Medicines Discovery Catapult (MDC)

Rafael C Jimenez

Accessing biomedical data for data analysis

B4BB, SME Forum

2022-07-06

MDC

ESTABLISHED IN 2015

BASED IN ALDERLEY PARK, CHESHIRE - THE UK'S LARGEST SINGLE SITE LIFE SCIENCES CAMPUS

PART OF THE CATAPULT NETWORK – ESTABLISHED BY INNOVATE UK

INDEPENDENT, NOT-FOR-PROFIT ORGANISATION

TEAM OF > 130 SPECIALISTS WITH EXTENSIVE INDUSTRY EXPERIENCE

To reshape drug discovery for the benefit of patients

SUPPORTING THE DRUG DISCOVERY INDUSTRY

SMEs and innovators to deliver growth for the drug discovery economy

LEVERAGING UK INFRASTRUCTURE

Connecting the community

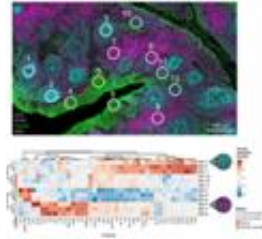**TACKLING INDUSTRY-LED CHALLENGES**

That limit today's discovery process

INDUSTRIALISING AND DRIVING ADOPTION
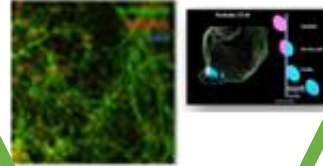
Of new scientific tools and modern techniques for discovering medicines
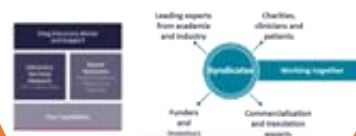
MDC capabilities

Target and pathway engagement

Drug Delivery Biodistribution

Biomarkers

Biomedical informatics

Cell model analysis and validation

Virtual R&D and Syndicates
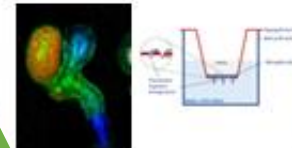
MDC Informatics

MDC Informatics

Analyse and interpret complex data to help you to make decisions about developing new medicines

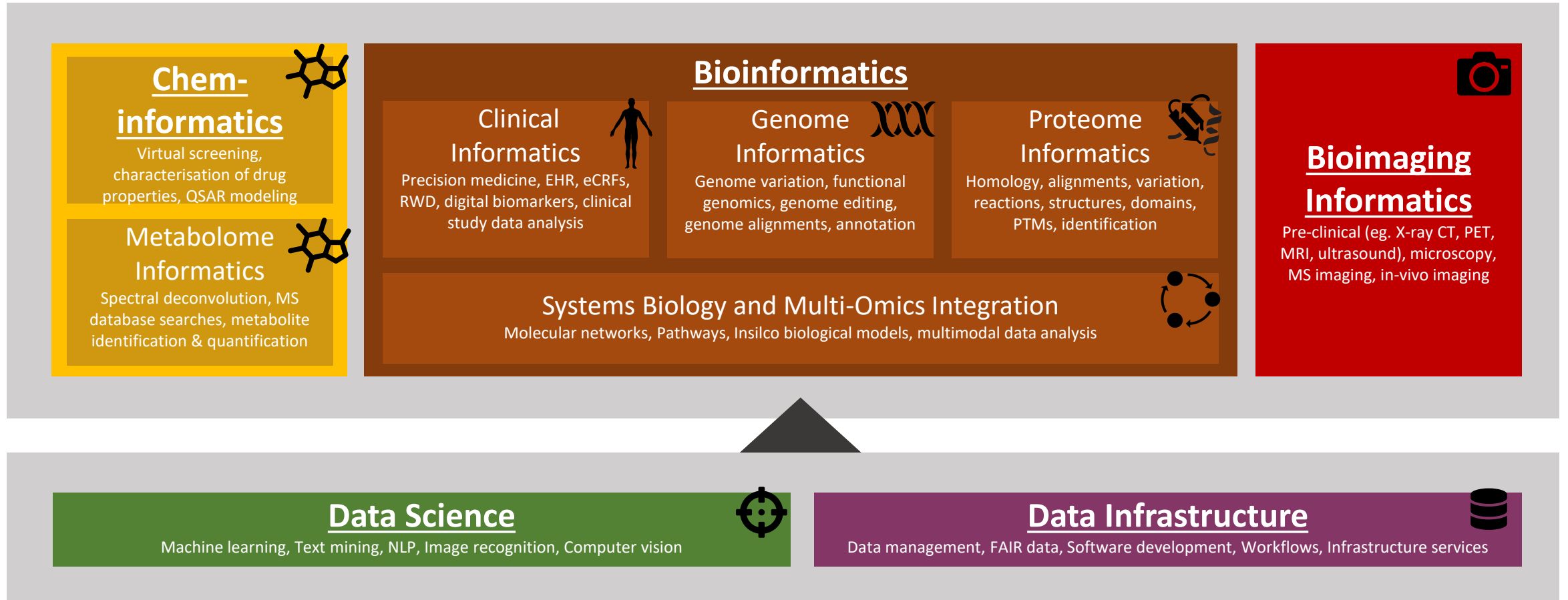| Multidisciplinary expertise | In-house developed assets | Integration of third party data | Specialised software | Data infrastructure |

