International Tables for Crystallography (2006). Vol. G, Chapter 3.7, pp. 199–205.

3.7. Classification and use of image data

BY H. J. BERNSTEIN

3.7.1. Introduction

This chapter describes the categories and organization of data items defined in the CBF/imgCIF dictionary. The classification of image data applies to both Crystallographic Binary File (CBF) and Image-supporting Crystallographic Information File (imgCIF) representations. An introduction to CBF data and construction is given in Chapter 2.3. Full details of the CBF/imgCIF dictionary are given in Chapter 4.6.

The main reason for introducing the new items defined in the CBF/imgCIF dictionary was to extend the mmCIF dictionary (Chapter 3.6) to allow the storage of synchrotron diffraction images. However, these items are also important in other fields that use binary image data, including the publication of articles, the creation of web pages and the production of movies.

Data categories in the CBF/imgCIF dictionary can describe one-, two- and three-dimensional array detectors that output data organized by time and/or wavelength. The categories defined at present support modular data that can be extended for future applications without having to make fundamental structural changes. For example, it is anticipated that additional data items will be needed soon to allow higher-dimensional data representations and more complex data structures; these should be accommodated easily.

The CBF/imgCIF dictionary consists of three groups of categories of data items: the ARRAY_DATA group, the AXIS group and the DIFFRN group (Table 3.7.1.1). All fall within the 'Experimental measurements' classification of Table 3.1.10.1. The DIFFRN group already exists in the mmCIF dictionary (Section 3.6.5.2; see also Section 3.2.2.2) and describes the diffraction data and their measurement. Definitions in the CBF/imgCIF dictionary extend and in some cases restate the definitions in the mmCIF dictionary.

The data categories defined in the CBF/imgCIF dictionary are described in this chapter. Table 3.7.1.1 lists the formal category groups declared in the dictionary and the sections of this chapter in which they are discussed. Each section is divided into subsections describing a single category or a small set of closely related categories. Within each subsection, the data names within the relevant categories are listed. Category keys, pointers to parent data items and aliases to data items in the mmCIF dictionary are indicated.

The data collected in an experiment are organized into scans. Each scan consists of one or more frames. Each frame consists of one or more data arrays. The logical data in the data arrays need to be described in terms of physical arrays of image elements. The axes of the laboratory coordinate system needed to describe the physical positions of the image elements and the positioning of the specimen are given in the AXIS category. The axes used for the positioning systems for the specimen and the detector are constructed in the same laboratory coordinate system.

Affiliation: HERBERT J. BERNSTEIN, Department of Mathematics and Computer Science, Kramer Science Center, Dowling College, Idle Hour Blvd, Oakdale, NY 11769, USA.

Table 3.7.1.1. *Category groups defined in the CBF/imgCIF dictionary*

Section	Category group	Subject covered
Experimental measurements		
3.7.2	ARRAY_DATA	Binary image data
3.7.3	AXIS	Axes required to specify the data collection
3.7.4	DIFFRN	Diffraction experiment

The DIFFRN_DETECTOR_AXIS category relates detector elements to axes. The DIFFRN_MEASUREMENT_AXIS category relates goniometers to axes. The DIFFRN_SCAN_AXIS and DIFFRN_SCAN_FRAME_AXIS categories relate scans to overall axis settings and individual frames to frame-by-frame axis settings, respectively.

The organization of the data in the collected arrays of data is given in the ARRAY_STRUCTURE_LIST category and the physical settings of axes for the centres of pixels that correspond to data points are given in the ARRAY_STRUCTURE_LIST_AXIS category.

3.7.2. Binary image data

The six categories that collectively define the relationship between the sequences of octets in arrays of binary data and the information in the images those octets represent are as follows:

ARRAY_DATA group
The image data (§3.7.2.1)
ARRAY_DATA
Array elements (§3.7.2.2)
ARRAY_ELEMENT_SIZE
Intensities (§3.7.2.3)
ARRAY_INTENSITIES
Organization and encoding of array data (§3.7.2.4)
ARRAY_STRUCTURE
ARRAY_STRUCTURE_LIST
ARRAY_STRUCTURE_LIST_AXIS

3.7.2.1. The image data

Data items in this category are as follows:

ARRAY_DATA
• **_array_data.array_id**
→ **_array_structure.id**
• **_array_data.binary_id**
_array_data.data

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

Each value of the **_array_data.data** data item is a sequence of octets representing a binary image. **_array_data.array_id** and **_array_data.binary_id**, taken together, uniquely identify each image. The value of **_array_data.array_id** is a pointer to **_array_structure.id** to provide the relationship between the sequence of octets and the logical structure of the image. Since multiple images may have the same logical structure, the purpose of **_array_data.binary_id** is to ensure that each image has a unique identifier.

3.7.2.2. Array elements

Data items in this category are as follows:

ARRAY_ELEMENT_SIZE
```
• _array_element_size.array_id
        → _array_structure.id
• _array_element_size.index
        → _array_structure_list.index
  _array_element_size.size
```

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

The value of the **_array_element_size.size** data item is a size in metres of an image element (a pixel or voxel). The direction of the measurement is given in each dimension by **_array_element_size.index**. The array structure specifying the organization of the dimensions is referenced by the value of **_array_element_size.array_id**, which is a pointer to **_array_structure.id**. The value of **_array_element_size.index** is a pointer to **_array_structure_list.index**. For data organized into rectangular arrays of pixels or voxels, this gives the spatial dimensions of the individual image elements.

