

Version 1.0

The PRISM Source Vocabulary PAM to PSV Guide

September 27, 2012



Copyright and Legal Notices

© 2001 – 2012 International Digital Enterprise Alliance, Inc. All Rights Reserved.

PRISM® and nextPub® are registered trademarks of the International Digital Enterprise Alliance, Inc. (IDEAlliance).

This document may be downloaded and copied provided that the above copyright notice and this Notice are included on all such copies. This document itself may not be modified in any way, except as needed for the purpose of developing International Digital Enterprise Alliance, Inc. ("IDEAlliance") specifications. Use of the specification or standard set forth in this document shall not create for the user any rights in or to such specification or standard or this document, which rights are exclusively reserved to IDEAlliance or its licensors or contributors.

Use of this document and any specification or standard contained herein is voluntary. By making use of this document or any specification or standard contained herein, the user assumes all risks and waives all claims against IDEAlliance, its licensors and contributors. By making this document available, IDEAlliance is not providing any professional services or advice to any person or entity. Any person or entity utilizing this document or any specification or standard contained herein should rely upon the advice of a competent professional before using any such information.

NO WARRANTY, EXPRESSED OR IMPLIED, IS MADE REGARDING THE ACCURACY, ADEQUACY, COMPLETENESS, LEGALITY, RELIABILITY OR USEFULNESS OF ANY INFORMATION CONTAINED IN THIS DOCUMENT OR IN ANY SPECIFICATION OR STANDARD OR OTHER PRODUCT MADE AVAILABLE BY IDEALLIANCE. THIS DOCUMENT AND THE INFORMATION CONTAINED HEREIN AND INCLUDED IN ANY SPECIFICATION OR STANDARD OR OTHER PRODUCT OR SERVICE OF IDEALLIANCE IS PROVIDED ON AN "AS IS" BASIS. IDEALLIANCE DISCLAIMS ALL WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, ANY ACTUAL OR ASSERTED WARRANTY OF NON-INFRINGEMENT OF PROPRIETARY RIGHTS, MERCHANTABILITY, OR FITNESS FOR A PARTICULAR PURPOSE.

IN NO EVENT SHALL IDEALLIANCE, ITS LICENSEES, CONTRIBUTORS OR THEIR RESPECTIVE OFFICERS, DIRECTORS, EMPLOYEES, AGENTS, REPRESENTATIVES, SUPPLIERS OR CONTENT OR SERVICE PROVIDERS BE LIABLE FOR DAMAGES OF ANY KIND, INCLUDING WITHOUT LIMITATION, DIRECT, INDIRECT, COMPENSATORY, SPECIAL, INCIDENTAL, PUNITIVE OR CONSEQUENTIAL DAMAGES (INCLUDING WITHOUT LIMITATION DAMAGES FROM DATA LOSS OR BUSINESS INTERRUPTION) EVEN IF MADE AWARE OF THE POSSIBILITY OF SUCH DAMAGES, WHETHER IN AN ACTION UNDER CONTRACT, TORT OR ANY OTHER THEORY, ARISING OUT OF OR IN CONNECTION WITH THE USE, INABILITY TO USE OR PERFORMANCE OF THIS DOCUMENT, THE SPECIFICATION OR STANDARD CONTAINED HEREIN, OR ANY OTHER DOCUMENT OR SPECIFICATION OR STANDARD MADE AVAILABLE BY IDEALLIANCE.

Some states do not allow the disclaimer or limitation of damages, so the disclaimers set forth above apply to the maximum extent permitted under applicable law.

IDEAlliance takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed or implicated with respect to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available. IDEAlliance does not represent that it has made any effort to identify any such rights. Information on IDEAlliance's procedures with respect to rights in IDEAlliance specifications can be found at the IDEAlliance website at www.idealliance.org. Copies of third-party claims of rights, assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification, can be obtained from the President of IDEAlliance at patent-disclosure@idealliance.org.

IDEAlliance requests interested parties to disclose any copyrights, trademarks, service marks, patents, patent applications, or other proprietary or intellectual property rights which may cover technology that may be required to implement this specification. Please address the information to the President of IDEAlliance at patent-disclosure@idealliance.org.