Team and expertise

Chem-informatics
Virtual screening, characterisation of drug properties, QSAR modeling

Metabolome Informatics
Spectral deconvolution, MS database searches, metabolite identification & quantification

Bioinformatics

Clinical Informatics
Precision medicine, EHR, eCRFs, RWD, digital biomarkers, clinical study data analysis

Genome Informatics
Genome variation, functional genomics, genome editing, genome alignments, annotation

Proteome Informatics
Homology, alignments, variation, reactions, structures, domains, PTMs, identification

Systems Biology and Multi-Omics Integration
Molecular networks, Pathways, Insilco biological models, multimodal data analysis

Bioimaging Informatics
Pre-clinical (eg. X-ray CT, PET, MRI, ultrasound), microscopy, MS imaging, in-vivo imaging

Data Science
Machine learning, Text mining, NLP, Image recognition, Computer vision

Data Infrastructure
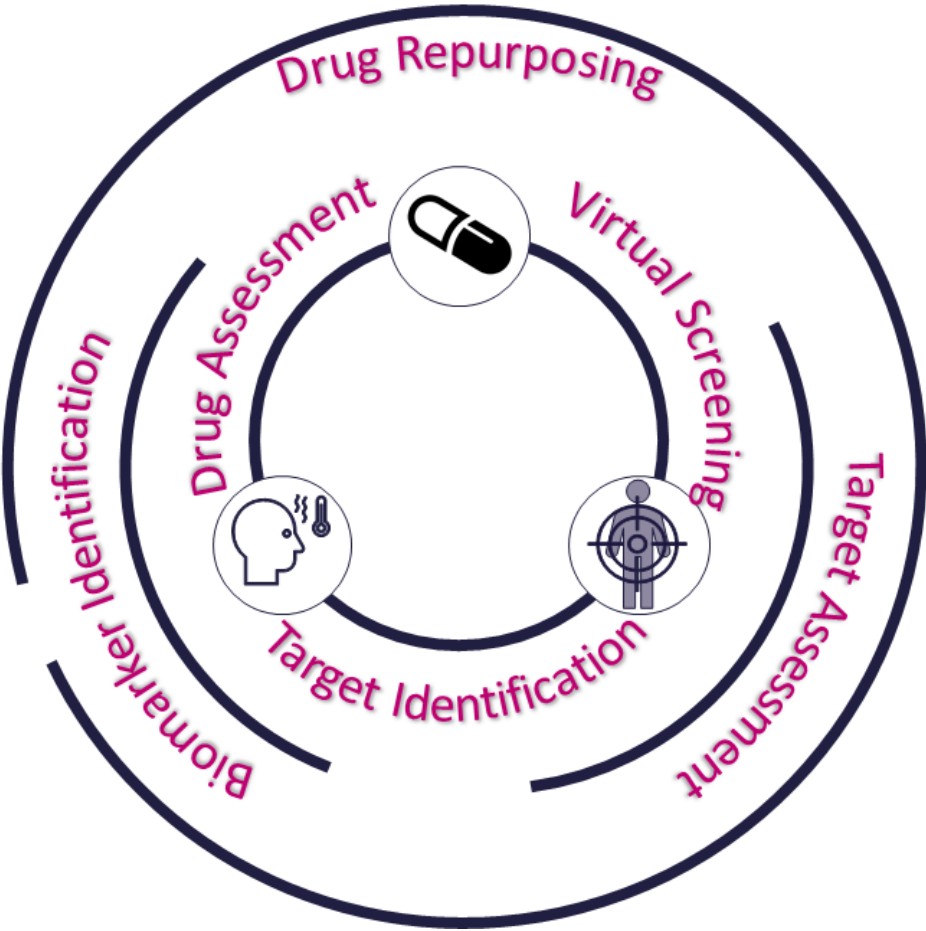Data management, FAIR data, Software development, Workflows, Infrastructure services