3.7.2.3. Intensities

Data items in this category are as follows:

ARRAY_INTENSITIES
```
• _array_intensities.array_id
        → _array_structure.id
• _array_intensities.binary_id
        → _array_data.binary_id
  _array_intensities.gain
  _array_intensities.gain_esd
  _array_intensities.linearity
  _array_intensities.offset
  _array_intensities.overload
  _array_intensities.scaling
  _array_intensities.undefined_value
```

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

The relationship between the data values for individual image elements and the number of incident photons can be complex. The data items in the ARRAY_INTENSITIES category provide information about this relationship. The value of **_array_intensities.linearity** states the type of relationship, and the values of **_array_intensities.array_id** and **_array_intensities.binary_id** identify the array structure and the image being discussed. The other items are used in different ways depending on the relationship. If the value of **_array_intensities.linearity** is `raw`, then the image elements hold uninterpreted raw data values from the detector, *e.g.* for calibration. If the value of **_array_intensities.linearity** is `linear`, then the count in an image element is proportional to the incident number of photons by the value of **_array_intensities.gain**. The standard uncertainty (estimated standard deviation) of the gain may be given in **_array_intensities.gain_esd**. The value used for this should be estimated from a good understanding of the physical characteristics of the experimental apparatus. If the value of **_array_intensities.linearity** is `offset`, then the value of **_array_intensities.offset** should be added to the image element value. If the value of **_array_intensities.linearity** is `scaling`, `scaling_offset`, `sqrt_scaled` or `logarithmic_scaled`, the necessary scaling factor is given by the value of **_array_intensities.scaling**. In all cases, the scaling factor is applied to the image element value before the other operations are applied. In the first case, only simple scaling is used. In the second case, the value of **_array_intensities.offset** is added after scaling. In the third case, the scaled value is squared. In the final case, 10 is taken to the power given by the scaled value.

3.7.2.4. Organization and encoding of array data

Data items in these categories are as follows:

(*a*) ARRAY_STRUCTURE
```
• _array_structure.id
  _array_structure.byte_order
  _array_structure.compression_type
  _array_structure.encoding_type
```

(*b*) ARRAY_STRUCTURE_LIST
```
• _array_structure_list.array_id
• _array_structure_list.index
        → _array_structure.id
  _array_structure_list.axis_set_id
  _array_structure_list.dimension
  _array_structure_list.direction
  _array_structure_list.precedence
```

(*c*) ARRAY_STRUCTURE_LIST_AXIS
```
• _array_structure_list_axis.axis_id
        → _axis.id
• _array_structure_list_axis.axis_set_id
        → _array_structure_list.axis_set_id
  _array_structure_list_axis.angle
  _array_structure_list_axis.angle_increment
  _array_structure_list_axis.angular_pitch
  _array_structure_list_axis.displacement
  _array_structure_list_axis.displacement_increment
  _array_structure_list_axis.radial_pitch
```

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

The data items in the ARRAY_STRUCTURE category show how the stream of octets in a binary image is to be reorganized into words of an appropriate size. Each possible encoding is identified by a value of **_array_structure.id**. In most cases, large images will have been compressed. The type of compression used is given by **_array_structure.compression_type**. Once a stream of octets has been decompressed, it can be organized into words. The type of each word is given by the value of **_array_structure.encoding_type** and the order of mapping octets onto words, most significant octet first ('big-endian') or least significant octet first ('little-endian'), is given by the value of **_array_structure.byte_order**.

The data items in the ARRAY_STRUCTURE_LIST category show how the list of words defined by the ARRAY_STRUCTURE category should be organized into image arrays. The value of **_array_structure_list.array_id** is a pointer to **_array_structure.id**. Each dimension (row, column, sheet *etc.*) of the image is identified by an index, counting from 1, given by **_array_structure_list.index**. The order of nesting of the indices is given by the values of **_array_structure_list.precedence**, with the index of precedence 1 varying most rapidly (*i.e.* having values stored sequentially). The direction of index change for increasing memory location is given by the value of **_array_structure_list.direction**. For a given index, the number of image elements in that direction is given by the value of **_array_structure_list.dimension**.

Data items in the ARRAY_STRUCTURE_LIST_AXIS category describe the physical settings of sets of axes for the centres of pixels that correspond to data points described in the ARRAY_STRUCTURE_LIST category.

In the simplest cases, the physical increments of a single axis correspond to the increments of a single array index. More complex organizations (*e.g.* spiral scans) may require coupled motions along multiple axes.

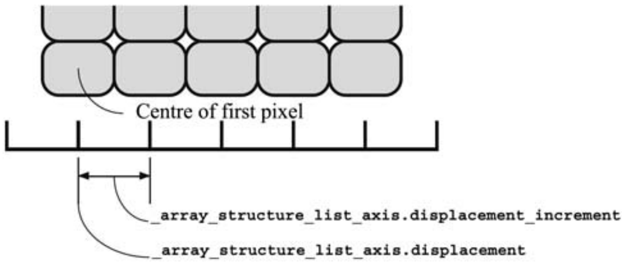

Fig. 3.7.2.1. ARRAY_STRUCTURE_LIST specification of linearly organized image elements.