Table of Contents

1 Int	roduction	
1.1	Purpose and Scope	1
1.2	Relationship of PSV to PRISM	1
1.3	Relationship of nextPub to PAM	1
1.4	The PSV Namespace	2
1.5	Diagrams	2
1.6	The PSV Documentation Package	3
1.7	The PRISM Documentation Package	4
1.7	7.1 General Documents	Error! Bookmark not defined.
1.7	7.2 PRISM Metadata Specifications	Error! Bookmark not defined.
1.7	7.3 PRISM Controlled Vocabulary Specifications	Error! Bookmark not defined.
1.7	7.4 PRISM Aggregator Message Markup Specification	Error! Bookmark not defined.
1.7	7.5 PRISM Inline Markup Specification	Error! Bookmark not defined.
1.8	PSV Content Management Schema	4
1.9	Other PSV Schemas	7
2 Ta	gging Overview	9
2.1	Unit of Content Tagging	9
2.1	.1 Content Types	9
2.1	.2 Genres	9
2.2	HTML5 PSV vs. XHTML PAM	10
2.2	2.1 Elimination of pam: and pim: Elements	10
2.2	2.2 Handling the <pim:quote element<="" th=""><th>11</th></pim:quote>	11
2.2	2.3 Addition of New HTML5 Elements	11
2.3	Comparison of High Level Structures	11
2.3	3.1 PAM Structure	11
2.3	3.2 PSV Structure	12
3 Ta	gging Article Metadata	13
3.1	Comparison of PSV Metadata to PAM Metadata	13

3.2	New	PSV Metadata Fields	14
3.2.1 Content Type Field			14
3.2	2.2 Un	ique ID Field	14
3.2	2.3 Re	lations Metadata Fields	15
3.2	2.4 Us	age Rights Metadata Fields	15
3.2	2.5 Co	mponent Metadata Fields	16
3.2	2.6 Wł	nere Used Metadata Fields	17
3	3.2.6.1	Aggregation Information Metadata	17
3	3.2.6.2	Publication Information	18
3	3.2.6.3	Issue Information	18
3	3.2.6.4	Article Information	19
3	3.2.6.5	Website Information	20
3	3.2.6.6	Book Information	20
3	3.2.6.7	Chapter Information	21
3	3.2.6.8	Blog Information	21
3	3.2.6.9	Blog Entry Information	21
3.2	2.7 PA	M to PSV Mapping Table	22
3.2	2.8 Exa	ample Metadata Tagging	24
4 Ta	gging	Content	27
4.1	Wha	t Does this Mean?	27
4.2	Samp	ole Coding for Body Content	27
5 Ta	gging	Media	29
5.1	Abou	ut PAM Media	29
5.2	Abou	ut nextPub Media	29
5.3	Samı	ole Coding	30
5.4	Abou	ut Tagging Media Metadata in nextPub	31
Арре	endix A	PAM Sample Tagging	35
Арре	endix B	Sample PSV Tagging	41

1 Introduction

1.1 Purpose and Scope

The IDEAlliance PSV Specification defines an encoding for XML elements and attributes in the psv: namespace to encode semantically rich source content. Metadata fields and values used in this specification are drawn from the IDEAlliance PRISM Metadata and Controlled Vocabulary Specifications. This Guide provides insights for those currently using the PRISM Aggregator Message about how PAM content can be expressed in PSV XML markup and about how these two markup schemas are different.

1.2 Relationship of PSV to PRISM

PSV builds upon the foundation of PRISM Specifications. PSV defines XML structures for tagging source content, but it *does not* define its own metadata fields or controlled vocabularies. nextPub is built upon PRISM and relies on the metadata fields and controlled vocabularies defined by PRISM. Of necessity, PSV will require an update to the current version of PRISM because new metadata fields and controlled vocabularies and terms must be added to support PSV functionality. Therefore PRISM 3.0 will be published simultaneously with the publication of nextPub 1.0 and will be highly referenced by the nextPub XML Source Specification. See Figure 1.1.

1.3 Relationship of nextPub to PAM

PAM is the PRISM Aggregator Message. PAM is an XML tag set built on the foundation of PRISM metadata and controlled vocabularies. The use case for PAM was originally to encode magazine articles in XML to deliver content to aggregators. While some publishers currently use PAM XML as a content source, that was not the original intent. PAM is an application of PRISM, but PAM and PRISM are not synonymous. PAM is an XML tag set that uses PRISM metadata for a very specific purpose while PRISM remains the core specification for metadata and controlled vocabularies. See Figure 1.1.



