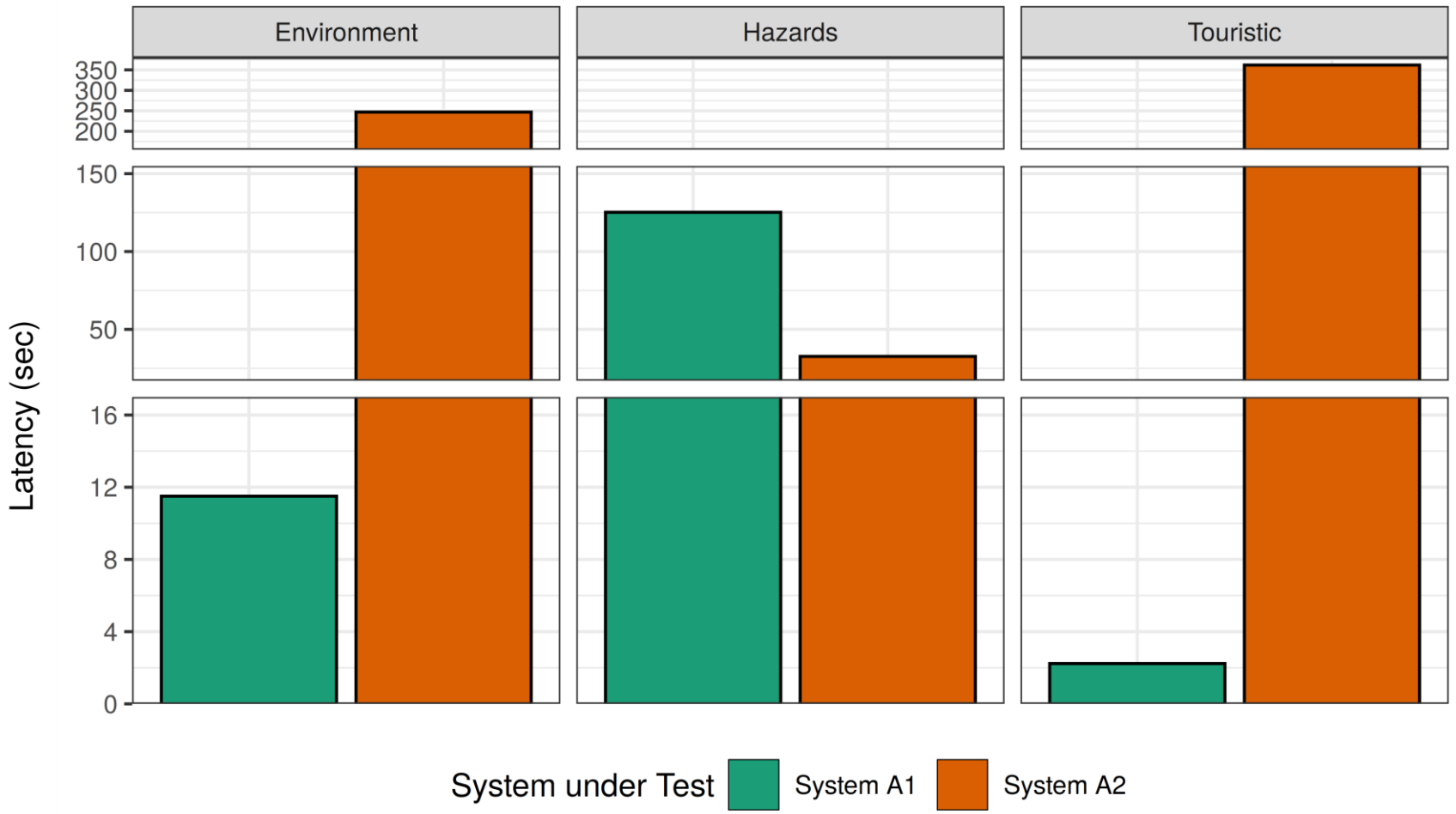
Automating Earth Observation Analytics Pipelines with Agent Raven

G. Dusella, H. Gavriilidis, B. Chen, B. Demir, V. Markl, E. Tzirita Zacharatou

Big Data from Space 2025

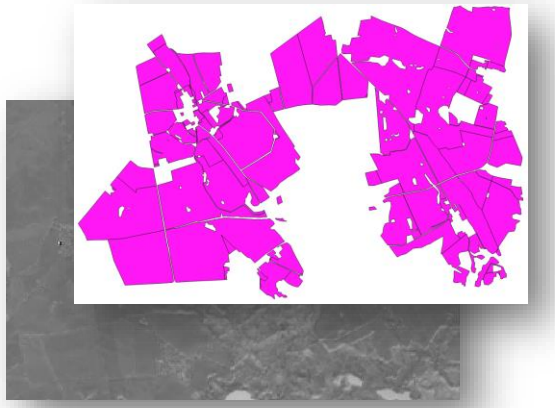
Variance in System Usage and Configurations



Optimizing Queries

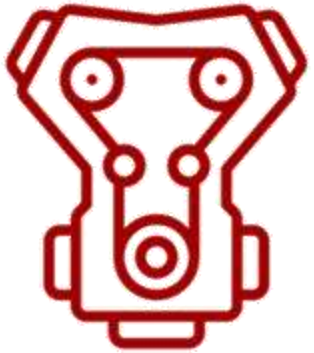
```
zs_result = ZSGen.build(
    raster="/data/s2a_ndvi_20m",
    vector="/data/ALKIS_nutz_MOL")
.group("oid")
.summarize({"max": ZSAgg.MAX,
            "avg": ZSAgg.AVG})
.filter("nutzart = 'Landwirtschaft'")
.system(System.PostGIS())
.vectorize_type(VecType.POINTS)
.align_to_crs(DataType.RASTER)
.vector_filter_at(Stage.EXECUTION)
.raster_clip(True)
```

