

ICE

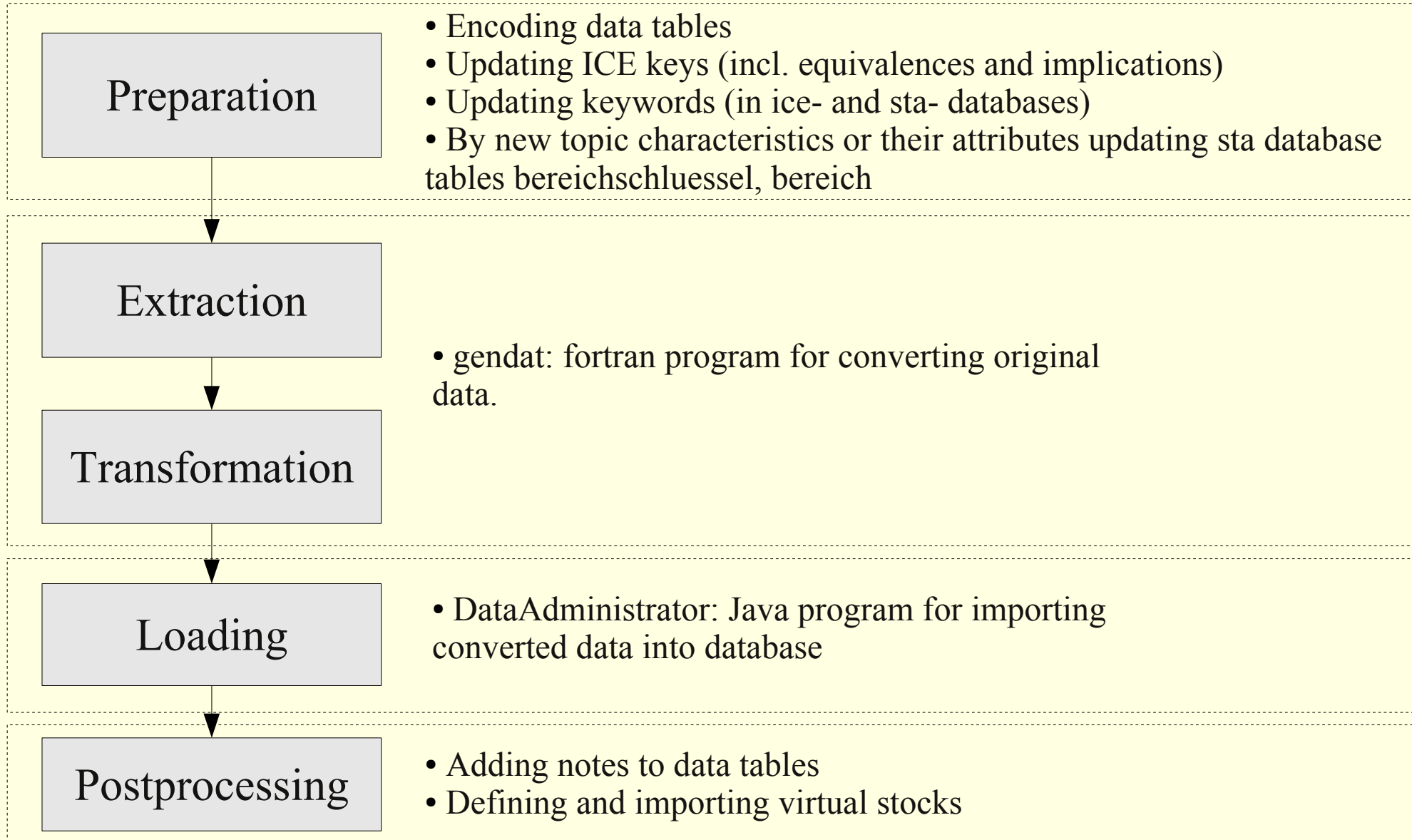
Importing Aggregated Data Into ICE System