Note that a spiral scan uses two coupled axes, one for the angular direction and one for the radial direction. This differs from a cylindrical scan for which the two axes are not coupled into one set.

Multiple related axes are gathered together into sets. Each set is identified by the value of the axis set identifier, `_array_structure_list_axis.axis_set_id`, and each axis within a set is identified by the value of `_array_structure_list_axis.axis_id`. Each set given by a value of `*.axis_set_id` is linked to a corresponding value for `_array_structure_list.axis_set_id` to relate settings of the axes in the axis set to particular image elements in ARRAY_STRUCTURE_LIST.

If axes are all independent, no value need be given for `_array_structure_list_axis.axis_set_id`, which is then implicitly given the corresponding value of `_array_structure_list_axis.axis_id`. Each axis given by a value of `_array_structure_list_axis.axis_id` is linked to a corresponding value for `_axis.id` to provide a physical description of the axis. `_array_structure_list_axis.axis_id` and `_array_structure_list_axis.axis_set_id` together uniquely identify a row of data in an ARRAY_STRUCTURE_LIST_AXIS table.

For the remaining data items, there are two important cases to consider: axes that step by Euclidean distance and axes that step by angle. Fig. 3.7.2.1 shows a portion of an array of image elements laid out on a rectangular grid. The starting point of an axis is specified in millimetres by the value of `_array_structure_list_axis.displacement` and the centre-to-centre distance between pixels is specified in millimetres by the value of `_array_structure_list_axis.displacement_increment`.

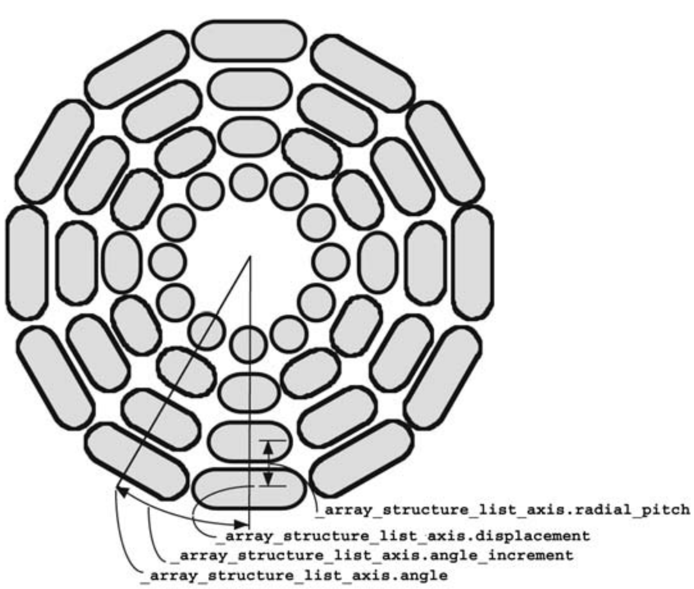

Fig. 3.7.2.2. ARRAY_STRUCTURE_LIST specification of 'constant-angle' image elements in a cylindrical scan. The angular and radial axes are independent. Note that outer-zone image elements are further apart, centre-to-centre, than inner-zone image elements.

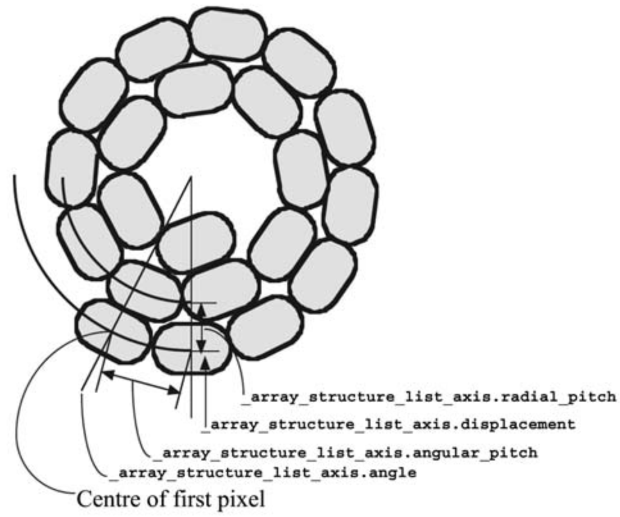

Fig. 3.7.2.3. ARRAY_STRUCTURE_LIST specification of 'constant-velocity' image elements in a cylindrical scan. The angular and radial axes are coupled. Note that outer-zone image elements are the same linear distance apart, centre-to-centre, as the inner-zone image elements.

Fig. 3.7.2.2 shows a portion of an array of image elements laid out in concentric cylinders. The starting point of the angular axis is specified in degrees by the value of `_array_structure_list_axis.angle` and the centre-to-centre angular distance between pixels is specified in degrees by the value of `_array_structure_list_axis.angle_increment`. The starting point of the radial axis is specified by the value of `_array_structure_list_axis.displacement` and the radial distance between cylinders of pixels is specified in millimetres by the value of `_array_structure_list_axis.radial_pitch`. Note that the image elements further from the centre are larger than the image elements closer to the centre.

