unit

- Monitoring system

Validation

What is important

Technological stack matters:

- Part of a product team or service (depends on if production environment is a part of the product)

Domain knowledge / DS domain

- Platform or in-house skills (capacity)

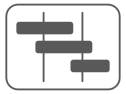
Data science needs skilled decision makers

- Watch who is your PO

Risk Model Governance (RMG) principles

DS core competence – model development / data analysis

Model owners are responsible as ultimate beneficiaries



Each model is treated as a separate product or as a part of product

Model owners are responsible for availability of core competences for a project

Matrix organizational structure for DS function: agile DS

Model owners have influence on DS motivation / evaluation

Three line of defense concept is introduced in model risk management

Model risk management implies consolidated responsibility

1st line of defense

Model sponsor, definition of business goals, preliminary list of hypotheses and model's success criteria

Developer, development of a model in accordance with the internal standards

Model owner, management of a model application process, monitoring of model's performance in processes

Model user, monitoring of errors in model application

2nd line of defense

Validation and MRM, model quality assessment, Model risk management

3rd line of defense

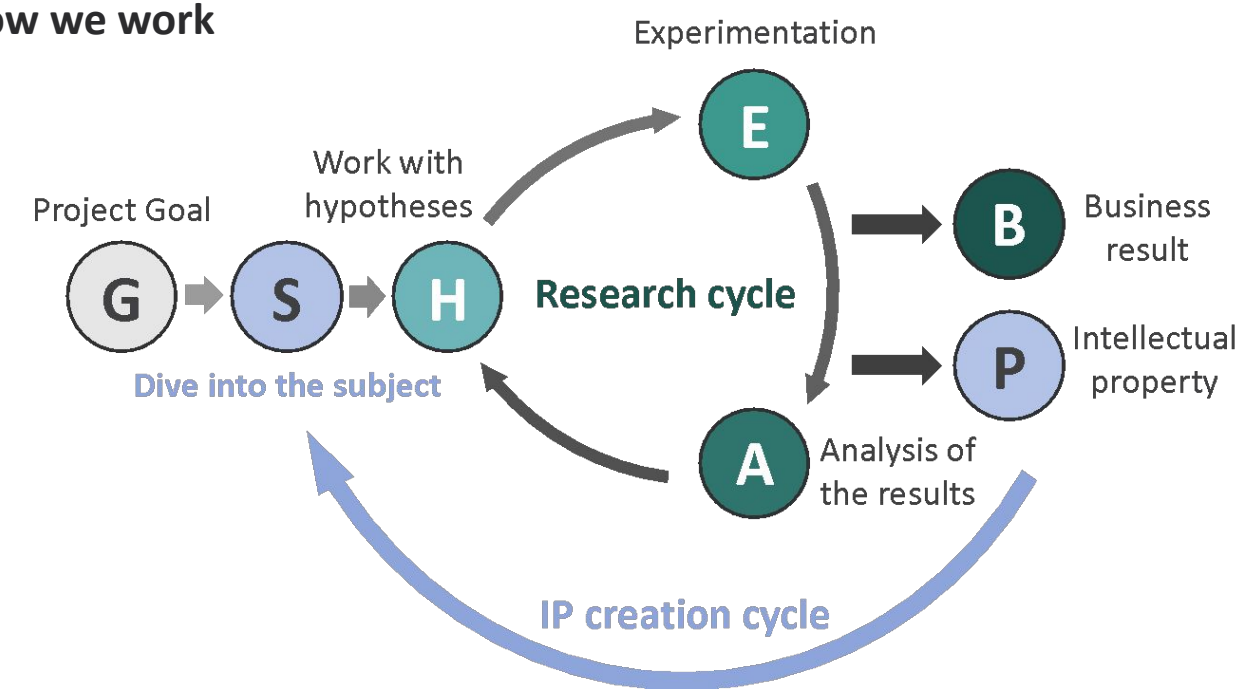
Internal audit, assessment of application and effectiveness of model risk management procedures

R&D function: opportunity for loonshots*

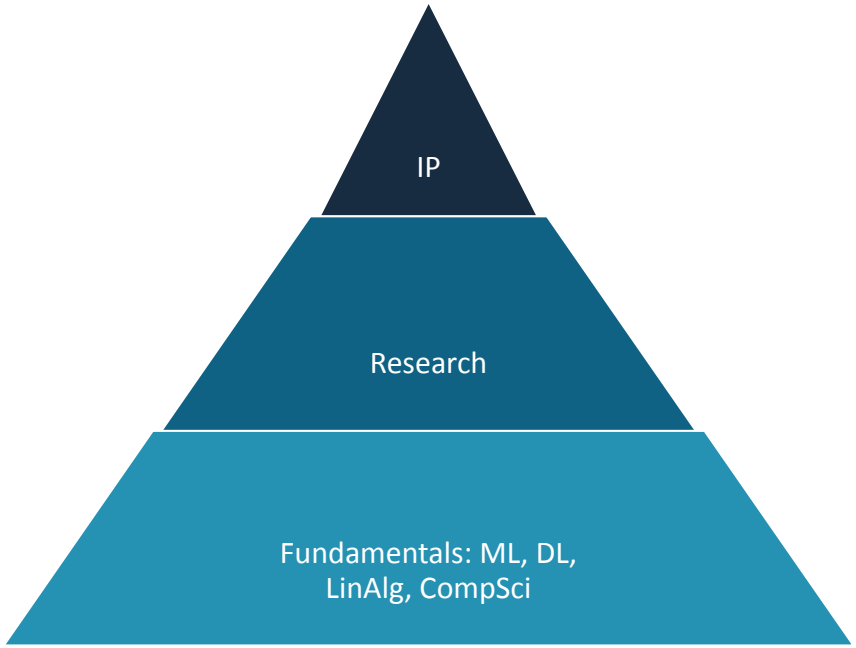
Role in the process

Search for new ideas in risk estimation to disrupt existing processes to get superprofit.

How we work



Skills



*Loonshot. Longshot — long perspective (high risk), and loon — crazy.
From book: ”Loonshots: How to Nurture the Crazy Ideas That Win Wars, Cure Diseases, and Transform Industries” by Safi Bahcall


Sberbank