Presentation on the 13th of May, 2008

Developed by *ICE* - Group, HIS Hochschul-Informations-System

HIS 

Aggregated Data Import Workflow



Encoding Initial File

Aggregated Data: Preparation

Table 05.01
Teaching staff of Universities according to staff category and academic stream

| University / Faculty | Professor | Associate Professor | Senior Lecturer | Lecturer / Prob. Lecturer | Academic Support Staff | Total |
|-------------------------|------------|---------------------|-----------------|---------------------------|------------------------|------------|
| COLOMBO | | | | | | |
| Arts | 14 | 11 | 57 | 34 | 32 | 148 |
| Education | 1 | 4 | 13 | 10 | 7 | 28 |
| Law | | 1 | 9 | 10 | 2 | 22 |
| Medicine | 25 | 5 | 47 | 45 | 2 | 124 |
| Science | 12 | 2 | 52 | 21 | | 87 |
| Management Studies | | 1 | 27 | 18 | | 46 |
| Sri Palee Campus | | | 1 | 11 | | 12 |
| Total | 52 | 24 | 206 | 149 | 36 | 467 |
| PERADENYA | | | | | | |
| Arts | 22 | 12 | 93 | 52 | | 179 |
| Agriculture | 22 | 1 | 55 | 26 | 3 | 135 |
| Dental Science | 6 | | 29 | 27 | | 62 |
| Vet. Science | 5 | | 16 | 19 | 4 | 40 |
| Engineering | 8 | 1 | 42 | 29 | 30 | 110 |
| Medicine | 13 | 4 | 40 | 35 | 1 | 92 |
| Science | 25 | 3 | 39 | 22 | 2 | 89 |
| Total | 101 | 21 | 314 | 210 | 61 | 707 |
| JAYEWAR DENEPURA | | | | | | |
| Arts | 11 | 6 | 51 | 30 | 8 | 106 |
| Applied Science | 13 | 2 | 45 | 14 | 1 | 74 |
| Management | 2 | 1 | 73 | 55 | 17 | 131 |
| Medical Science | 14 | 2 | 32 | 49 | 2 | 97 |
| Total | 40 | 11 | 201 | 148 | 8 | 408 |
| KELANIYA | | | | | | |
| Humanities | 5 | 5 | 45 | 19 | | 74 |
| Social Sciences | 4 | 4 | 47 | 20 | 3 | 75 |
| Science | 8 | 9 | 50 | 25 | 38 | 92 |
| Medicine | 11 | 2 | 38 | 38 | 48 | 89 |
| Commerce & Mgt | | | 21 | 23 | 9 | 44 |
| Total | 28 | 20 | 201 | 125 | 81 | 374 |
| MORATUVA | | | | | | |
| Architecture | 4 | | 24 | 11 | | 39 |
| Engineering | 1 | 23 | 4 | 90 | 48 | 167 |
| Information Technology | | 4 | 4 | 3 | | 7 |
| Total | 1 | 27 | 4 | 118 | 61 | 213 |
| JAFFNA | | | | | | |
| Arts | 7 | 8 | 54 | 30 | 20 | 119 |
| Medicine | 3 | 9 | 12 | 17 | 17 | 24 |
| Science | 3 | 23 | 17 | 31 | 21 | 63 |
| Agriculture | | 8 | 10 | 10 | 1 | 21 |
| Management | | | 11 | 14 | | 25 |
| Yavunja Campus | | | | | | 0 |
| Applied Science | | | 4 | 1 | 8 | 14 |
| Business Studies | | | 3 | 2 | 6 | 9 |
| Total | 13 | 10 | 112 | 207 | 59 | 275 |
| RUKUNA | | | | | | |
| Agriculture | 8 | 2 | 14 | 31 | 3 | 58 |
| Humanities & Social So. | 5 | 3 | 43 | 14 | 12 | 79 |
| Medicine | 11 | | 33 | 35 | 4 | 83 |
| Engineering | | | 9 | 30 | 27 | 39 |
| Science | 5 | 1 | 43 | 35 | 24 | 84 |
| Management & Finance | | | 11 | 7 | 12 | 23 |
| Total | 0 | 29 | 0 | 153 | 21 | 168 |

Describing content with ICE-Keys and defining new ICE characteristics or attributes if applicable.

Subject: Teaching staff

102 Staff in Universities
02 Scient. Staff

Time point

605 Time point: yearly
20050 Year 2005

According to Universities

490 Universities
1 Colombo

...

According to academic stream

491 Faculty
1 Art

...