Fig. 3.7.2.3 shows a portion of a spiral scan array in which the angular and radial axes are coupled. This example represents a 'constant-velocity' scan, in which the size of the image elements does not depend on the distance from the centre. The starting point of the angular axis is again specified in degrees by the value of `_array_structure_list_axis.angle`, but the centre-to-centre distance between pixels is specified in millimetres by the value of `_array_structure_list_axis.angular_pitch`. The coupled radial axis is handled in much the same way as for the uncoupled radial axis in the cylindrical array.

These examples show some of the more common two-dimensional data structures. By coupling an additional axis not in the plane of the first two, regular three-dimensional arrays of data can be represented without additional tags. The categories in the DIFFRN group allow arrays of data to be associated with frames and thereby with time and/or wavelength. More general data structures, for example ones based on dope vectors or hash tables, would require the definition of additional tags, but any data structure (see Aho *et al.*, 1987) that can be handled by a modern computer should be manageable within this framework.

3.7.3. Axes

The category describing the axes required to specify the data collection is as follows:

AXIS group
  AXIS

Data items in this category are as follows:

AXIS
- • _axis.equipment
- • _axis.id
- _axis.depends_on
  → _axis.id
- _axis.offset[1]
- _axis.offset[2]
- _axis.offset[3]
- _axis.type
- _axis.vector[1]
- _axis.vector[2]
- _axis.vector[3]

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

Data items in the AXIS category record the information required to describe the goniometer, detector, source and other axes needed to specify a data collection. The location of each axis is specified by two vectors: the axis itself, given as a unit vector, and an offset to the base of the unit vector. These vectors are referenced to a right-handed laboratory coordinate system with its origin in the specimen, as shown in Fig. 3.7.3.1.

The X axis is aligned to the mechanical axis pointing from the specimen along the principal axis of the goniometer.

The Z axis is defined next. The Z axis is derived from the source axis (the axis running from the sample to the source). If the source axis is orthogonal to the X axis, the source axis is the Z axis. If the source axis is not orthogonal to the X axis, the Z axis is the component of the source axis orthogonal to the X axis. The direction is chosen to form an acute angle with the source axis.

The Y axis is defined last. The Y axis completes an orthogonal right-handed system defined by the X axis and the Z axis (see below).

These axes are based on the goniometer, not on the orientation of the detector, gravity *etc.* The vectors necessary to specify all other axes are given by sets of three components in the order (X, Y, Z). If the axis involved is a rotation axis, it is right-handed, *i.e.* as one views the object to be rotated from the origin (the tail) of the unit vector, the rotation is clockwise. If a translation axis is specified, the direction of the unit vector specifies the sense of positive translation.

Note: This choice of coordinate system is similar to but significantly different from the choice in *MOSFLM* (Lesley & Powell, 2003), in which X is along the X-ray beam (our Z axis) and Z is along the rotation axis.

All rotations are given in degrees and all translations are given in millimetres.

Axes may be dependent on one another. The X axis is the only axis that is strictly connected to the hardware. All other axes are specified by the positions they would assume when the axes upon which they depend are at their zero points.

When specifying detector axes, each axis is specified relative to the beam centre. The location of the beam centre on the detector should be given in the DIFFRN_DETECTOR category in millimetres from the (0, 0) corner of the detector and should be corrected for distortion.

It should be noted that many different origins arise in the definition of an experiment. In particular, as noted above, we need to specify the location of the beam centre on the detector in terms of the origin of the detector, which is, of course, not coincident with the centre of the sample.

Each axis is uniquely identified by the values of `_axis.id` and of `_axis.equipment`. An axis may be a translation axis, a rotation axis or an axis for which the mode of motion is not relevant. The type of axis is specified by the value of `_axis.type`. The base of the axis is specified by the point in the laboratory coordinate system given by the values of `_axis.offset[1]`, `_axis.offset[2]` and `_axis.offset[3]`, and the direction of the axis from that base, as a dimensionless unit vector, is given by `_axis.vector[1]`, `_axis.vector[2]` and `_axis.vector[3]`.

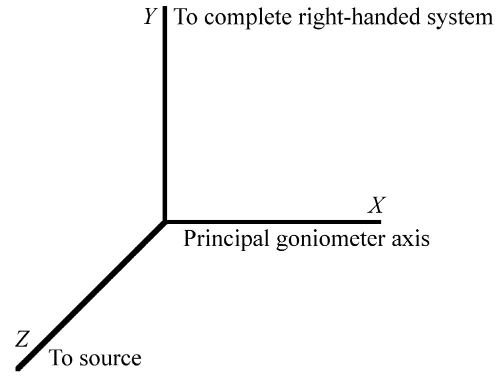

Fig. 3.7.3.1. AXIS laboratory coordinate system. The origin is centred in the specimen.