Figure 1.1 Relationship of PRISM to PSV and PAM

1.4 The PSV Namespace

Dublin Core, PRISM, PRISM Usage Rights, PRISM Recipe Metadata and other relevant PRISM metadata namespaces along with the new psv: namespace will be utilized as appropriate. XML Structures unique to this specification for content encoding will be given the namespace nextPub:

The recommended namespace for nextPub markup is: xmlns:psv ="http://prismstandard.org/namespaces/psv/1.0/"

1.5 Diagrams

In this Specification, the PSV XML model and the PAM XML models are illustrated using model diagrams produced with the XML Spy product. These diagrams show the element and attribute structures.

The legend for reading XML model diagrams is shown in Figure 1.2. Elements that are required by the model are shown in a solid box. Elements that are optional are shown in a dotted box. Likewise attributes may be required (solid box) or optional (dotted box). A repeatable occurrence of elements is indicated by numbers below each element box to the right.

The diagrams also indicate how elements are assembled. When building some models, elements may occur in a sequence with a specified order. Other models provide a choice from among a number of elements. The legend in Figure 1.2 shows the connectors for sequence and choice.

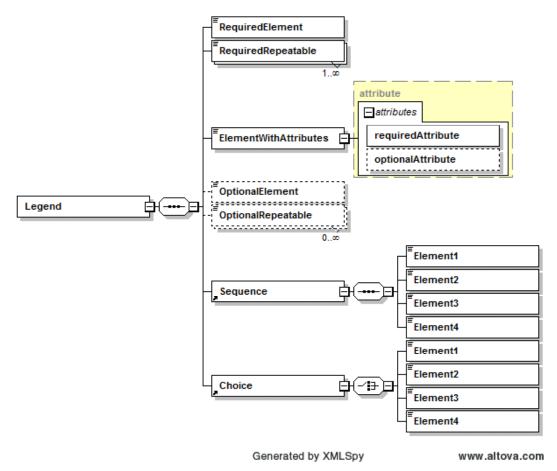


Figure 1.2 Legend for XML Diagrams

Note: There is no official schema for HTML5. However, nextPub has developed a schema for the PRISM Source Vocabulary Specification. This schema is specifically designed to be more restrictive than a true HTML5 schema would be in order to make source content encoding and transformations to delivery channel formats much more straightforward. The design goal is to define a valid HTML5 subset for the content encoding portion of nextPub.

1.6 The PSV Documentation Package

The nextPub Working Group has developed a series of specifications collectively known as the PRISM Source Vocabulary. The use case for PSV is to encode semantically rich content that can be transformed and delivered to any platform in that platform-native format. This Specification is made up of a modular documentation package that builds on PRISM 3.0 and HTML5. Over time new modules may be added to the documentation package. The documentation package for the nextPub PRISM Source Vocabulary Specification Version 1.0 consists of:

Document	Description
PRISM Source Vocabulary Specification Overview	The Introduction to the PRISM Source
[PSVSO]	Vocabulary provides an introduction
http://www.prismstandard.org/specifications/psv/	and a non-technical overview of the
1.0/PSV_overview.pdf	PRISM Source Vocabulary.
or	
http://www.prismstandard.org/specifications/psv/	
1.0/PSV_overview.htm	
PRISM Source Vocabulary Specification [PSVS]	The PRISM Source Vocabulary
http://www.prismstandard.org/specifications/psv/	Specification defines semantically rich
1.0/PSV.pdf	for source metadata and content
or	markup that can be transformed and
http://www.prismstandard.org/specifications/psv/	served to a wide variety of output
1.0/PSV.htm	devices including eReaders, mobile
	tablet devices, smart phones and print.
PRISM Source Vocabulary Markup Specification	The PSV Markup Specification
[PSVMS]	documents the XML tags in the PSV
http://www.prismstandard.org/specifications/psv/	namespace that are used to encode
1.0/PSV_markkup.pdf	XML Source Content.
or	
http://www.prismstandard.org/specifications/psv/	
1.0/PSV_markup.htm	
PAM to PSV_Guide [PAMPSVGUIDE]	This Guide documents mappings from
http://www.prismstandard.org/specifications/psv/	PAM XML to PSV XML. It is normative
1.0/PAM_PSV.pdf	only.
or	
http://www.prismstandard.org/specifications/psv/	
1.0/PAM_PSV.htm	

1.7 The PRISM Documentation Package

Because PSV is built on PRISM 3.0, there is a close relationship between the two specifications. In fact, access to the PRISM 3.0 Documentation Package is critical to the implementation of PSV. The PRISM 3.0 Documentation Package consists of:

1.8 The PRISM 3.0 Documentation Package

The PRISM Documentation Package consists of:

Because PSV is built on PRISM 3.0, there is a close relationship between the two specifications. In fact, access to the PRISM 3.0 Documentation Package is critical to the implementation of PSV. The PRISM 3.0 Documentation Package consists of:

1.8.1 General Documents

This is a set of general or overview documents that apply to PRISM.