According to staff category

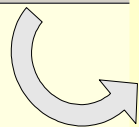
Encoding Initial File

| University Faculty | Professor | Associate Professor | Senior Lecturer | Lecturer/ Prob. Lecturer | Academic Support Staff | Total |
|-------------------------|-----------|---------------------|-----------------|--------------------------|------------------------|-------|
| COLOMBO | 14 | 11 | 57 | 34 | 32 | 148 |
| PERADENYA | 22 | 12 | 93 | 68 | 52 | 179 |
| JAYAWARDE PERERA | 11 | 6 | 11 | 14 | 11 | 54 |
| KELANIYA | 5 | 5 | 5 | 5 | 5 | 25 |
| HOBARTLYN | 4 | 4 | 4 | 4 | 4 | 20 |

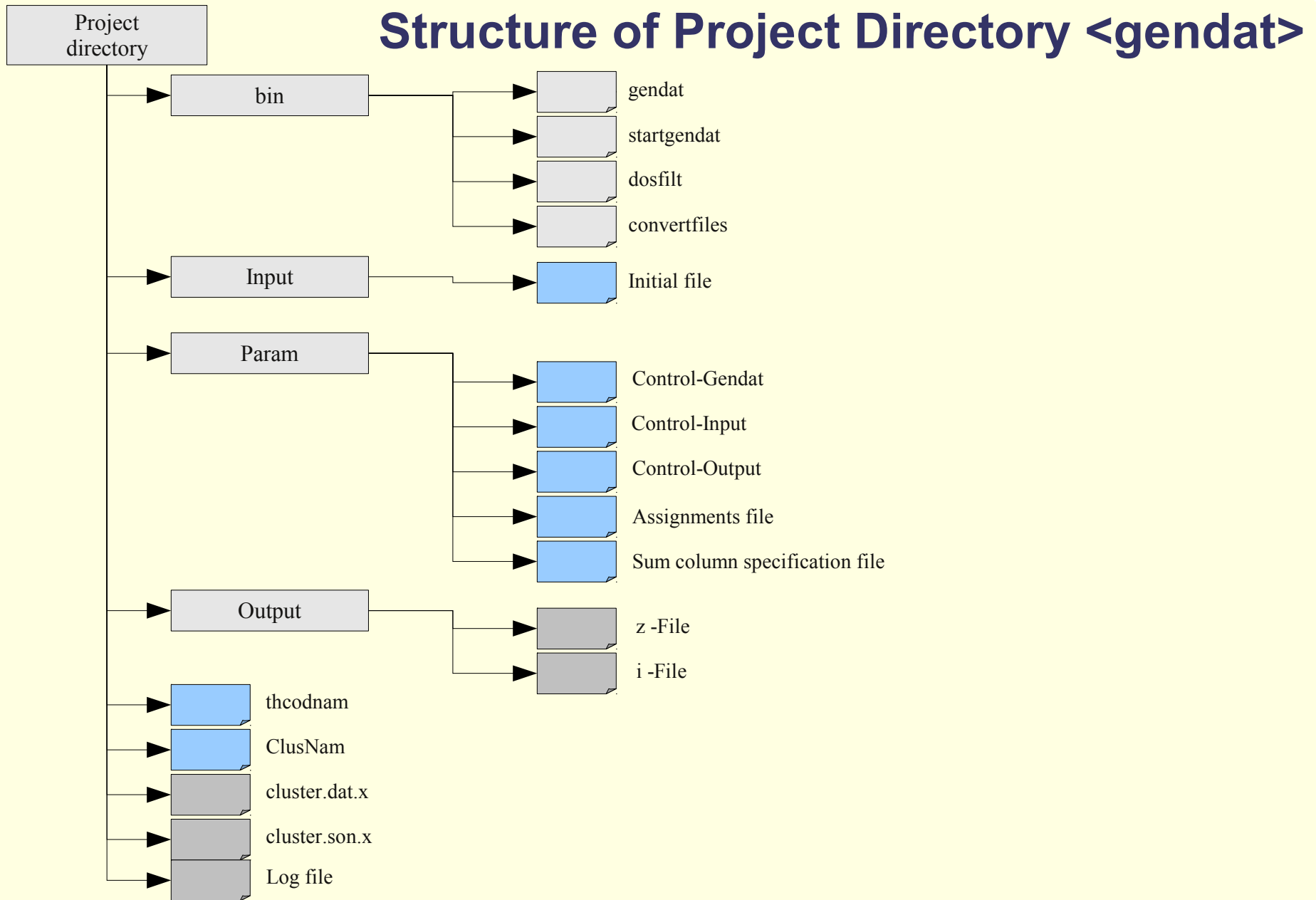
Each sum is unambiguously coded with ICE keys ...