Workload / Query



Datasets

Optimizer



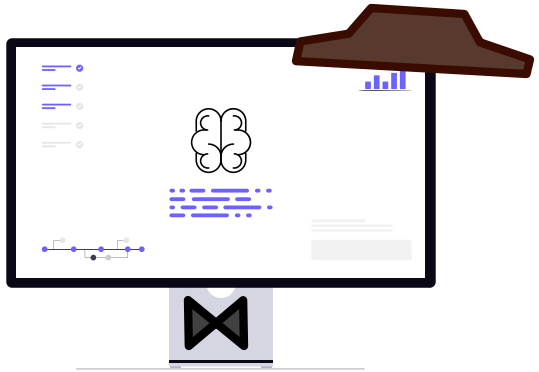
Execution Engine

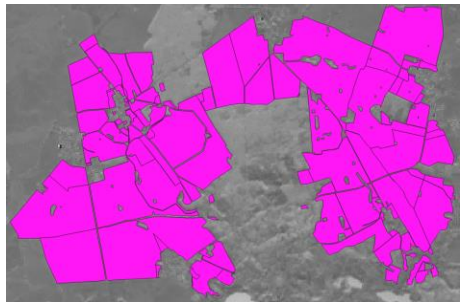
oid	nutzart	bez	is_forest	count(*)
DEBBAL010007ZkSBTN	Landwirtschaft	Ackerland	0	12312
DEBBAL010007ZkSBTN	Landwirtschaft	Ackerland	1	2
DEBBAL01000ceYk7TN	Landwirtschaft	Ackerland	0	4534
DEBBAL010007ZkSDTN	Landwirtschaft	Ackerland	0	3448
DEBBAL010007ZkSDTN	Landwirtschaft	Ackerland	1	32
DEBBAL01000caTdMTN	Landwirtschaft	Brachland	0	45
DEBBAL01000caTdMTN	Landwirtschaft	Brachland	1	1
DEBBAL01000caTe2TN	Landwirtschaft	Brachland	0	34384
DEBBAL01000caTe4TN	Landwirtschaft	Ackerland	1	9287
DEBBAL01000cc3ATTN	Landwirtschaft	Ackerland	0	4354
DEBBAL01000cc3ATTN	Landwirtschaft	Ackerland	1	47