Document	Description
PRISM Introduction [PRISMINT]	Overview, background, purpose and

Document	Description
http://www.prismstandard.org/specifications/3.0/ PRISM_introduction_3.0.pdf	scope of PRISM; examples; contains no normative material.
or	
http://www.prismstandard.org/specifications/3.0/	
PRISM_introduction_3.0.htm	
PRISM Compliance [PRISMCOMP]	Describes three profiles of PRISM
http://www.prismstandard.org/specifications/3.0/	compliance for content and systems;
PRISM_compliance_3.0.pdf	includes normative material.
or	
http://www.prismstandard.org/specifications/3.0/	
PRISM_compliance_3.0.htm	

1.8.2 PRISM Metadata Specifications

This is the set of documents that outline the prism metadata fields and values by PRISM metadata category. PRISM has modularized its metadata specification by namepace so users may pick those modules that meet their unique business requirements without having to implement the entire PRISM specification.

Document	Description
The PRISM Basic Metadata Specification [PRISMBMS]	Describes the basic metadata elements
http://www.prismstandard.org/specifications/3.0/	contained in the PRISM namespace to
PRISM_Basic_Metadata_3.0.pdf	describe article content; includes
or	normative material.
http://www.prismstandard.org/specifications/3.0/	
PRISM Basic Metadata 3.0.htm	
PRISM Advertising Metadata Specification	Describes advertising metadata
[PRISMADMS]	elements including those drawn from
http://www.prismstandard.org/specifications/3.0/	AdsML, GWG and Ad-ID; includes
PRISM_Advertising_Metadata_3.0.pdf	normative material.
or	
http://www.prismstandard.org/specifications/3.0/	
PRISM_Advertising_Metadata_3.0.htm The PRISM Subset of Dublin Core Metadata	Describes the constant data also such forms
Specification [PRISMDCMS]	Describes the metadata elements from
http://www.prismstandard.org/specifications/3.0/	the Dublin Core namespace that are
PRISM_Dublin_Core_Metadata_3.0.pdf	included in PRISM; includes normative material.
or	material.
http://www.prismstandard.org/specifications/3.0/	
PRISM_Dublin_Core_Metadata_3.0.htm	
The PRISM Image Metadata Specification [PRISMIMS]	Describes the metadata elements
http://www.prismstandard.org/specifications/3.0/	contained in the PRISM Metadata for
PRISM_Image_Metadata_Specification_3.0.pdf	Images Namespace and other related
or	image namespaces, includes normative
http://www.prismstandard.org/specifications/3.0/	material.
PRISM_Image_Metadata_Specification_3.0.htm	

Document	Description
The PRISM Recipe Metadata Specification [PRISMRMS] http://www.prismstandard.org/specifications/3.0/ PRISM_Recipe_Metadata_3.0.pdf or http://www.prismstandard.org/specifications/3.0/ PRISM_Recipe_Metadata_3.0.htm	Describes the metadata elements contained in the PRISM Recipe Metadata Namespace, includes normative material
The PRISM Usage Rights Metadata Specification [PRISMURMS] http://www.prismstandard.org/specifications/3.0/ PRISM_Usage_Rights_Metadata_3.0.pdf	Describes the metadata elements contained in the PRISM Usage Rights Namespace; includes normative material. This namespace will supersede
or http://www.prismstandard.org/specifications/3.0/ PRISM Usage Rights Metadata 3.0.htm	elements in both the prism: and prl: namespaces in version 3.0 of the specification.

1.8.3 PRISM Aggregator Message Markup Specification

This module documents the PRISM Markup Elements and Attributes for use with the PRISM Aggregator Message. At the time of the publication of the Introduction to PRISM, the PAM Message remains at version 2.1. This set of documents includes:

Document	Description
The PRISM PAM Markup Specification [PRISMPAMMS] http://www.prismstandard.org/specifications/2.1/ PRISM_PAM_Markup_2.1.pdf	Describes the XML elements and attributes used to encode the PRISM Aggregator Message from both the pam: and pim: namespaces; includes
or http://www.prismstandard.org/specifications/2.1/ PRISM_PAM_Markup_2.1.htm	normative material.