3.7.4. The diffraction experiment

The categories relating to the diffraction experiment are as follows:

DIFFRN group
- *Frames of data* (§3.7.4.1)
  - DIFFRN_DATA_FRAME
- *The detector apparatus* (§3.7.4.2)
  - DIFFRN_DETECTOR
  - DIFFRN_DETECTOR_AXIS
  - DIFFRN_DETECTOR_ELEMENT
- *Apparatus and instrumentation at the crystal* (§3.7.4.3)
  - DIFFRN_MEASUREMENT
  - DIFFRN_MEASUREMENT_AXIS
- *The radiation source* (§3.7.4.4)
  - DIFFRN_RADIATION
- *Intensity measurements* (§3.7.4.5)
  - DIFFRN_REFLN
- *Diffraction scans* (§3.7.4.6)
  - DIFFRN_SCAN
  - DIFFRN_SCAN_AXIS
  - DIFFRN_SCAN_FRAME
  - DIFFRN_SCAN_FRAME_AXIS

The CBF/imgCIF dictionary extends the mmCIF categories in the DIFFRN group, which are very similar to their corresponding categories in the core CIF dictionary. The DIFFRN group is introduced in the description of the core CIF dictionary in Section 3.2.2.2. Its use in the mmCIF dictionary is described in Section 3.6.5.2, from which we quote: 'The categories in the DIFFRN category group describe the diffraction experiment. Data items in the DIFFRN category itself can be used to give overall information about the experiment, such as the temperature and pressure. Examples of the other categories are DIFFRN_DETECTOR, which is used for describing the detector used for data collection, and DIFFRN_SOURCE, which is used to give details of the source of the radiation used in the experiment. Data items in the DIFFRN_REFLN category can be used to give information about the raw data and data items in the DIFFRN_REFLNS category can be used to give information about all the reflection data collectively.' In this chapter we focus on the CBF/imgCIF extensions.

3.7.4.1. Frames of data

Data items in this category are as follows:

DIFFRN_DATA_FRAME
- • _diffrn_data_frame.detector_element_id
  → _diffrn_detector_element.id
- • _diffrn_data_frame.id
  _diffrn_data_frame.array_id
  → _array_structure.id
  _diffrn_data_frame.binary_id
  → _array_data.binary_id

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

Data items in the DIFFRN_DATA_FRAME category record details about each frame of data. An experiment may produce multiple frames of data and each frame may be constructed from data provided by multiple detector elements. Each complete frame of data is uniquely identified by the value of **_diffrn_data_frame.id**. The detector elements used are specified by values of **_diffrn_data_frame.detector_element_id**, which forms the category key together with **_diffrn_data_frame.id**. **_diffrn_data_frame.detector_element_id** is a pointer to **_diffrn_detector_element.id** in the DIFFRN_DETECTOR_ELEMENT category. The structure of the data in the frame is completed by giving values for **_diffrn_data_frame.array_id** (a pointer to **_array_structure.id**). The particular blocks of data in the frame are specified by giving values of **_diffrn_data_frame.binary_id** (a pointer to **_array_data.binary_id**).

3.7.4.2. The detector apparatus

Data items in these categories are as follows:

(*a*) DIFFRN_DETECTOR
- • _diffrn_detector.diffrn_id
  → _diffrn.id
- • _diffrn_detector.id
  _diffrn_detector.details
  _diffrn_detector.detector
  _diffrn_detector.dtime
  _diffrn_detector.number_of_axes
  _diffrn_detector.type

(*b*) DIFFRN_DETECTOR_AXIS
- • _diffrn_detector_axis.axis_id
  → _axis.id
- • _diffrn_detector_axis.detector_id
  → _diffrn_detector.id

(*c*) DIFFRN_DETECTOR_ELEMENT
- • _diffrn_detector_element.id
- • _diffrn_detector_element.detector_id
  → _diffrn_detector.id
  _diffrn_detector_element.center[1]
  _diffrn_detector_element.center[2]

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item. Items in italics are defined in the mmCIF dictionary.

The DIFFRN_DETECTOR category is defined in the mmCIF dictionary (Section 3.6.5.2; see the detailed discussion in Section 3.2.2.2.4). The CBF/imgCIF dictionary restates the DIFFRN_DETECTOR category, adding new tags. Data items in the DIFFRN_DETECTOR category describe the detector used to measure the scattered radiation, including any analyser and post-sample collimation. In order to allow for multiple detectors, the category key has been extended to include **_diffrn_detector.id** to uniquely identify each detector. If there is only one detector, **_diffrn_detector.id** need not be specified, and it will implicitly default to the value of **_diffrn_detector.diffrn_id** (a pointer to **_diffrn.id** in the DIFFRN category in the mmCIF dictionary). The general class of detector is given by the value of **_diffrn_detector.detector** with the make and model given by the value of **_diffrn_detector.type**. Any special aspects of the detector not covered elsewhere are given by the value of **_diffrn_detector.details**. As in mmCIF, the value of **_diffrn_detector.dtime** gives the deadtime of the detector. Additional data items may need to be added in the future for complex inhomogeneous deadtime situations. In addition, the number of axes can be specified using **_diffrn_detector.number_of_axes**.

Data items in the DIFFRN_DETECTOR_AXIS category associate axes with detectors. Each axis is associated with a detector through the value of **_diffrn_detector_axis.detector_id** (a pointer to **_diffrn_detector.id**). The value of ***.axis_id** (a pointer to **_axis.id**) identifies an axis. Together ***.detector_id** and ***.axis_id** form the category key.