Result

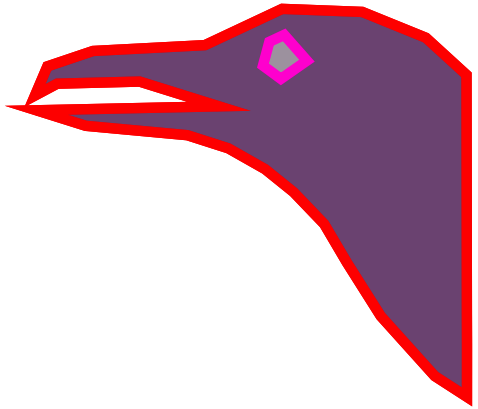
Specialized AI Agents for EO

AI Agent
Optimizer





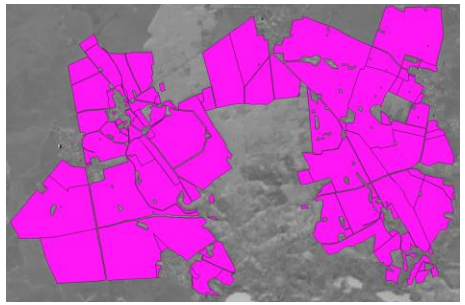
Zonal Statistics



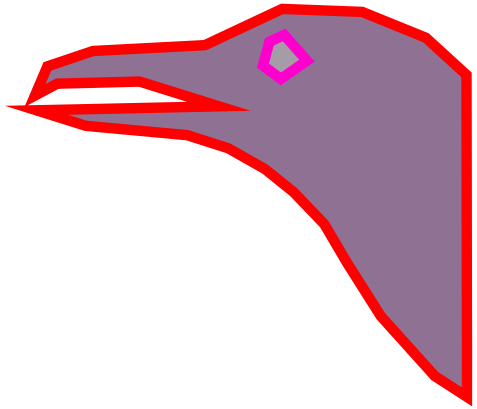
Zonal Statistics with Raven

Specialized AI Agents for EO

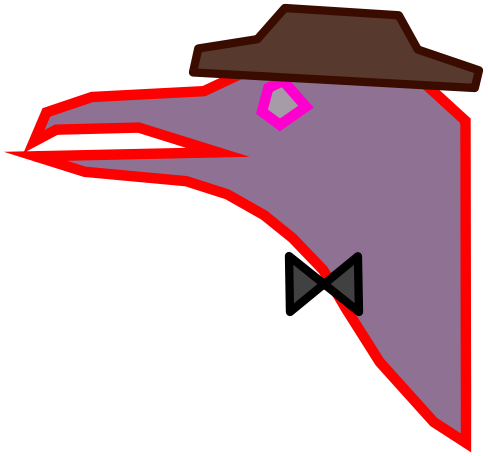
AI Agent
Optimizer



Zonal Statistics

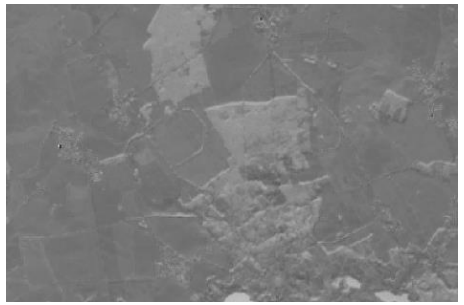


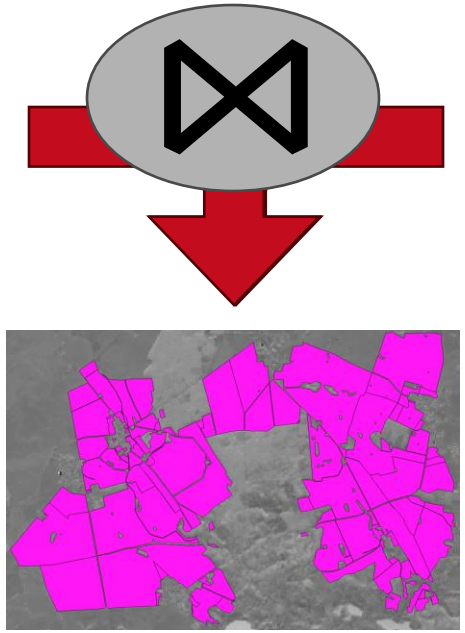
Zonal Statistics with Raven



Agent Raven

Zonal Statistics Queries





oid	nutzart	bez	is_forest	count(*)
DEBBAL010007ZkSBTN	Landwirtschaft	Ackerland	0	12312
DEBBAL010007ZkSBTN	Landwirtschaft	Ackerland	1	2
DEBBAL01000ceYk7TN	Landwirtschaft	Ackerland	0	4534
DEBBAL010007ZkSDTN	Landwirtschaft	Ackerland	0	3448
DEBBAL010007ZkSDTN	Landwirtschaft	Ackerland	1	32
DEBBAL01000caTdMTN	Landwirtschaft	Brachland	0	45
DEBBAL01000caTdMTN	Landwirtschaft	Brachland	1	1
DEBBAL01000caTe2TN	Landwirtschaft	Brachland	0	34384
DEBBAL01000caTe4TN	Landwirtschaft	Ackerland	1	9287
DEBBAL01000cc3ATTN	Landwirtschaft	Ackerland	0	4354
DEBBAL01000cc3ATTN	Landwirtschaft	Ackerland	1	47

ZS Query Parameters

Load Datasets

Select Vector Features

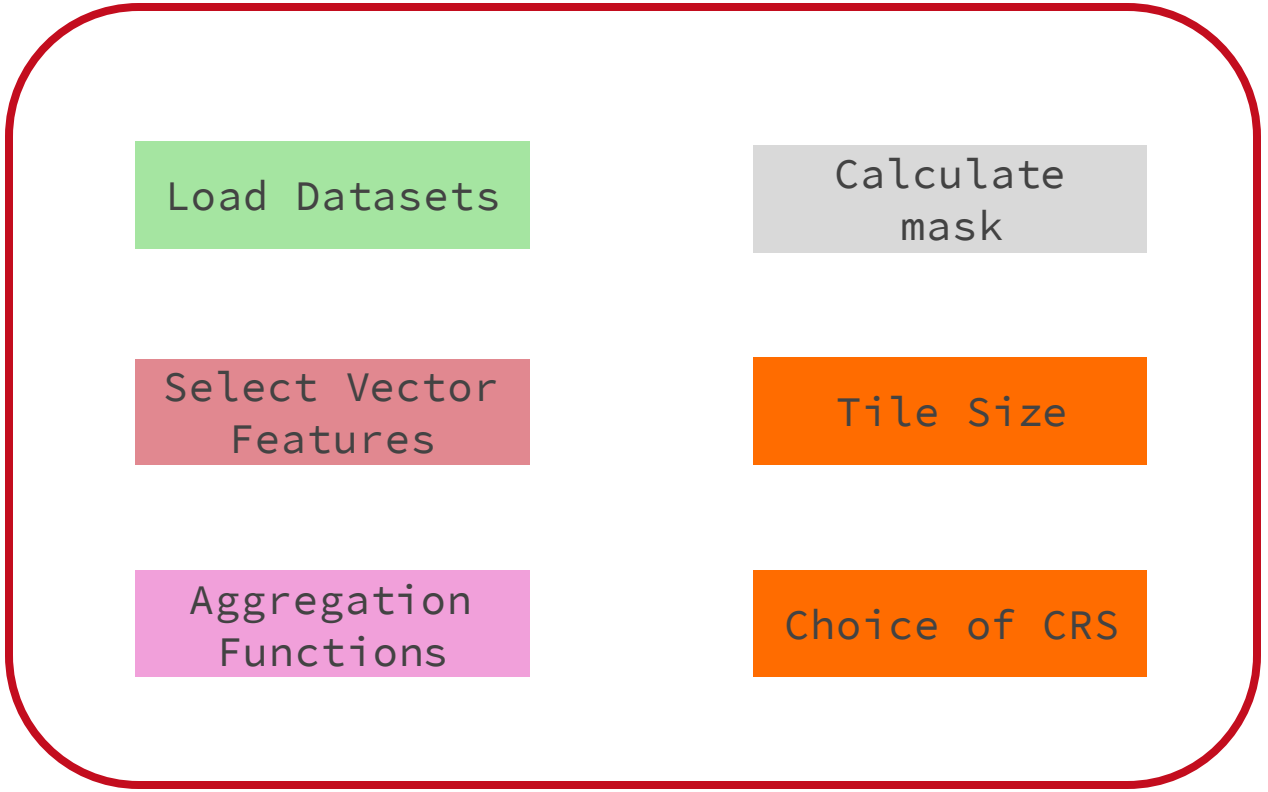
Aggregation Functions

Calculate mask

Tile Size

Choice of CRS

ZS Query Meta Parameters



Which system should I choose?