Aggregated Data: Preparation

... file is saved as text file



| | | | | | | | | | | | | | |
|-------|-----|-----|-----|---|-----|---|----|---|-----|-----|-----|----|----|
| 492 | 1 | 1 | 2 | 2 | 3 | 3 | 4 | 4 | 5 | 5 | | | |
| 493 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | | | |
| 605 | 102 | 490 | 491 | | | | | | | | | | |
| 20050 | 2 | 1 | 1 | - | 14 | - | 11 | - | 57 | - | 34 | - | 32 |
| 20050 | 2 | 1 | 1 | - | 1 | - | 4 | - | 13 | 7 | 10 | - | - |
| 20050 | 2 | 1 | 1 | - | - | - | 1 | - | 9 | - | 10 | - | 2 |
| 20050 | 2 | 1 | 5 | - | 25 | - | 5 | - | 47 | - | 45 | - | 2 |
| 20050 | 2 | 1 | 2 | - | 12 | - | 2 | - | 52 | - | 21 | - | - |
| 20050 | 2 | 1 | 3 | - | - | - | 1 | - | 27 | - | 18 | - | - |
| 20050 | 2 | 1 | 3 | - | - | - | - | - | 1 | - | 11 | - | - |
| 20050 | 2 | 1 | 0 | - | 52 | - | 24 | - | 206 | 7 | 149 | - | 36 |
| 20050 | 2 | 2 | 1 | - | 22 | - | 12 | - | 93 | 68 | 52 | - | - |
| 20050 | 2 | 2 | 6 | - | 22 | - | 1 | - | 55 | - | 26 | 3 | 31 |
| 20050 | 2 | 2 | 5 | - | 6 | - | - | - | 29 | 28 | 27 | - | - |
| 20050 | 2 | 2 | 6 | - | 5 | - | - | - | 16 | 4 | 19 | - | - |
| 20050 | 2 | 2 | 4 | - | 8 | - | 1 | - | 42 | - | 29 | 74 | 30 |
| 20050 | 2 | 2 | 5 | - | 13 | - | 4 | - | 40 | 42 | 35 | 1 | - |
| 20050 | 2 | 2 | 2 | - | 25 | - | 3 | - | 39 | 107 | 22 | 2 | - |
| 20050 | 2 | 2 | 0 | - | 101 | - | 21 | - | 314 | 249 | 210 | 80 | 61 |
| 20050 | 2 | 3 | 1 | - | 11 | - | 6 | - | 51 | 3 | 30 | 9 | 8 |
| 20050 | 2 | 3 | 2 | - | 13 | - | 2 | - | 45 | 1 | 14 | 1 | - |
| 20050 | 2 | 3 | 3 | - | 2 | - | 1 | - | 73 | 4 | 55 | 17 | - |
| 20050 | 2 | 3 | 5 | - | 14 | - | 2 | - | 32 | 2 | 49 | - | - |
| 20050 | 2 | 3 | 0 | - | 40 | - | 11 | - | 201 | 10 | 148 | 27 | 8 |



Gendat Parameter Files: Control-Gendat

Gives names to files to convert and assigns them to specific Fortran channel numbers

```

F 30      Input/staff_2005.prn
F 22      Param/staff_2005.cout
F 41      Param/staff_2005.cinp
F 20      44      Output/staff_2005.z
F 21      Output/staff_2005.i
F 14      72      ClusNam
F 40      Param/staff_2005.assign
F 46      Param/staff_2005.colspec
    
```

| Channel number | Description |
|----------------|----------------------------------|
| 14 | ClusNam |
| 20 | z-File (Name max. 20 characters) |
| 21 | i-File (Name max. 20 characters) |
| 22 | Control-Output file |
| 23 | s-File (Name max. 20 characters) |
| 30 | Input-File |
| 40 | Assignments file |
| 41 | Control-Input file |
| 46 | Sum column specification file |

- First name parts of z-, i- and s-files should be equal
- Rows starting with blank or semicolon are treated as comment rows
- Paths can be absolute or relative to project directory

Gendat Parameter Files: Control-Input

```

;Reading-Block
;Reading format for value columns
(48X,12I12)
;Reading format for variables columns
(4A12)
;Number of decimal places
NACHKOMMASTELLEN 0
;Special value
;SONDERZEICHEN - - -
;Number of variables for coding,
;Number of rows in the beginning that should not be read
  6+   3
;Variables in detail
  1  605   32005020050
  2  102   3   2   2
  3  490   3   1   1   2   2   3   3   4   4
  3  490   3   5   5   6   6   7   7   8   8
  3  490   3   9   9  10  10  11  11  12  12
  3  490   3  14  14  13  13   0   0
  4  491   3   1   1   5   5   2   2   3   3
  4  491   3   0   0   6   6   4   4
  5 -492   1   1   1   2   2   3   3   4   4
  5 -492   1   5   5   0   0
  6 -493   1   2   2   1   1
  6 -493   4   0   0
99
;Dummy-Block
99
;Aggregation-Block
99
;Summations-Block
493   0 s +   1 +   2
99
;Divisions-Block
99