Informatics capabilities

capabilities

insights

CartaPharma
Resource integrating ext & int data assets

CrISPy
DB with insights about drug discovery UK companies

Aurac
Plugin enriching and visualising biochem entities

AssayNet
Data mining resource of assays, drugs & targets

Ares
Virtual screening pipeline suing structural methods

DeepADMET
Data mining resource of ADMET properties

DREaM.ai
DB ranking drugs based on its properties

Apollo
DB of antimicrobials & their gold-standard annotations

MDC-PPDMS
Database of druggable domains

DSP ATLAS
Tool to manage & analyse spatial expression data

DARTS
DB of real drugs and biological targets

Asclepius
DB of Clinical Trials & Drug-Target-Disease annotations

VESPA
Tool to predict the impact of mutations in proteins

Narcissus
Tool to handle metadata & support analysis workflows

Polyphemus
DB with compounds from vendors and external DBs

Pan
Enriched library of chemical structures using FBLD

Dionysus
Dataset of drugged and druggable genes

DocLib
Software components for text mining

NER
Tool for named entity recognition for biochem

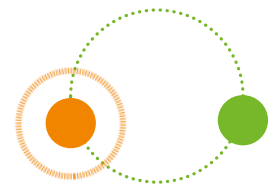**AMBER+**
Toolkit to use and integrate AMBER

Drug Repurposing

Virtual Screening

Drug Assessment

Biomarker Identification

Target Identification

Target Assessment

Disease mechanisms

Biomarkers

Targets

Druggability

Drugs

Efficacy

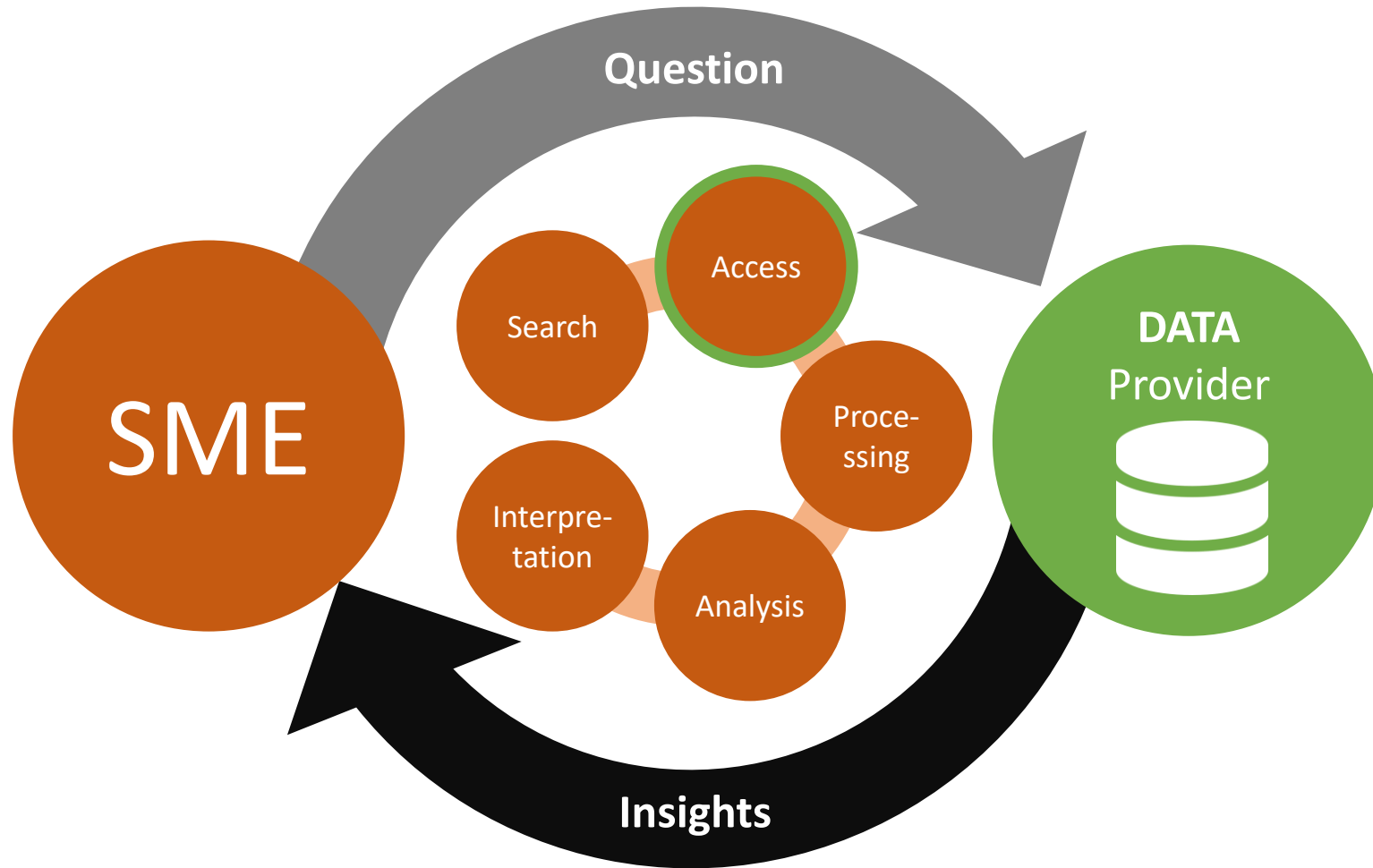Central Nervous System | Infectious Diseases | Oncology | Complex medicines