Linguistic Clutter

Load Datasets

Select Vector Features

Aggregation Functions

Calculate mask

PostGIS

```
SELECT v.id,
   SUM(t.count) AS count,
   SUM(t.count * t.value) / SUM(t.count) AS mean,
FROM plots v JOIN sentinel r
   ON ST_Intersects(raster.rast, vector.geom),
   ST_ValueCount(
   ST_Clip(raster.rast, vector.geom), 1) AS t
GROUP BY v.id
```

Apache Sedona

```
SELECT v.id,
   RS_ZonalStats(r.rast, v.geom, 'avg') as mean,
   RS_ZonalStats(r.rast, v.geom, 'count') as count,
FROM plots v JOIN sentinel r
   ON RS_Intersects(v.geom, r.rast)
```

```
def main(args): Unit = {
  // Setup Spark + Beast
  // load raster and vector datasets
  val join = raster.raptorJoin(vector)

  val rdd_raw = join.map(v => (v.feature
      .getAs[String]("id"), v.m))
  spark.createDataset(rdd_raw)
    .toDF(Seq("id", "value"): _*)
    .createOrReplaceTempView("raptorjr")

  val zonal_stats = spark.sql("""
  SELECT rj.id, COUNT(rj.value) AS count,
    AVG(rj.value) AS avg
  FROM raptorjr rj GROUP BY rj.id
  """)
  // Return result
}
```

BEAST

rasdaman raster data manager

Execute |*features*| times

```
for p in (sentinel)
  return avg( clip( c,
    Polygon((14.0061 52.7301, 14.0404 52.7300,
             14.0341 52.7420, 14.0464 52.7522,
             14.0131 52.7526, 14.0061 52.7301)),
     "http://loc:80/rasdaman/def/crs/EPSG/0/3587"
  ))
```


HPI BIFOLD Universität Potsdam berlin

Raven: Unifying Zonal Statistics

System-Agnostic Queries

Flexible Configuration

Easy Optimizations

Compare Systems

Raven Query

```
zs_result = ZSGen.build(
        raster="/data/s2a_ndvi_20m",
        vector="/data/ALKIS_nutz_MOL")
  .group("oid")
  .summarize({"max": ZSAgg.MAX,
              "avg": ZSAgg.AVG})
  .filter('nutzart="Landwirtschaft"')
  .system(System.PostGIS)
  .vectorize_type(VecType.POINTS)
  .align_to_crs(DataType.RASTER)
  .vector_filter_at(Stage.EXECUTION)
  .raster_clip(True)
```