```

Describes input file and controls reading and importing processes of it

Reading-Block

- Reading format for value columns
- Reading format for variables columns
- Number of decimal places
- Number of variables for coding
- Variables in detail (without administration characteristics)
 - Running number of variable
 - Variable number
 - Action type
 - 1- Variable value is not read, but assigned
 - 2- Variable value is read (external attributes are defined over assignments table)
 - 3- Variable value is read as ICE-Key
 - 4- Variable value is a result of summation or division
 - 6- Variable value are read from separate files
 - From - Attribute, to - Attribute

Dummy-Block

Dummies are variable values, which are used for converting, but never displayed.

Aggregation-Block

Aggregations are defined, if sum should be coded with several ICE keys.

Summations-Block

Divisions-Block

Gendat Parameter Files: Control-Output

```
05 11 (i2,i5,i2,i1,i1)
2007180000100001
2002010000100001
2007010000200002
2001020000200002
5006052005020050
2004900000000014
1004910000000006
1004920000000005
1 -49300000000002
```

Describes output file

- May contain neither empty rows nor rows with comments
- 1. Row
 - Number of variables in Index
 - Number of characters in Index
 - Fortran-Format for Index-Output
- 2. Row – Data source
- 3. Row - Presentation art
- 4. Row – Data quality
- 5. Row – Topic area
- ... other variables ...
- Last row – column variable

Gendat Parameter Files: Assignments File

```

; Reading format for the file
(7x,i7,i10,x,i10)
;      charact ICE_Attrib External_key
STB    502      1          0
STB    502      2          1
STB    305     11         05
STB    305     15         10
STB    305     21         15
STB    305     22         25
STB    305     23         35
STB    305     24         40
STB    305     25         45
STB    305     26         50
STB    305     18         55
STB    305     30         60
STB    305     40         65
STB    403     11          1
STB    403     12          2
STB    403     13          3
STB    403     14          4
STB    403     15          5
STB    403     16          6
STB    403     17          7
STB    403     18          8
STB    403     19          9
STB    403     26         16
-1
    
```

Assigns external (foreign) attributes to ICE keys

Assignments Table Structure

- 1. Column – ICE-Characteristic number
- 2. Column – ICE-Attribute number
- 3. Column – External attribute

First file row contains Fortran-Format for reading assignments table

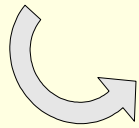
Gendat Parameter Files: Sum Column Specification File

Describes how output file columns are coded

File rows structure

- 1. Column – ICE-Characteristic number
- Other columns - ICE-Attributes
- All column values are 7-digit
- max.10 attributes per row
- max.100 value columns

| | A | B | C | D | E | F | G | H | I |
|----|---------|--------------------|-----------|---------------------|-----------------|--------------------------|------------------------|-------|--------|
| 1 | | 492 | 1 | 2 | 3 | 4 | 5 | 0 | 0 |
| 2 | | 501 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| 3 | | | | | | | | | |
| 4 | | University Faculty | Professor | Associate Professor | Senior Lecturer | Lecturer/ Prob. Lecturer | Academic Support Staff | Total | |
| 5 | | | | | | | | male | female |
| 6 | | | | | | | | | |
| 7 | | | | | | | | | |
| 8 | | | | | | | | | |
| 9 | COLOMBO | Arts | | | | | | | |
| 10 | | Education | | | | | | | |
| 11 | | Law | | | | | | | |
| 12 | | Medicine | | | | | | | |
| 13 | | Science | | | | | | | |
| 14 | | Management Studies | | | | | | | |
| 15 | | Sri Palee Campus | | | | | | | |



```

; Reading format for this file is (i7,7i7)
  492      1      2      3      4      5      0      0
  501      0      0      0      0      0      1      2
    
```

Gendat thcodnam File

```
...
10100001 Persons qualified to enter higher education
10100002 Undergraduate admitted
10100003 New entrants (1st subject-related semester)
10100004 Enroled students
10100005 Examinations
10100006 Graduate output
10100007 Persons doing their doctorate
10100008 Persons doing their postdoctoral lecture qualification
10100009 New entrants (1st subject-related semester) relating to graduates
10100011 Changer of degree programmes
10100012 Changer of academic subjects
10100013 Changer of institutes of higher education
10100014 Student dropouts
10100016 Synopses/mixed groups
10100021 Unemployed graduates
10100022 German students abroad
10100023 German graduates with study experience abroad
10100024 Students who passed final examinations
10100025 Students to UNESCO-recordal
10100026 Students abroad to UNESCO-recordal
10100027 New entrants with completed vocational training
...
```