Workshop expectations

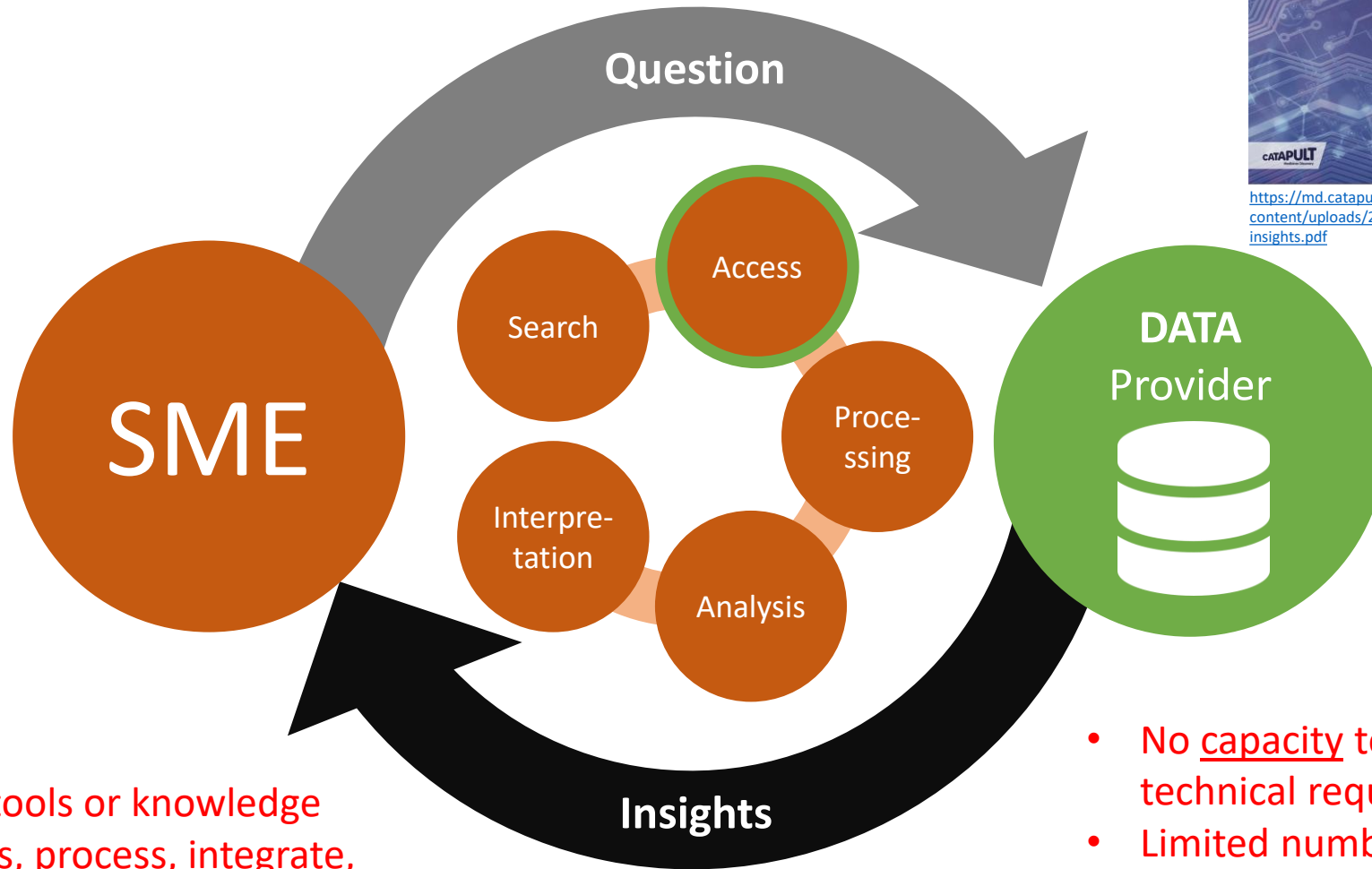**Learn** more about **health data resources** and **solutions to access, integrate, and analyse** health data