Dataset

Aggregation

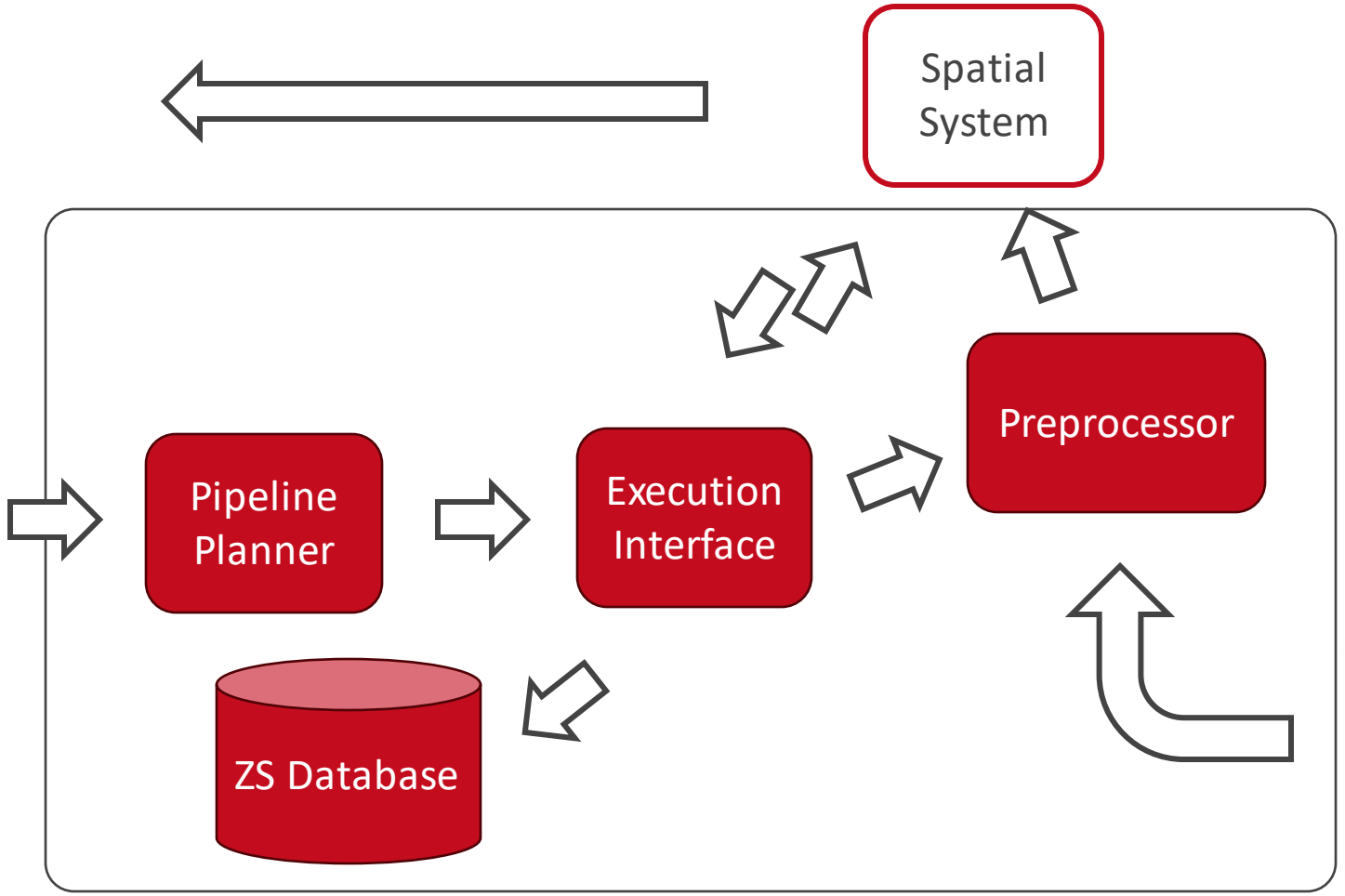
System

Parameters

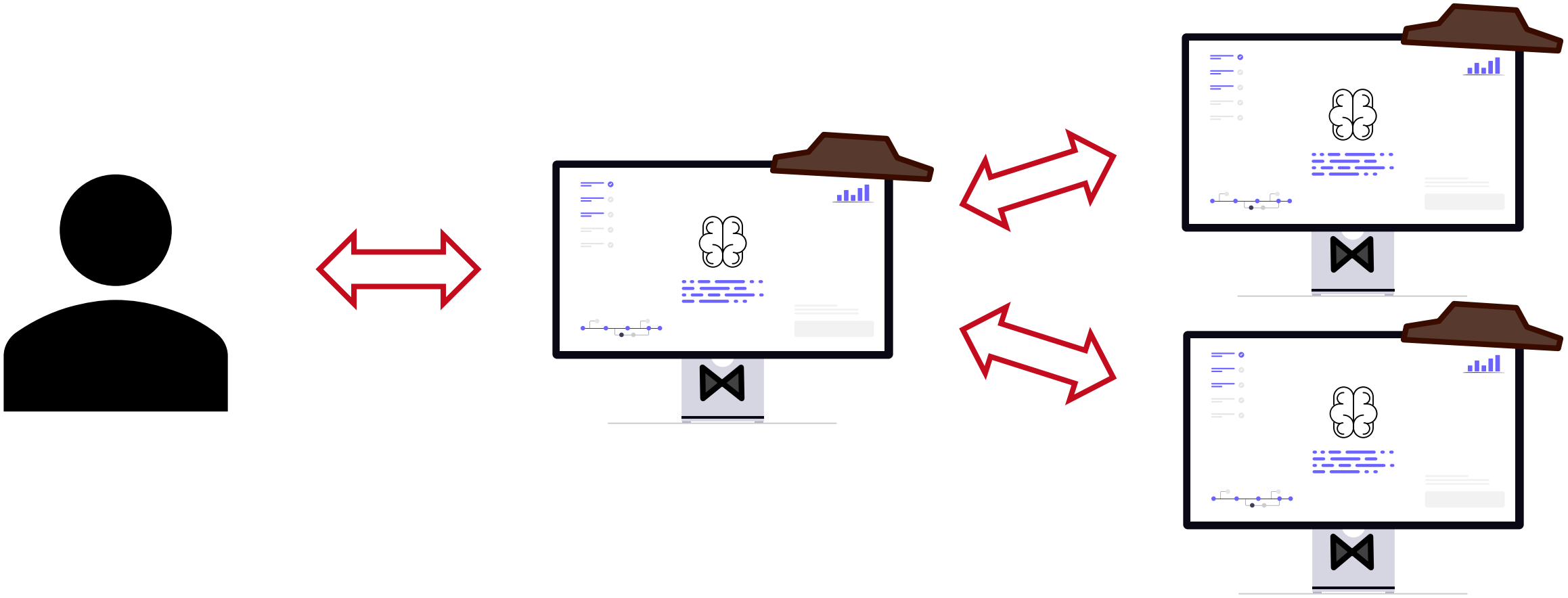
Raven Query Execution

oid	nutzart	bez	is_forest	count(*)
DEBBAL010007ZkSBTN	Landwirtschaft	Ackerland	0	12312
DEBBAL010007ZkSBTN	Landwirtschaft	Ackerland	1	2
DEBBAL01000ceYk7TN	Landwirtschaft	Ackerland	0	4534
DEBBAL010007ZkSDTN	Landwirtschaft	Ackerland	0	3448
DEBBAL010007ZkSDTN	Landwirtschaft	Ackerland	1	32
DEBBAL01000caTdMTN	Landwirtschaft	Brachland	0	45
DEBBAL01000caTdMTN	Landwirtschaft	Brachland	1	1
DEBBAL01000caTe2TN	Landwirtschaft	Brachland	0	34384
DEBBAL01000caTe4TN	Landwirtschaft	Ackerland	1	9287
DEBBAL01000cc3ATTN	Landwirtschaft	Ackerland	0	4354
DEBBAL01000cc3ATTN	Landwirtschaft	Ackerland	1	47

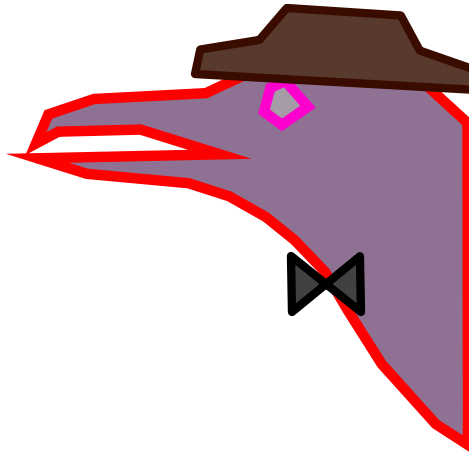
```
zs_result = ZSGen.build(
        raster="/data/s2a_ndvi_20m",
        vector="/data/ALKIS_nutz_MOL")
  .group("oid")
  .summarize({"max": ZSAgg.MAX,
              "avg": ZSAgg.AVG})
  .filter('nutzart="Landwirtschaft"')
  .system(System.PostGIS())
  .vectorize_type(VecType.POINTS)
  .align_to_crs(DataType.RASTER)
  .vector_filter_at(Stage.EXECUTION)
  .raster_clip(True)
```