1.8.4 PRISM Inline Markup Specification

This module documents the PRISM Inline Markup Elements and Attributes for use with the PRISM Aggregator Message. This set of documents includes:

Document	Description
The PRISM Inline Markup Specification [PRISMIMS]	Describes the XML elements used to
http://www.prismstandard.org/specifications/2.1/	encode the inline markup for the PRISM
PRISM_PIM_Markup_Specification 3.0.pdf	Aggregator Message. Includes
or	normative material.
http://www.prismstandard.org/specifications/2.1/	
PRISM PIM Markup Specification 3.0.htm	

1.8.5 PRISM Controlled Vocabulary Specifications

These modules are new with PRISM 3.0. All controlled vocabularies and their terms are documented in this publication set.

Document	Description
The PRISM Controlled Vocabulary Markup	Describes the metadata fields in the
Specification [PRISMCVMS]	PRISM Controlled Vocabulary
http://www.prismstandard.org/specifications/3.0/	Namespace that can be used to describe
PRISM_Controlled_Vocabulary_Markup_3.0.pdf	a controlled vocabulary. Actual PRISM
or	controlled vocabularies are now placed
http://www.prismstandard.org/specifications/3.0/	in the PRISM Controlled Vocabularies
PRISM Controlled Vocabulary Markup 3.0.htm	Specification [PRISMCVS]
The PRISM Controlled Vocabularies Specification	The PRISM Controlled Vocabularies are
[PRISMCVS]	now documented in this document.
http://www.prismstandard.org/specifications/3.0/	
PRISM_CV_Spec_3.0.pdf	
or	
http://www.prismstandard.org/specifications/3.0/	
PRISM CV Spec 3.0.htm	

1.9 PSV Content Management Schema

In order to assist implementers develop a PSV-based federated content management solution, the nextPub Working Group is providing an XML Schema (XSD) that can serve as the basis for the design of a PSV content repository.

Note: The PSV CM schema is not designed for tagging content. It is provided simply to serve as a basis for the design of a content repository. Metadata building blocks from this schema can be combined with HTML5 by publishers who wish to develop a hybrid PSV metadata and content tagging schema.

1.10 Other PSV Schemas

Because PSV is a flexible framework, it supports many different use case scenarios. A different schema, using the PSV metadata fields and content encoding can be developed for each different use case. In order to assist PSV implementers, the nextPub Working Group is planning to provide a number of XML Schemas (XSDs) to support common use cases including tagging an article and transmitting articles to content aggregators. These PSV sample schemas will be available from the nextPub website (http://www.nextpub.org) and documented in the nextPub PSV Implementation Guide that will be published following the publication of this specification.

2 TAGGING OVERVIEW

This section provides an overview of source encoding for PAM and PSV. The philosophy for moving from the existing PRISM Aggregator Message will be discussed. Then a high-level overview of the source structure will be highlighted.

2.1 Unit of Content Tagging

When defining PSV source content tagging, a new concept, the content type was developed. In fact PSV XML tagging requires that the unit of storage, or prism:contentType must be specified. In addition the prism:genre has been updated so that prism:genre is used to refine the prism:contentType. For the purposes of mapping PAM to PSV a direct mapping with the article as the default unit of content is assumed.

Note: The use cases for nextPub have expanded the scope of PRISM metadata fields and controlled vocabularies beyond the scope of magazines, newsletters and journals. PAM mapping will only consider those fields and controlled vocabulary terms that relate to the use case for PAM and that is magazines.

2.1.1 Content Types

The use cases for nextPub have expanded the scope of PRISM beyond the article. This, in turn, has introduced the concept of a prism:contentType to the PRISM Specification. Content Type has not been included in PAM for the purposes of marking up magazine content for delivery to aggregators as the assumption is that PAM will remain as a method for delivering article content to aggregators and will not extend to other content types.

2.1.2 Genres

The PRISM Genre Controlled Vocabulary has been enhanced to refine the intellectual description of core content units in PRISM 3.0 and for PSV. Because PAM assumes only the delivery of article content, only the genre that can refine the description of an article need be considered.

