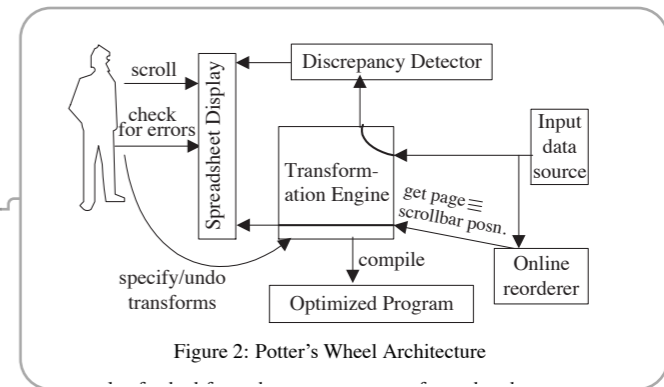


# Potter's Wheel

matchWithConfidence Error Detection

Architecture



Interactive Auditing/Transformation

- tight loop of audit/xform
- direct manipulation
- code output not data
- undo lineage

Visual algebra of transforms

Based in SchemaSQL

- Format
- Add column
- Drop column
- Copy Column
- Merge columns
- Split column (and auto-split from example)
- Divide column
- Fold
- Select
- Unfold (not visual)

MDL type induction

Intuition

- Precision: few false positives
- Recall: few false negatives
- Conciseness

MDL

- Bits to encode format
- Cardinality
- Bits to encode matches in format
- Bits to encode non-matches explicitly

Search

Prune via isRedundantAfter

Split by example

- MDL on the example subsplits
- LeftRight
- DecSpecificity
- How to split subsequent strings?