Data items in the DIFFRN_DETECTOR_ELEMENT category record details about the spatial layout and other characteristics of each element of a detector which may have multiple elements, giving the X and Y coordinates of the position of the beam centre relative to the lower left corner of each detector element. Each detector element is identified by the value of **_diffrn_detector_element.id** and the detector of which it is an element is identified by the value of **_diffrn_detector_element.detector_id** (a pointer to **_diffrn_detector.id**).

In most cases, it would be preferable to use the more detailed information provided in the ARRAY_STRUCTURE_LIST and ARRAY_STRUCTURE_LIST_AXIS categories rather than simply specifying the coordinates of the centre of the beam relative to the lower left corner of each detector element.

3.7.4.3. Apparatus and instrumentation at the crystal

Data items in these categories are as follows:

(*a*) DIFFRN_MEASUREMENT
- • *_diffrn_measurement.diffrn_id*
  → _diffrn.id
- • *_diffrn_measurement.device*
- • _diffrn_measurement.id
  _diffrn_measurement.details
  _diffrn_measurement.device_details
  _diffrn_measurement.device_type
  _diffrn_measurement.method
  _diffrn_measurement.number_of_axes
  _diffrn_measurement.specimen_support

(*b*) DIFFRN_MEASUREMENT_AXIS
- • _diffrn_measurement_axis.axis_id
  → _axis.id
- • _diffrn_measurement_axis.measurement_device
  → _diffrn_measurement.device
- • _diffrn_measurement_axis.measurement_id
  → _diffrn_measurement.id

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item. Items in italics are defined in the mmCIF dictionary.

The DIFFRN_MEASUREMENT category is defined in the mmCIF dictionary (Section 3.6.5.2; see the detailed discussion in Section 3.2.2.2.3). The CBF/imgCIF dictionary restates the DIFFRN_MEASUREMENT category, adding new tags. Data items in the DIFFRN_MEASUREMENT category record details about the device used to orient and/or position the crystal during data measurement and the manner in which the diffraction data were measured. To allow for multiple measurement devices, **_diffrn_measurement.id** has been added to the category key. The number of axes is given by the value of **_diffrn_measurement.number_of_axes**. The axes should be described using entries in DIFFRN_MEASUREMENT_AXIS.

Data items in the DIFFRN_MEASUREMENT_AXIS category associate axes with goniometers, just as data items in the DIFFRN_DETECTOR_AXIS category associate axes with detectors.

3.7.4.4. The radiation source

Data items in this category are as follows:

DIFFRN_RADIATION
- • *_diffrn_radiation.diffrn_id*
  - → _diffrn.id
- *_diffrn_radiation.collimation*
- _diffrn_radiation.div_x_source
- _diffrn_radiation.div_y_source
- _diffrn_radiation.div_x_y_source
- *_diffrn_radiation.filter_edge*
- *_diffrn_radiation.inhomogeneity*
- *_diffrn_radiation.monochromator*
- *_diffrn_radiation.polarisn_norm*
- *_diffrn_radiation.polarisn_ratio*
- _diffrn_radiation.polarizn_source_norm
- _diffrn_radiation.polarizn_source_ratio
- *_diffrn_radiation.probe*
- *_diffrn_radiation.type*
- *_diffrn_radiation.wavelength_id*
  - → _diffrn_radiation_wavelength.id
- *_diffrn_radiation.xray_symbol*

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item. Items in italics are defined in the mmCIF dictionary.

The DIFFRN_RADIATION category is defined in the mmCIF dictionary (Section 3.6.5.2; see the detailed discussion in Section 3.2.2.2.2). The CBF/imgCIF dictionary adds the items _diffrn_radiation.div_x_source, *.div_y_source and *.div_x_y_source to specify beam crossfire, and the items _diffrn_radiation.polarizn_source_norm and *.polarizn_source_ratio to provide a definition of polarization relative to the laboratory coordinate system rather than relative to the diffraction plane. The value of the beam crossfire component _diffrn_radiation.div_x_source is the mean deviation in degrees of the X-ray beam from being parallel to the X axis as it illuminates the sample. The value of the beam crossfire component _diffrn_radiation.div_y_source is the mean deviation in degrees of the X-ray beam from being parallel to the Y axis as it illuminates the sample. The value of the beam crossfire component _diffrn_radiation.div_x_y_source is the correlation of the X and Y components. The value of the normal component of the polarization _diffrn_radiation.polarizn_source_norm is the angle in degrees, as viewed from the specimen, between the normal to the polarization plane and the laboratory Y axis as defined in the AXIS category. The dimensionless value of _diffrn_radiation.polarisn_ratio is the ratio $(I_p - I_n)/(I_p + I_n)$, where I_n is the intensity (amplitude squared) of the electric vector of the illumination of the sample normal to the polarization and I_p is the intensity of the electric vector of the illumination of the sample in the plane of polarization. With suitable choices of laboratory axes, the definitions conform to synchrotron conventions. See Chapter 4.6 for a detailed description of these items.