Again, since the scope of PSV is broader than the scope for PAM, only a limited number of the new PRISM 3.0 genre terms make sense for use with PAM markup. Those fields include:

- abstract (refines article)
- adaptation (refines article)
- analysis (refines article)
- autobiography (refines article)
- biography (refines article)
- calendar (refines article)
- chronology (refines article)
- clarification (refines article)
- column (refines article)
- correction (refines article)
- coverStory (refines article)
- coverPackageArticle (refines article)
- department (refines article)

- electionResults (refines article)
- essay (refines article)
- excerpt (refines article)
- fashionShoot (refines article)
- faq (refines article)
- feature (refines article)
- featurePackageArticle (refines article)
- fiction (refines article)
- financialStatement (refines article)
- interview (refines article)
- letters (refines article)
- list (refines article)
- newsBulletin (refines article)
- notice (refines article)
- obituary (refines article)
- opinion (refines article)
- photoEssay (refines article)
- profile (refines article)
- qAndA (refines article)
- quotation (refines article)
- ranking (refines article)
- reprint (refines article)
- resources (refines article)
- response (refines article)
- review (refines article)
- schedule (refines article)
- supplementArticle (refines article)
- transcript (refines article)
- userComments (refines article)
- userReview (refines article)
- wireStory (refines article)

2.2 HTML5 PSV vs. XHTML PAM

When work began on the nextPub XML Source Specification, the majority of participants favored retaining PRISM/PAM XHTML as the source format and extending that source to take into account functionality required by the new use cases for content delivered to tablets, eReaders, smart phones and beyond. However, as work progressed and transforms into the two key target delivery formats were considered, it became apparent that nextPub needed to be based on HTML5 rather than XHTML as PAM is. Benefits of basing nextPub on HTML5 included the ability to easily transform into key delivery formats for new digital devices and the enhanced ability to handle rich media.

So what does this mean?

2.2.1 Elimination of pam: and pim: Elements

When PAM tagging was designed, the PRISM Working Group extended XHTML with new XML elements in the pam: namespace.

For PSV we made the decision that we will NOT extend the HTML5 tag set that native HTML5 browsers and processors are expecting. Hence this leads to some changes in how PSV article content is tagged when compared to PAM tagging. And it means we no longer have use of <pam: or <pre><ppm: elements</pre>.

2.2.2 Handling the <pim:quote Element

HTML5 has a quote (<q) element that is used to mark quotations. This should be used to represent <pim:quote>. Note that browsers often put quotation marks in the display when <q is used.

2.2.3 Addition of New HTML5 Elements

When designing PSV tagging schema, the decision was made to leverage several new HTML5 tags. These new tags include:

- <article: independent, self-contained content. Used as the root element for an article. Replaces <pam:article
- <aside: defines some content aside from the content it is placed in. Can be used for coding a sidebar or box
- <figure and <figcaption: self-contained content, like illustrations, diagrams, photos, code listings, etc. Replaces <pam:media</p>
- <nav: containing element for a set of navigation links, either to other pages, or fragment identifiers in the current page.

2.3 Comparison of High Level Structures

PAM is all about tagging article level content to send to content aggregators. nextPub is about tagging content to store in a content management system for assembly into publication products.

2.3.1 PAM Structure

PAM tagging supports sending a message containing one or more articles. Articles allow metadata entry in the <head and the content is coded with XHTML-based body tags. See Figure 2.1.

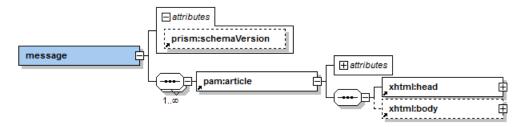


Figure 2.1 PAM Structure

2.3.2 PSV Structure

PSV has a different XML model than does PAM. Since PSV is designed as a storage format for source content, it has a very well-organized and robust metadata block followed by a block for content which is a parsable subset of HTML5.

Note: Although there is an optional <meta tag in the HTML5 <head structure, it is not to be used to store metadata about the article. The recommended PSV HTML5 subset definition for the <head only allows structures such as <link and <styles but does not allow for the encoding of metadata. Metadata is expected to be consolidated in the PSV<metadata block.

