1) Semantically-rich Tabular Data

Gender	Age	Heart rate (bpm)	Drug	Label
1	0.10	0.239	0	0
0	0.81	0.95	2	1
1	0.20	0.851	1	0

<Existing Tabular AD Benchmark>

✅ Anomaly definition

Gender	Age	Heart rate (bpm)	Drug	Label
Female	25	81	None	0
Male	76	174	Warfarin	1
Female	32	161	Aspirin	0

<ReTabAD Benchmark>

2) Structured Text Metadata

Original Source

Domain knowledge

Manual verification

✅ Dataset name
✅ Dataset description
✅ [Column names]
✅ [Column descriptions]
✅ Label description

- Source: {Dataset original source}
- Dataset : {Dataset name}
  - {Dataset description}
- Columns
  - {Column name}: {Column description}
  - {Column name}: {Column description}
  - {Column name}: {Column description}
  - {Column name}: {Column description}
- Label
  - {Label description}

ReTabAD

CSV

JSON