Contains all ICE keys, incl. those being used for coding data block to convert

Thcodnam – Row Structure

- 3-digit: ICE-Characteristic number
- 5-digit: ICE-Attribute number
- 8. Column - Blank
- from 9. Column – Attribute label

Gendat ClusNam File

| | | | |
|------|---|---------------|---------------|
| 0101 | 1 | cluster.dat.1 | cluster.son.1 |
| 0102 | 2 | cluster.dat.2 | cluster.son.2 |
| 0103 | 3 | cluster.dat.3 | cluster.son.3 |
| 0104 | 4 | cluster.dat.4 | cluster.son.4 |
| 0105 | 5 | cluster.dat.5 | cluster.son.5 |
| 0106 | 6 | cluster.dat.6 | cluster.son.6 |
| 0107 | 7 | cluster.dat.7 | cluster.son.7 |
| 0108 | 8 | cluster.dat.8 | cluster.son.8 |
| 0109 | 9 | cluster.dat.9 | cluster.son.9 |

Contains all topic areas available in ICE-Key System and also manages Cluster files assigned to them

ClusNam – Row Structure

- 4-digit: Number of rows in cluster.son file
- 3-digit: ICE-topic area characteristic number
- 2-digit: Blanks
- 1-digit: Topic area number
- 14-digit: Name of cluster.dat-File
- 30-digit: Name of cluster.son-File
- 18-digit: Name of cluster.son-File

Important! Check size of cluster.son files prior converting

Performing Converting

- Adjust number of rows entries in ClusNam to actual number of rows in cluster.son files prior converting

- If ClusNam not available, create it from ClusNam.ori

```
cp ClusNam.ori ClusNam
```

- Format initial- and parameter files in Unix

```
./bin/convertfiles
```

- Start Gendat

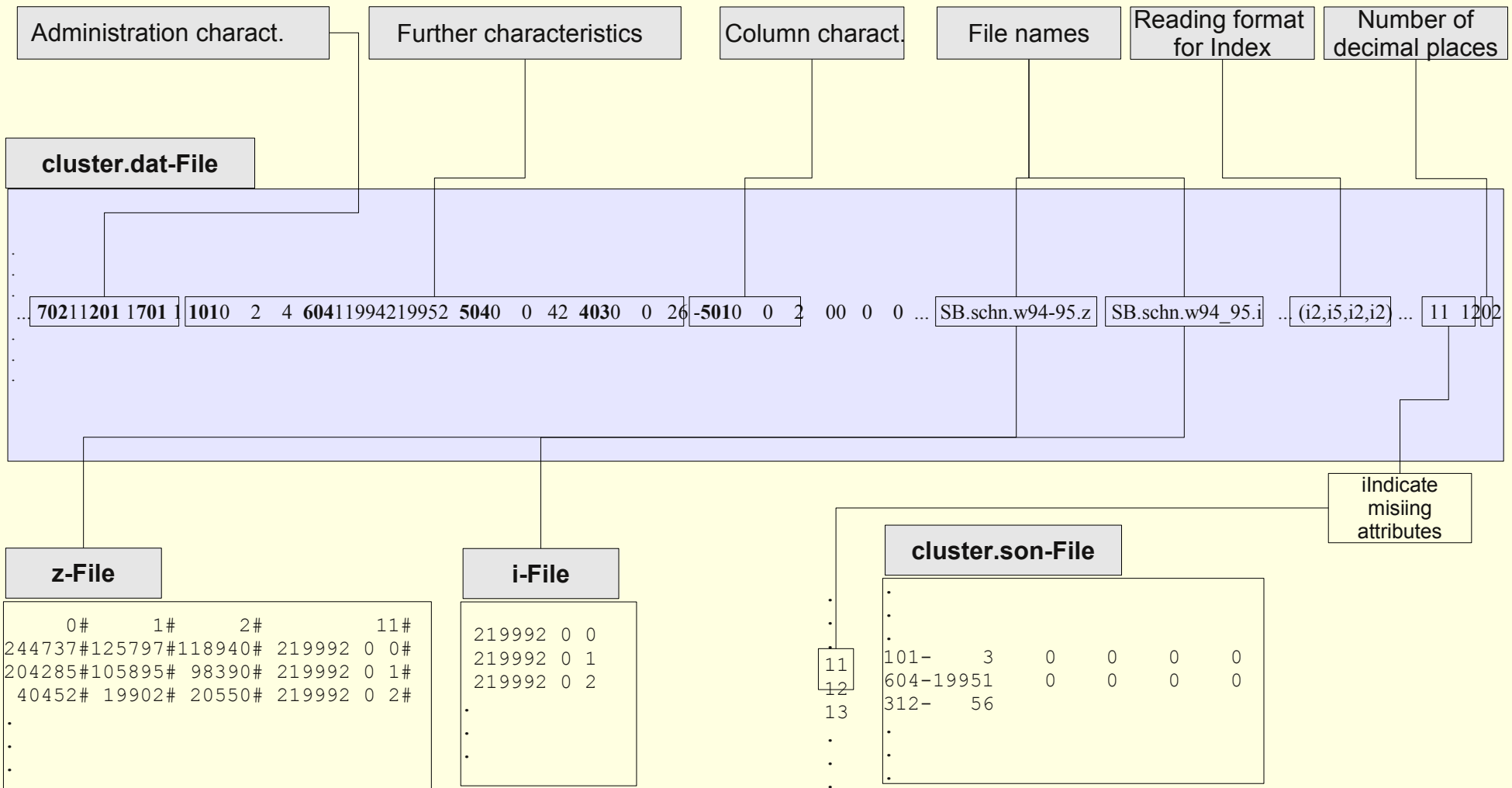
```
./bin/startgendat <NameOfControlGendatFile>
```