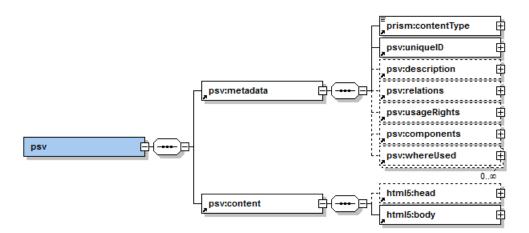


Figure 2.2 PSV Structure

3 TAGGING ARTICLE METADATA

This section provides an overview of metadata tagging for PAM and PSV. Differences between the two approaches to tagging metadata will be discussed. Elements currently in PAM that are not used in PSV will be discussed.

3.1 Comparison of PSV Metadata to PAM Metadata

Because PSV metadata is built on PRISM but <u>not</u> directly on PAM and because PSV has a completely different use case, there are differences in the metadata blocks describing an article. The major differences between PAM and PSV metadata encoding are:

- In PAM, all metadata is included in the XHTML < head element. In PSV, metadata is included in a new PSV < metadata block.
- In PAM, all the metadata is included at the same level in a prescribed order. In PSV we have partitioned the metadata to provide for easier access and evaluation.
- In PAM, metadata identifies the article by the publication where it appears. In PSV we list all instances of use by platform and device and eliminate the idea of "origin platform."

See Figure 3.1 to view the PAM metadata structure.

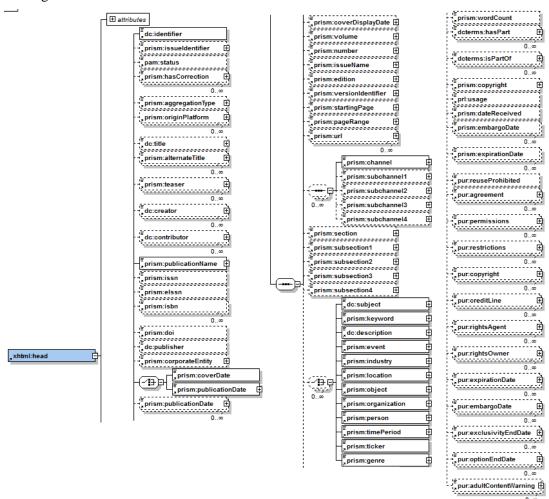


Figure 3.1 PAM Metadata

The nextPub use cases rely on the metadata in the <psy:metadata block of the article to enable content management and the automated aggregation units of content for any purpose, from creating an issue of a magazine to creating a bookazine or even to new collection types in the future. When nextPub moved away from the PRISM/PAM XML model, the Working Group developed high-level organizational structures to provide for easier query, evaluation and access. See Figure 3.2.

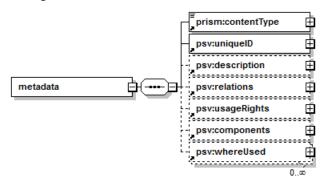


Figure 3.2 PSV Metadata

3.2 New PSV Metadata Fields

PSVadds prism:contentType to indicate the of the nature of the asset that is tagged and stored. This is being added to PAM 2.2 as well. See Section 2.1.1 and 2.1.2 for information about the content types and genre that apply to PAM tagging for magazine content.

In addition, PSV adds new categories of metadata.

3.2.1 Content Type Field

PSV has many publication types and content units within its scope. The same is not true for PAM. To map PAM to PSV the prism:contentType field should be set to the value "article".

3.2.2 Unique ID Field

In PAM 2.2, a unique ID in the form of dc:identifier is required. This maps directly into the psv:uniqueID block of PSV. See Figure 3.3.

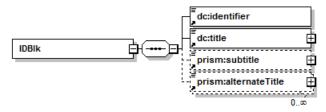


Figure 3.3 ID Block Structure

3.2.3 Relations Metadata Fields

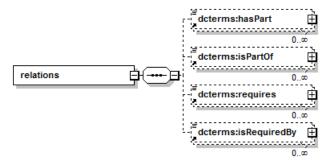


Figure 3.4 Relations Metadata Structure

3.2.4 Usage Rights Metadata Fields



Figure 3.5 Usage Rights Metadata

3.2.5 Component Metadata Fields

New metadata found in PSV but not in PRISM is component metadata. In PAM, metadata can be coded in the body of the article using <pim: markup or within the pam:media structure. In PSV, the content is compliant with HTML5 and no metadata may be included directly in the body. Hence the metadata is coded in the "component" metadata block and linked to the structure in the body by matching the id= in the body to a refines= attribute in the components block. Note that the refines= syntax matches that used in EPUB 3 where refines= is a "URI". This means that the value is preceded by a # sign to indicate the path for the link being specified. In this way PSV metadata is much more sophisticated than metadata coded with PAM markup.