3.7.4.5. Intensity measurements

Data items in this category are as follows:

DIFFRN_REFLN
- • _diffrn_refln.frame_id
  - → _diffrn_data_frame.id
- • *_diffrn_refln.id*
- • *_diffrn_refln.diffrn_id*
- *_diffrn_refln.angle_chi*
- *_diffrn_refln.angle_kappa*
- *_diffrn_refln.angle_omega*
- *_diffrn_refln.angle_phi*
- *_diffrn_refln.angle_psi*
- *_diffrn_refln.angle_theta*
- *_diffrn_refln.attenuator_code*
- *_diffrn_refln.counts_bg_1*
- *_diffrn_refln.counts_bg_2*
- *_diffrn_refln.counts_net*
- *_diffrn_refln.counts_peak*
- *_diffrn_refln.counts_total*
- *_diffrn_refln.detect_slit_horiz*
- *_diffrn_refln.detect_slit_vert*
- *_diffrn_refln.elapsed_time*
- *_diffrn_refln.index_h*
- *_diffrn_refln.index_k*
- *_diffrn_refln.index_l*
- *_diffrn_refln.intensity_net*
- *_diffrn_refln.intensity_sigma*
- *_diffrn_refln.scale_group_code*
- *_diffrn_refln.scan_mode*
- *_diffrn_refln.scan_mode_backgd*
- *_diffrn_refln.scan_rate*
- *_diffrn_refln.scan_time_backgd*
- *_diffrn_refln.scan_width*
- *_diffrn_refln.sint_over_lambda*
- *_diffrn_refln.standard_code*
- *_diffrn_refln.wavelength*
- *_diffrn_refln.wavelength_id*

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item. Items in italics are defined in the mmCIF dictionary.

The DIFFRN_REFLN category is defined in the mmCIF dictionary (Section 3.6.5.2; see the detailed discussion in Section 3.2.2.2.2). Data items in the DIFFRN_REFLN category record details of the intensities measured in the diffraction data set identified by _diffrn_refln.diffrn_id. The CBF/imgCIF dictionary extends the key with _diffrn_refln.frame_id (a pointer to _diffrn_data_frame.id), so that multiple data sets may be recorded.

3.7.4.6. Diffraction scans

Data items in these categories are as follows:

(*a*) DIFFRN_SCAN
- • _diffrn_scan.id
- _diffrn_scan.date_end
- _diffrn_scan.date_start
- _diffrn_scan.frame_id_start
  - → _diffrn_data_frame.id
- _diffrn_scan.frame_id_end
  - → _diffrn_data_frame.id
- _diffrn_scan.frames
- _diffrn_scan.integration_time

(*b*) DIFFRN_SCAN_AXIS
- • _diffrn_scan_axis.axis_id
  - → _axis.id
- • _diffrn_scan_axis.scan_id
  - → _diffrn_scan.id
- _diffrn_scan_axis.angle_start
- _diffrn_scan_axis.angle_range
- _diffrn_scan_axis.angle_increment
- _diffrn_scan_axis.angle_rstrt_incr
- _diffrn_scan_axis.displacement_start
- _diffrn_scan_axis.displacement_range
- _diffrn_scan_axis.displacement_increment
- _diffrn_scan_axis.displacement_rstrt_incr

(*c*) DIFFRN_SCAN_FRAME
- _diffrn_scan_frame.date
- • _diffrn_scan_frame.frame_id
  - → _diffrn_data_frame.id
- • _diffrn_scan_frame.scan_id
  - → _diffrn_scan.id
- _diffrn_scan_frame.frame_number
- _diffrn_scan_frame.integration_time

(*d*) DIFFRN_SCAN_FRAME_AXIS
- • _diffrn_scan_frame_axis.axis_id
  - → _axis.id
- • _diffrn_scan_frame_axis.frame_id
  - → _diffrn_data_frame.id
- _diffrn_scan_frame_axis.angle

```
_diffrn_scan_frame_axis.angle_increment
_diffrn_scan_frame_axis.angle_rstrt_incr
_diffrn_scan_frame_axis.displacement
_diffrn_scan_frame_axis.displacement_increment
_diffrn_scan_frame_axis.displacement_rstrt_incr
```

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

Data items in the DIFFRN_SCAN category describe the parameters of one or more scans, relating axis positions to frames. Each scan is uniquely identified by the value of `_diffrn_scan.id`. The data items in this category give overall information for the scan. The detailed frame-by-frame data are given in DIFFRN_SCAN_FRAME and DIFFRN_SCAN_FRAME_AXIS. The values of `_diffrn_scan.date_start` and `*.date_end` give the starting and ending time for a scan. The original definition of the *yyyy-mm-dd* data type, which includes date and time, has been extended in the CBF/imgCIF dictionary. This allows the seconds part of the time to include an optional decimal fraction. The approximate average integration time for each step of the scan is given by the value of `_diffrn_scan.integration_time`. The scan is tied to individual frame IDs by the values of `_diffrn_scan.frame_id_start` and `*.frame_id_end`. The number of frames in the scan is given by the value of `_diffrn_scan.frames`.