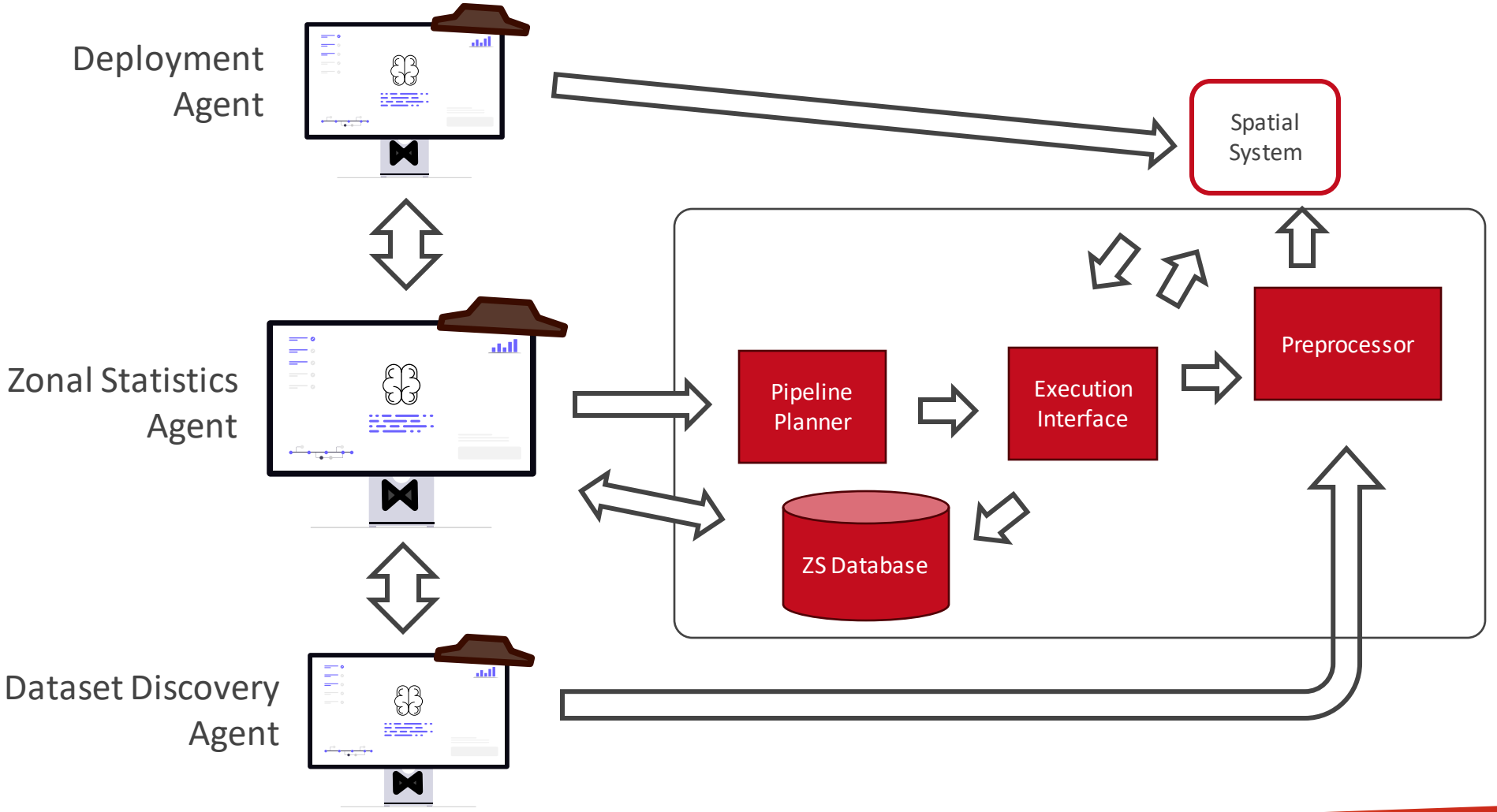


Why use an AI Agent?

