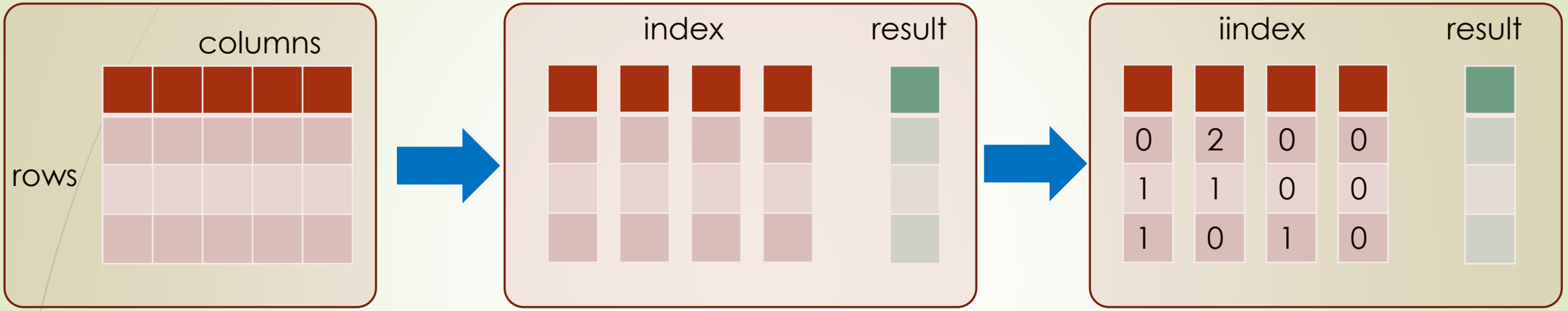
# How to convert any csv to an Xarray with chosen dimension ?