Data items in the DIFFRN_SCAN_AXIS category describe the settings of axes for particular scans. Unspecified axes are assumed to be at their zero points. The vector of each axis is not given here, because it is provided in the AXIS category. By making `_diffrn_scan_axis.scan_id` and `_diffrn_scan_axis.axis_id` keys of the DIFFRN_SCAN_AXIS category, an arbitrary number of scanning and fixed axes can be specified for a scan. The value of `_diffrn_scan_axis.scan_id` (a pointer to `_diffrn_scan.id`) identifies the scan and the values of `_diffrn_scan_axis.axis_id` (a pointer to `_axis.id`) associate particular axes with that scan. The steps of each axis are specified by `*_start`, `*_range`, `*_increment` and `*_rstrt_incr` values for angles or for displacements. The `*_start` value is the setting of the relevant axis at the start of the scan. The `*_range` value is the total change in the axis setting through the scan. The `*_increment` value is the increment in the axis setting for each step of the scan. The `*_rstrt_incr` value is the increment in the axis setting after each step of the scan.

Data items in the DIFFRN_SCAN_FRAME category describe the relationship of particular frames to scans. The value of `_diffrn_scan_frame.frame_id` (a pointer to `_diffrn_data_frame.id`) identifies the frame. The value of `_diffrn_scan_frame.scan_id` (a pointer to `_diffrn_scan.id`) identifies the scan of which the frame is a part. Together `_diffrn_scan_frame.frame_id` and `*.scan_id` form the category key. The value of `_diffrn_scan_frame.date` gives the date and time of the start of the data collection for the frame. The value of `_diffrn_scan_frame.frame_number` gives the number of the frame (starting with 1). The value of `_diffrn_scan_frame.integration_time` gives the precise time in seconds to integrate this step of the scan.

Table A3.7.1.1. *Categories in the CBF/imgCIF dictionary*

Numbers in parentheses refer to the section of this chapter in which each category is described in detail.

ARRAY_DATA group (§3.7.2)	DIFFRN_DETECTOR_AXIS (§3.7.4.2(*b*))
ARRAY_DATA (§3.7.2.1)	DIFFRN_DETECTOR_ELEMENT (§3.7.4.2(*c*))
ARRAY_ELEMENT_SIZE (§3.7.2.2)	DIFFRN_MEASUREMENT (§3.7.4.3(*a*))
ARRAY_INTENSITIES (§3.7.2.3)	DIFFRN_MEASUREMENT_AXIS (§3.7.4.3(*b*))
ARRAY_STRUCTURE (§3.7.2.4(*a*))	DIFFRN_RADIATION (§3.7.4.4)
ARRAY_STRUCTURE_LIST (§3.7.2.4(*b*))	DIFFRN_REFLN (§3.7.4.5)
ARRAY_STRUCTURE_LIST_AXIS (§3.7.2.4(*c*))	DIFFRN_SCAN (§3.7.4.6(*a*))
AXIS group (§3.7.3)	DIFFRN_SCAN_AXIS (§3.7.4.6(*b*))
AXIS (§3.7.3)	DIFFRN_SCAN_FRAME (§3.7.4.6(*c*))
DIFFRN group (§3.7.4)	DIFFRN_SCAN_FRAME_AXIS (§3.7.4.6(*d*))
DIFFRN_DATA_FRAME (§3.7.4.1)	
DIFFRN_DETECTOR (§3.7.4.2(*a*))	

Data items in the DIFFRN_SCAN_FRAME_AXIS category describe the settings of axes for particular frames. Unspecified axes are assumed to be at their zero points. If for any given frame non-zero values apply for any of the data items in this category, those values should be given explicitly in this category and not simply inferred from values in DIFFRN_SCAN_AXIS. Since the collection for a given frame may involve multiple axes, the frame involved is identified by the value of `_diffrn_scan_frame_axis.frame_id` (a pointer to `_diffrn_data_frame.id`) and each axis is identified by the value of `_diffrn_scan_frame_axis.axis_id` (a pointer to `_axis.id`). Together `_diffrn_scan_frame_axis.frame_id` and `*.axis_id` form the category key. If the axis is an axis of rotation, the axis settings for the frame are given by the values of `_diffrn_scan_frame_axis.angle`, `*.angle_increment` and `*.angle_rstrt_incr`. If the axis is a translation axis, the axis settings for the frame are given by the values of `_diffrn_scan_frame_axis.displacement`, `*.displacement_increment` and `*.displacement_rstrt_incr`. The integration begins at the setting given by the value of `_diffrn_scan_frame_axis.angle` or of `*.displacement`. The `*_increment` value gives the change of axis setting during the scan. At the end of the integration, the axis may need to be repositioned by an additional amount. That amount is given by `*_rstrt_incr`.

Appendix 3.7.1
Category structure of the CBF/imgCIF dictionary

Table A3.7.1.1 provides an overview of the structure of the CBF/imgCIF dictionary by category group and member categories.

We are grateful to Frances C. Bernstein, Paula Fitzgerald and Bob Sweet for their helpful comments and suggestions.

References

Aho, A. V., Hopcroft, J. E. & Ullman, J. D. (1987). *Data structures and algorithms*. Reading, MA: Addison–Wesley.

Leslie, A. G. W. & Powell, H. (2003). *MOSFLM V6.11*. http://www.ccp4.ac.uk/dist/x-windows/Mosflm.