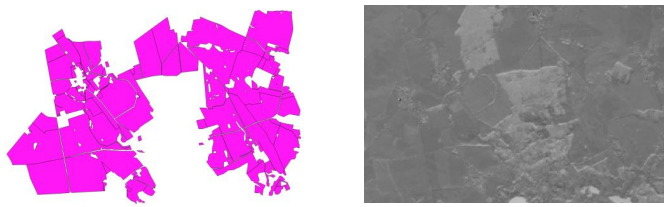


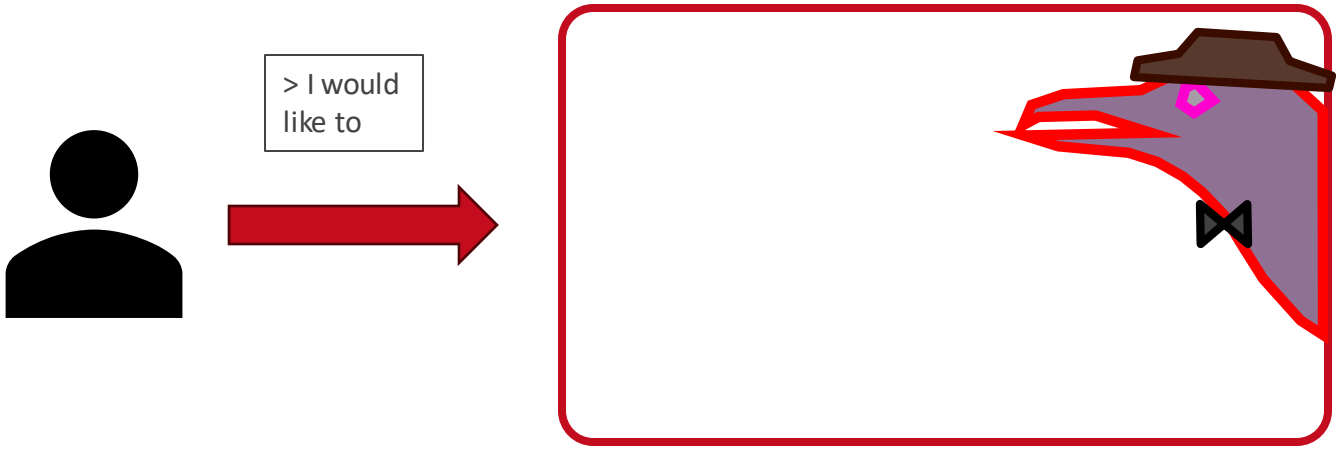
Our Vision: Agent Raven



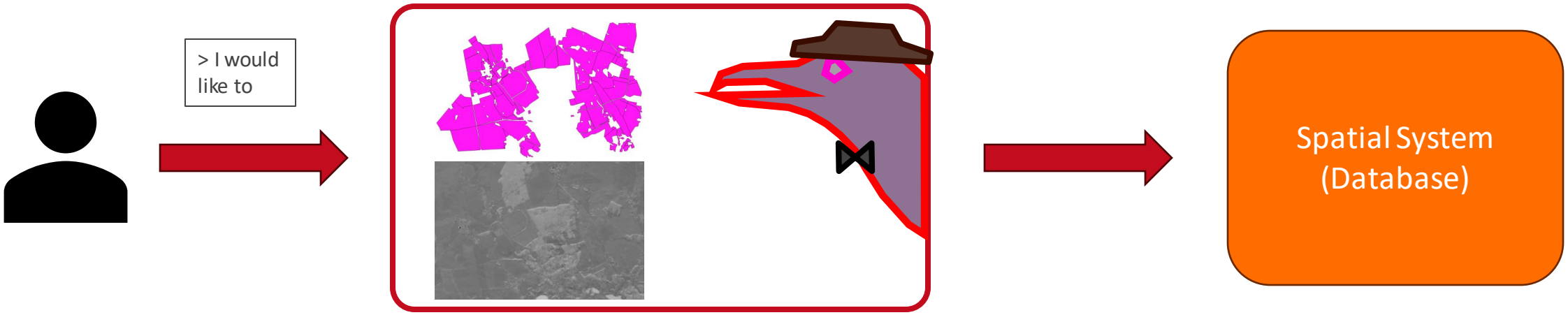


Multi-Query Optimization with Agent Raven

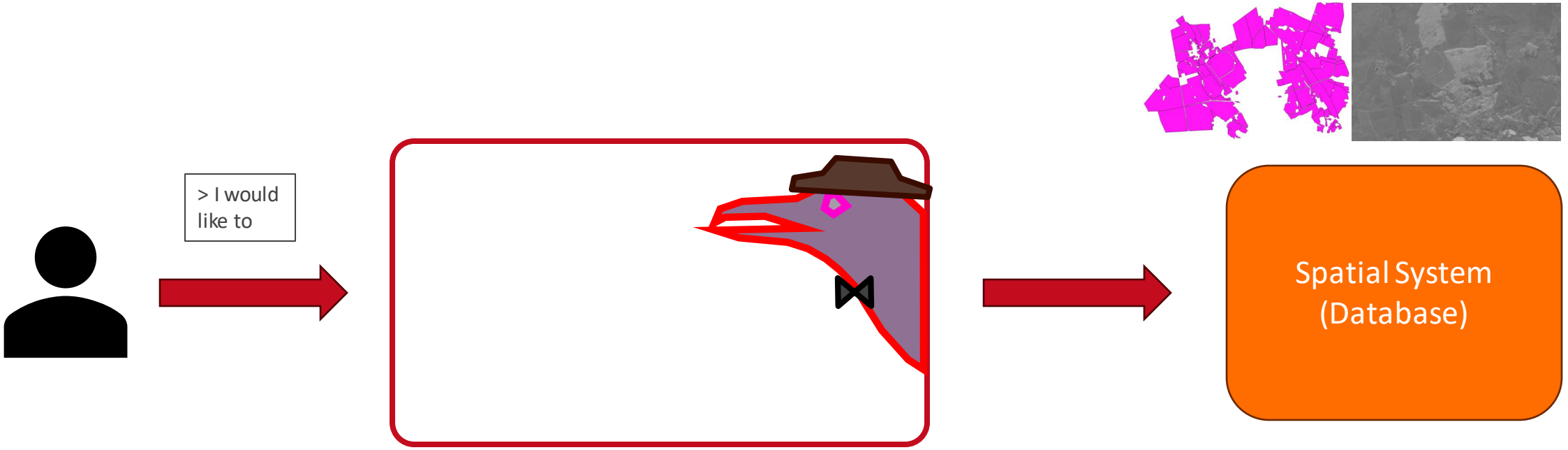




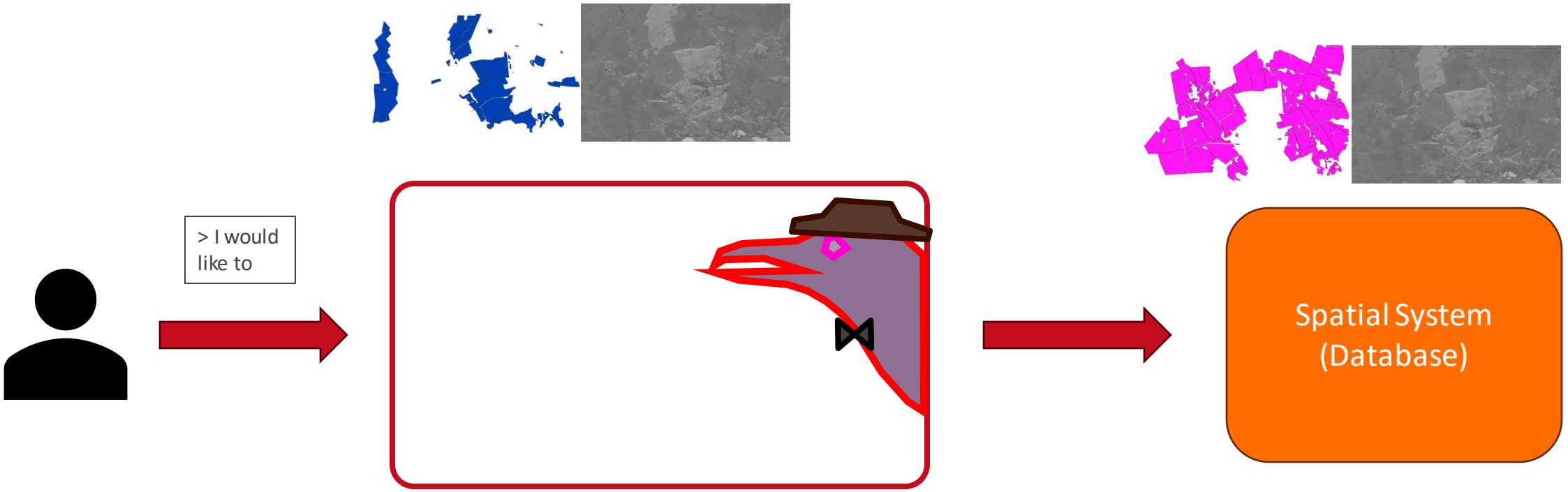
Multi-Query Optimization with Agent Raven