Explore the **opportunity** for public funded service providers to **work as a collective** to support SMEs

Data to insights

Data to insights

Question

Insights

SME

Search | Access | Processing | Analysis | Interpretation

DATA Provider

https://md.catapult.org.uk/wp-content/uploads/2019/05/sodn19-ai-insights.pdf

https://md.catapult.org.uk/wp-content/uploads/2019/10/health-data-report.pdf

issues

- No capacity, tools or knowledge to find, access, process, integrate, analyse and interpret 3rd party data

- No capacity to help SMEs with their technical requirements
- Limited number of tools and refence data to support data analysis
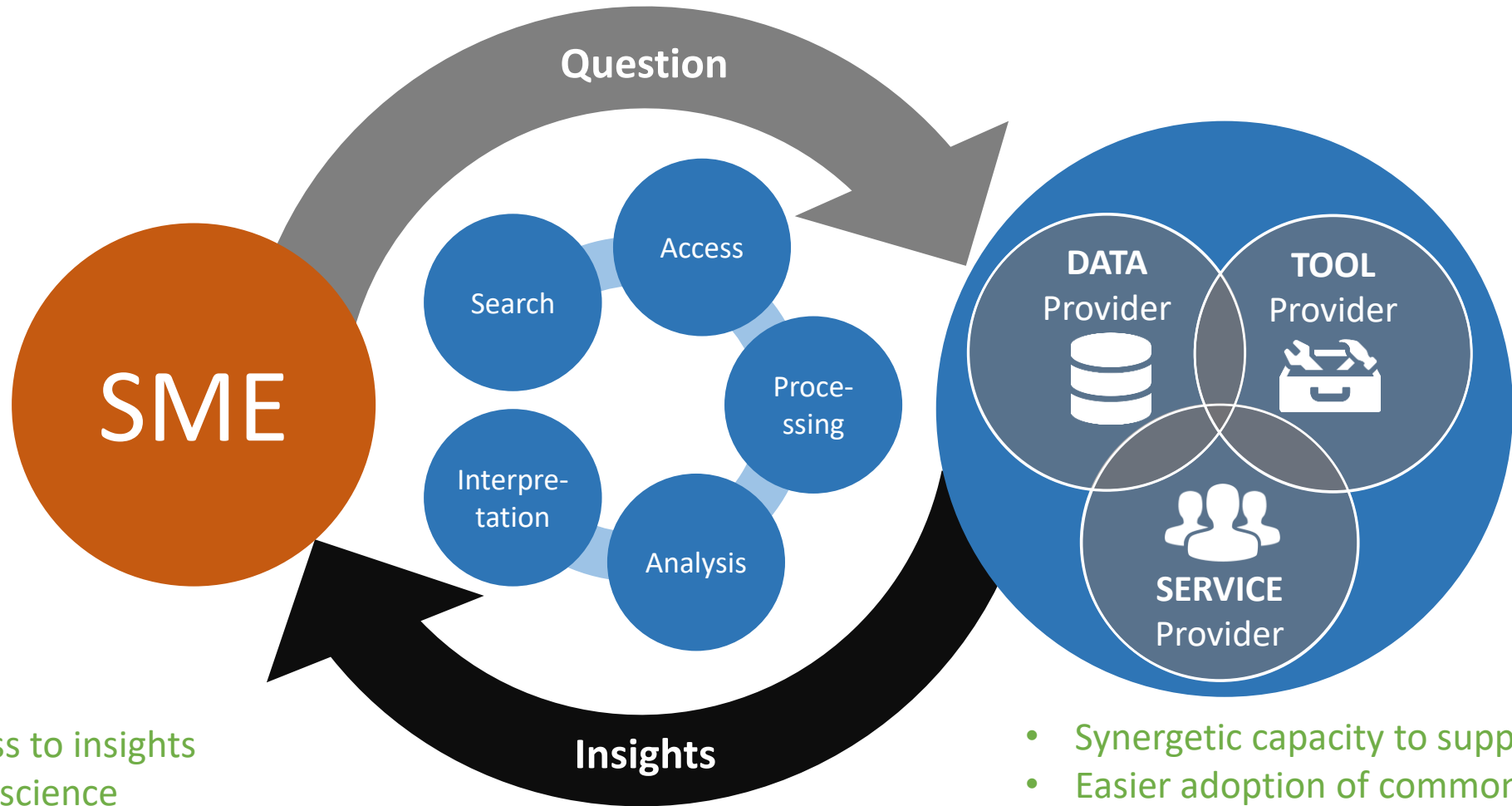- Limited data analysis support or domain knowledge to help with data analysis
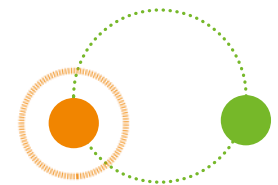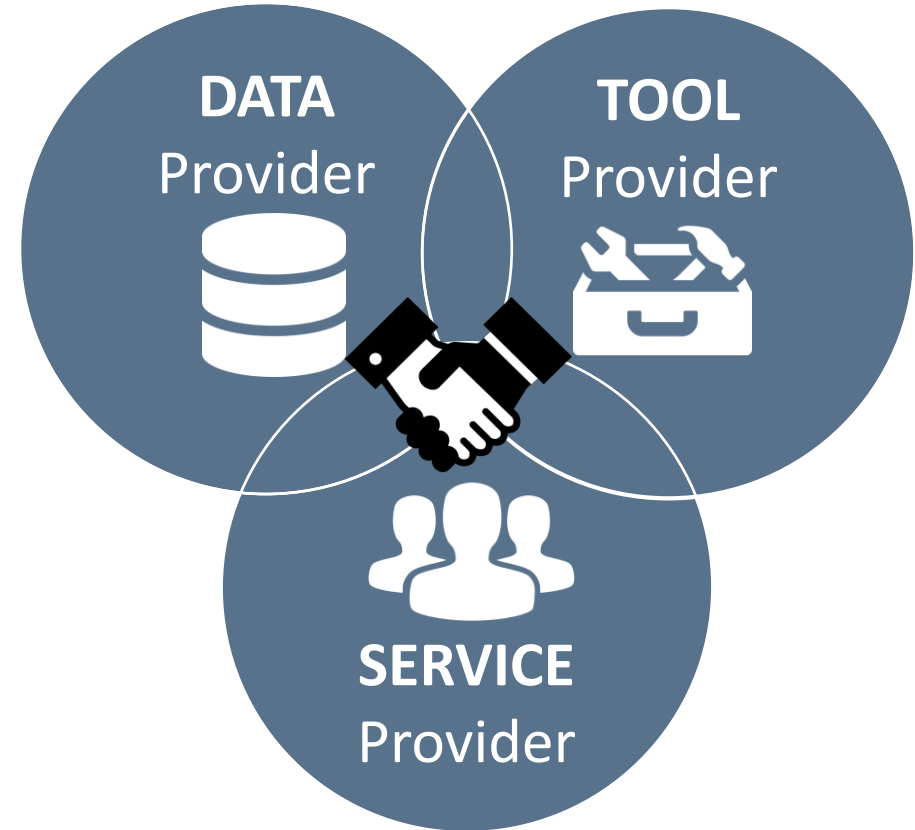
Work as a collective

Question

Search · Access · Proce-ssing · Analysis · Interpre-tation

SME

DATA Provider · **TOOL** Provider · **SERVICE** Provider

Insights

benefits

- Quicker access to insights
- Focus on the science
- No need to worry about legal or technical related issues

- Synergetic capacity to support SMEs
- Easier adoption of common solutions
- Better integration of public funded initiatives
- Not sharing data but insights

Questions

- What are the **needs** (and bottlenecks) of SMEs to **exploit** cohort data?

- What practical **solutions** could be easily implemented **across organisations** to support SMEs?

- What **agreements** **between organisations** could help to better support SMEs?

Thanks for your attention

Follow up proposal

Work as a collective to support SMEs to get insights from health data

B4BB workshop	→	SME consultation	→	White paper	→	Stakeholder survey	→	Follow up workshop
Exploring interest & solutions		Learning from SMEs		Proposing solutions		Validating solutions		Working on implementation