# 1 - CSV data -> llist data



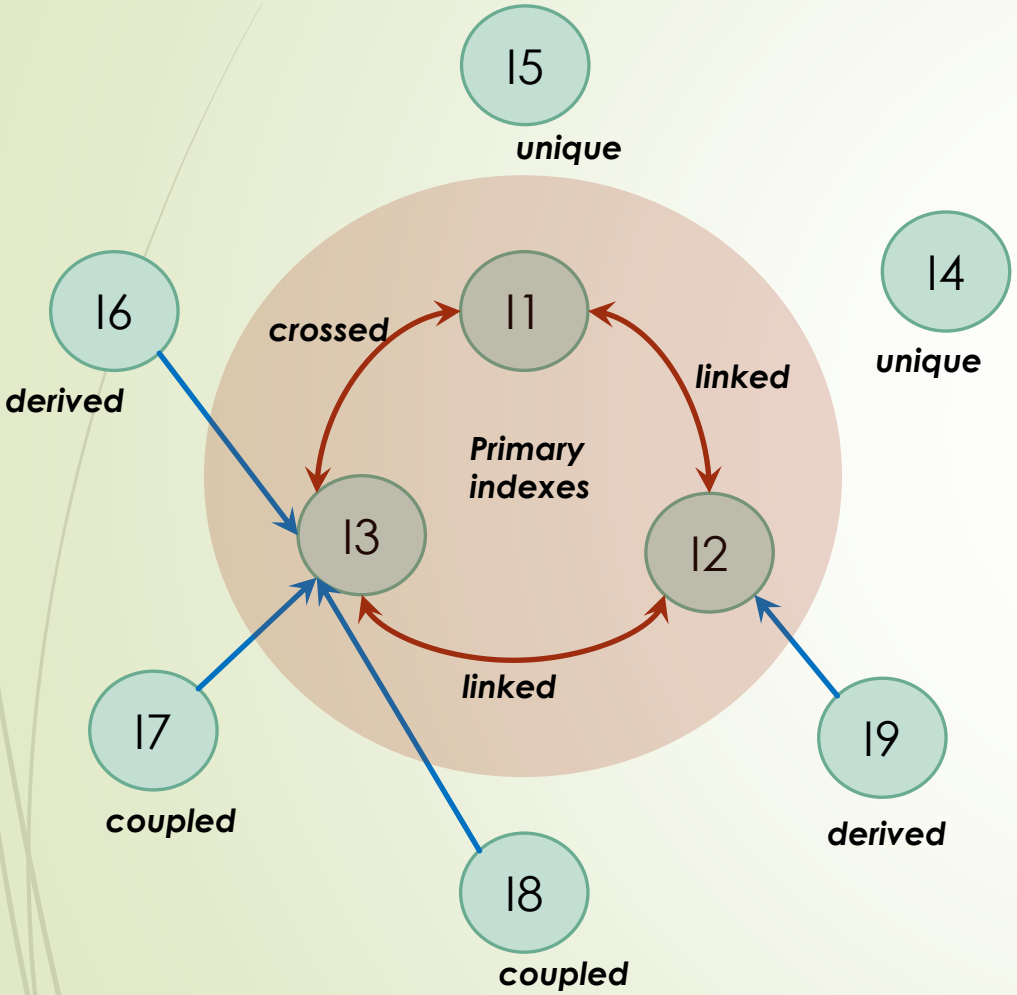

**Example :**

| Name | City | Age | Subject | score |
|---|---|---|---|---|
| Paul | Paris | 16 | Math | 10 |
| Lea | Paris | 15 | Math | 12 |
| Lea | Paris | 15 | English | 15 |

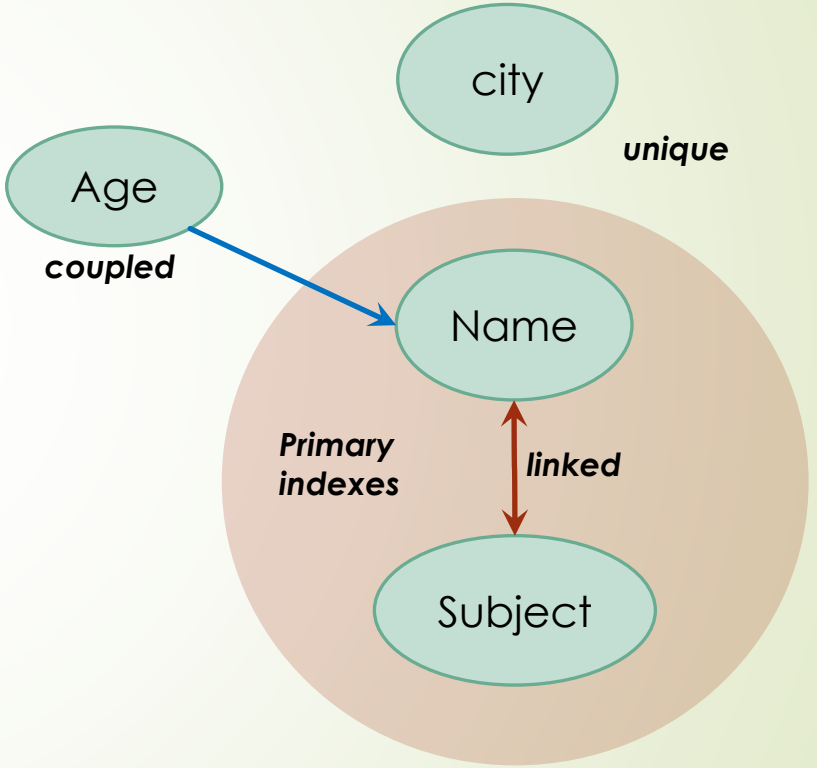
iindex / result

| Name | City | Age | Subject | score |
|---|---|---|---|---|
| 1 | 0 | 0 | 1 | 10 |
| 0 | 0 | 1 | 1 | 12 |
| 0 | 0 | 1 | 0 | 15 |