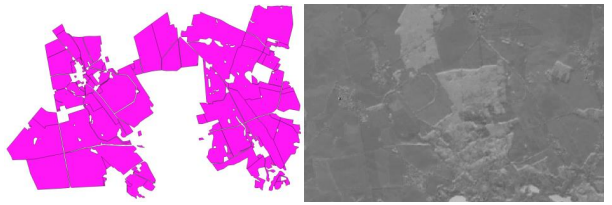
Multi-Query Optimization with Agent Raven

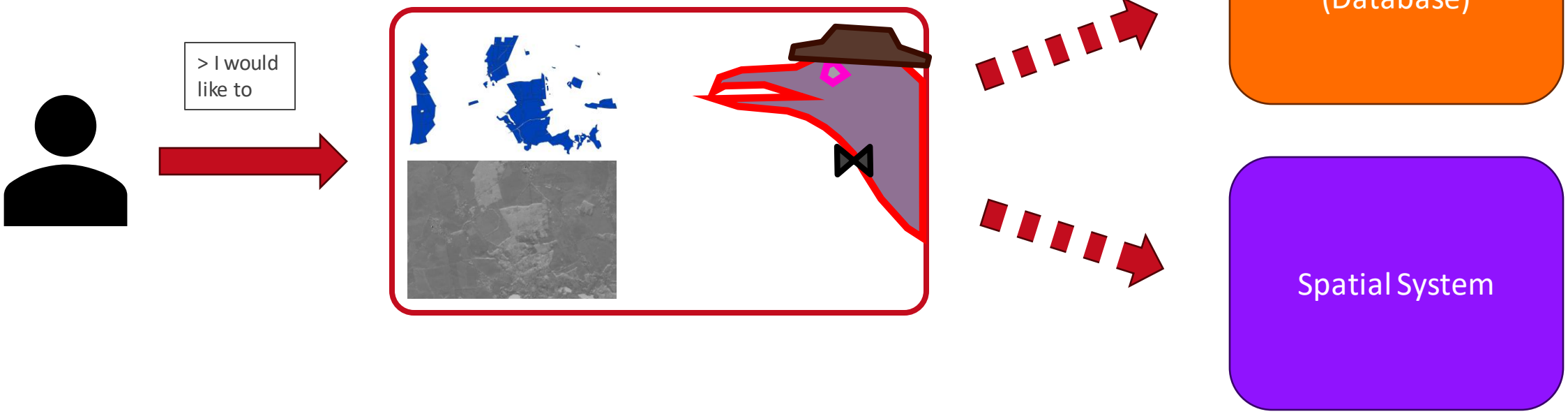


Multi-Query Optimization with Agent Raven

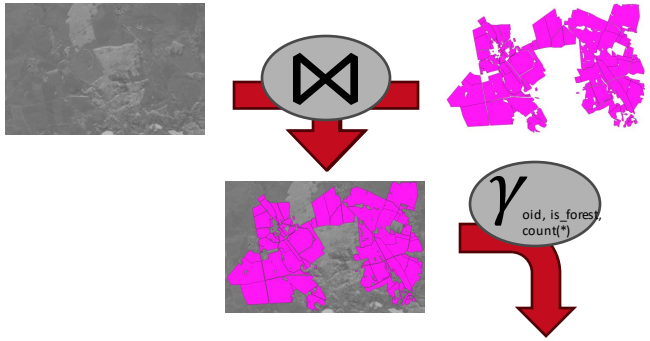


Multi-Query Optimization with Agent Raven



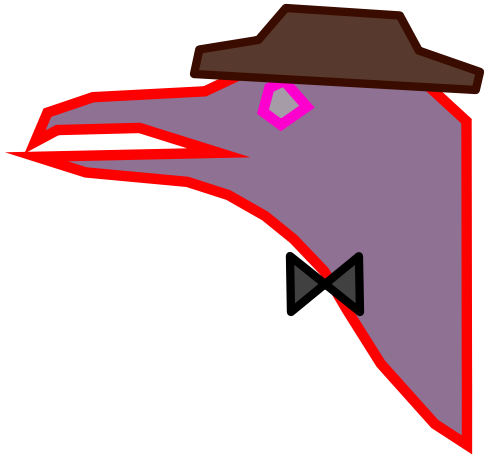


Our Vision for Agent Raven



oid	nutzart	bez	is_forest	count(*)
DEBBAL010007ZkSBTN	Landwirtschaft	Ackerland	0	12312
DEBBAL010007ZkSBTN	Landwirtschaft	Ackerland	1	2
DEBBAL01000ceYk7TN	Landwirtschaft	Ackerland	0	4534
DEBBAL010007ZkSDTN	Landwirtschaft	Ackerland	0	3448
DEBBAL010007ZkSDTN	Landwirtschaft	Ackerland	1	32
DEBBAL01000caTdMTN	Landwirtschaft	Brachland	0	45
DEBBAL01000caTdMTN	Landwirtschaft	Brachland	1	1
DEBBAL01000caTe2TN	Landwirtschaft	Brachland	0	34384
DEBBAL01000caTe4TN	Landwirtschaft	Ackerland	1	9287
DEBBAL01000cc3ATTN	Landwirtschaft	Ackerland	0	4354
DEBBAL01000cc3ATTN	Landwirtschaft	Ackerland	1	47

Zonal Statistics Queries are complex



Agent Raven will Optimize them