- Review log file

- Check result files (z-file, i-file, cluster.dat-file, cluster.son-file)

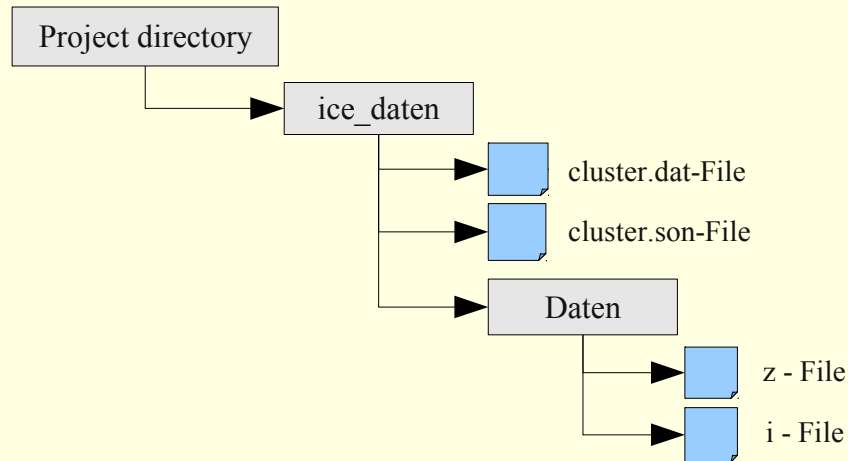
Results of Converting

Aggregated Data: Converting



Importing into Database

- Create project directory with the following structure:



- Change to directory with ICE-Server:

```
cd %TOMCAT_HOME%/webapps/iceproject/WEB-INF/classes
```

- Set classpath:

```
export CLASSPATH=./:/path/to/jdbcdriver:/path/to/log4j.jar
```

- Run DataAdministrator application in package dbimport:

```
/path/to/java dbimport.DataAdministrator [/path/to/project]  
[number of cluster.dat] [part of database] [check details  
j/n] [tabnr to start (optional)] [>name of log file]
```

Deleting Data Stocks

- Change to directory with ICE-Server:

```
cd %TOMCAT_HOME%/webapps/iceproject/WEB-INF/classes
```

- Then JDBC drivers and logging libraries (log4j) must be added to classpath. You can do it either by setting a system environment variable CLASSPATH or by using an option `-classpath` of java binary:

```
export CLASSPATH=./path/to/jdbcdriver:/path/to/log4jjar
```

- Using application `ImportRueckgaengig` in package `dbimport` you can delete either single (`tabid`) or all lastly imported (`all`) (specified in table `tabellenimportiert`) data tables :

```
/path/to/java dbimport.ImportRueckgaengig [tabid|all]  
[database segment]
```

or (with `-classpath` option):

```
/path/to/java -classpath  
./path/to/jdbcdrivers.jar:/path/to/log4.jar  
dbimport.ImportRueckgaengig [tabid|all] [database segment]
```

Adding Notes

- Notes are additional text information that could be added to a physical stock.
- Notes refer to a key in data table and always have time reference in addition.