# 2 - Ilist -> Canonical format

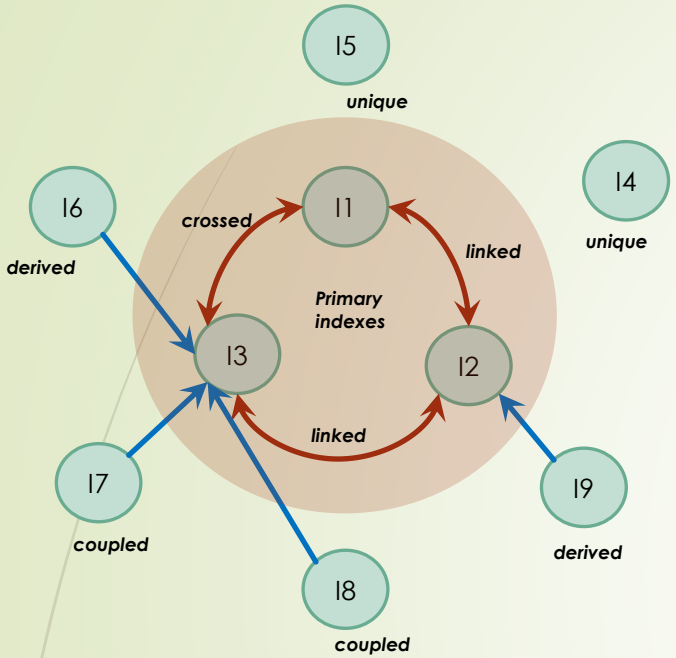



**Example :**

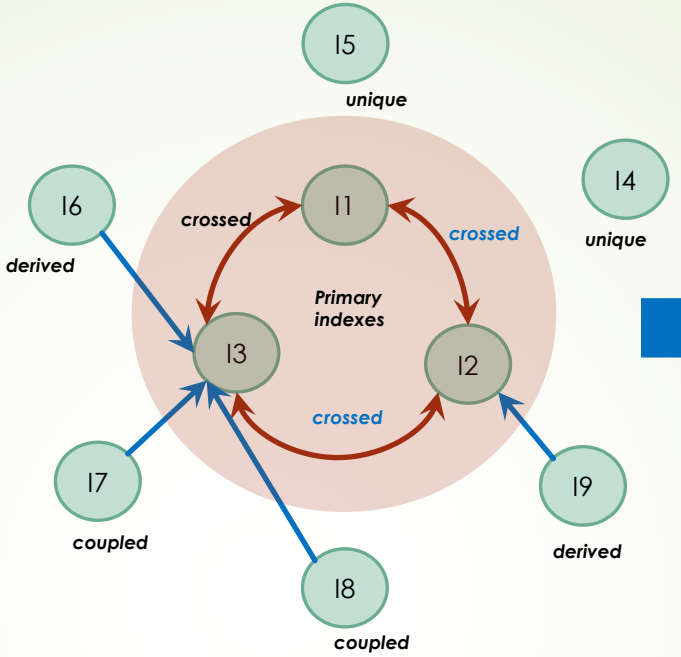



*Explanation : See Appendix*

# 3 - Xarray generation

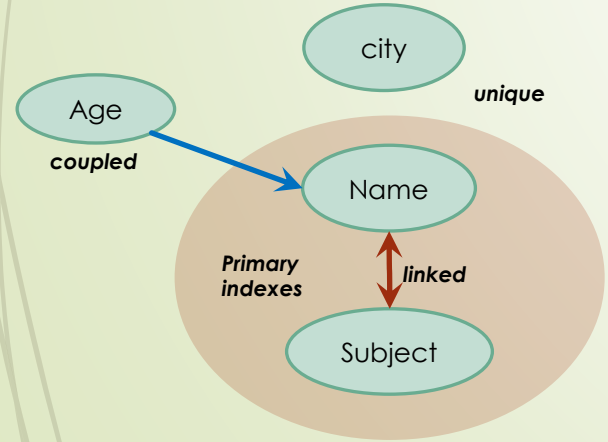



Full




- Primary indexes
  -> **Xarray dims**
- Derived/coupled indexes
  -> **Xarray coords**
- Indexed value
  -> **Xarray data**
- Unique index
  -> **Xarray attrs**

**Example :**




```
In [353]: il = Ilist.Iedic({'score'   : [10,12,15]},
     ...:                  {'name'    : ['Paul', 'Lea', 'Lea'],
     ...:                   'city'    : ['Paris', 'Paris', 'Paris'],
     ...:                   'age'     : [16,15,15],
     ...:                   'subject' : ['math', 'math', 'english']})

In [354]: il.to_xarray(fillvalue=math.nan)
Out[354]:
<xarray.DataArray 'score' (name: 2, subject: 2)>
array([[15., 12.],
       [nan, 10.]])
Coordinates:
  * name     (name) <U4 'Lea' 'Paul'
    age      (name) int32 15 16
  * subject  (subject) <U7 'english' 'math'
Attributes:
    city:     Paris
```

# Appendix – Indexed List

*https://github.com/loco-philippe/Environnemental-Sensing/blob/main/documentation/Ilist_technical.pdf*

https://loco-philippe.github.io/ES/ilist.html