- Java - Tool for inserting notes NotesAdmin.jar can be run on any machine:

```
path/to/java -jar NotesAdmin.jar
```

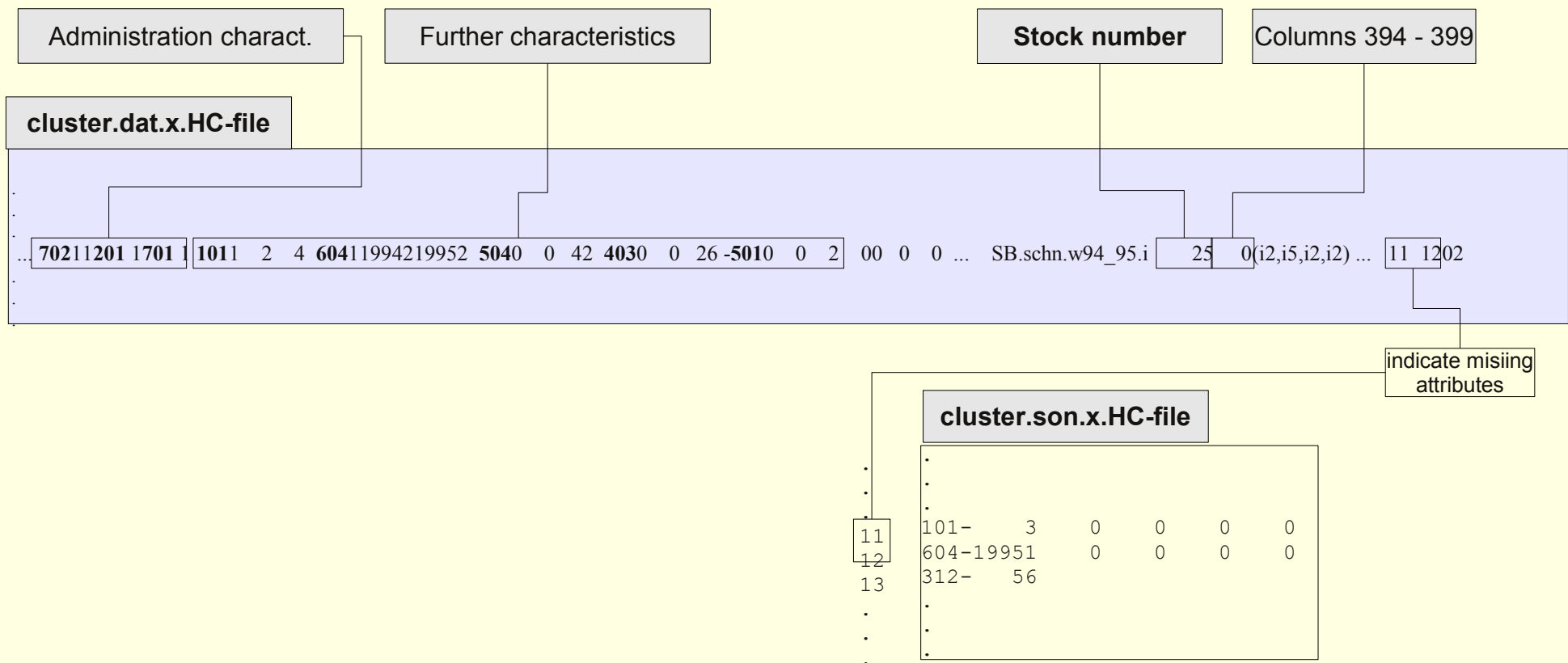
- Tool requires ICE-Listener server-side, listening to Port 50 per default. You start listener like this:

```
path/to/java server.TabServer [portnr]
```

Configuration file `config/ICETab.cfg` server-side contains database connection settings and should be customized prior running the program.

Virtual Stocks

- represent abstract view to imported data tables
- are used by user while defining requests
- are created on basis of cluster-files and saved as .HC-files



Updating Virtual Stocks

- Change to directory with ICE-Server:

```
cd %TOMCAT_HOME%/webapps/iceproject/WEB-INF/classes
```

- Set classpath:

```
export CLASSPATH=./path/to/jdbcdriver:/path/to/log4j.jar
```

- Use VirtualStockGenerator application in package tools to create BestandSchluessel file from HC-files:

```
/path/to/java tools.VirtualStockGenerator [/path/to/stock  
summary] [database segment]
```

- Use VirtualStockReader application in package tools to import BestandSchluessel into database:

```
/path/to/java tools.VirtualStockReader [/path/to/stock  
summary] [database segment]
```