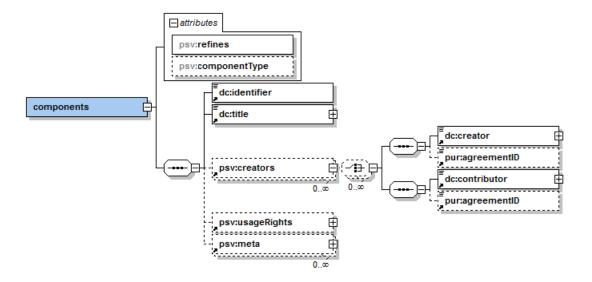


Figure 3.6 Component Metadata

3.2.6 Where Used Metadata Fields

Because PSV is about tagging content to store in a content management system for assembly into publication products, the prism:originPlatform does not make sense for inclusion in the PSV tag set. In its place a metadata block, <psv:whereUsed, has been developed. Metadata in this block tracks usage of content over time. See Figure 3.7. Note that some PRISM metadata associated with an article in PAM is now included in the PSV "where used" block.

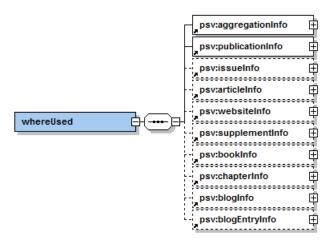


Figure 3.7 Where Used Metadata

3.2.6.1 Aggregation Information Metadata

The prism:aggregationType field is included in PAM and in PSV. It should be included in the psv:aggregationInfo block. If prism:platform is available, it will be mapped into this block as well. See Figure 3.7.

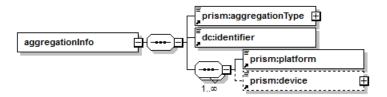


Figure 3.7 Aggregation Info Structure

3.2.6.2 Publication Information

PAM includes numerous fields that specify information about the publication in the <head element. These all should be mapped into the psv:publicationInfo block. See Figure 3.8.

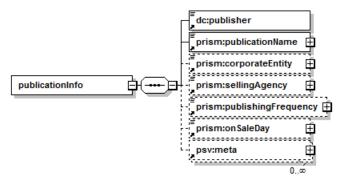


Figure 3.8 Publication Info Structure

3.2.6.3 Issue Information

PAM includes numerous fields that specify information about the issue in the <head element. These all should be mapped into the psv:issueInfo block. See Figure 3.9.

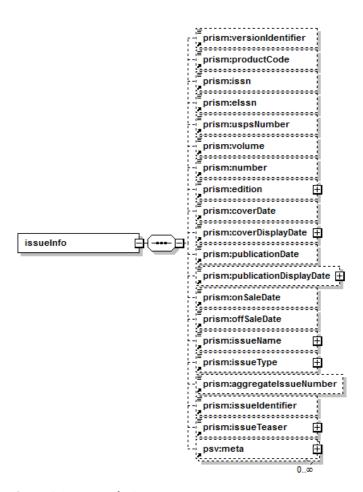


Figure 3.9 Issue Info Structure

3.2.6.4 Article Information

PAM includes numerous fields that specify information about the article in the <head element. These all should be mapped into the psv:articleInfo block. See Figure 3.10.

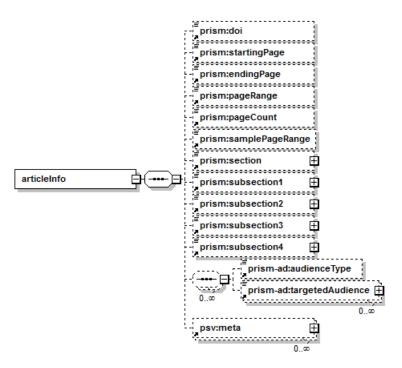


Figure 3.10 Article Info Structure

3.2.6.5 Website Information

PAM includes numerous fields that specify information about a website in the <head element. These all should be mapped into the psv:websiteInfo block. See Figure 3.11.

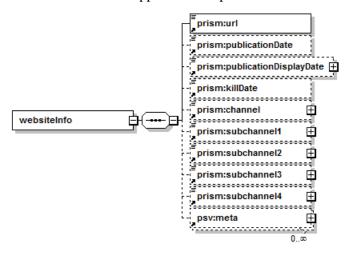


Figure 3.11 Website Info Structure

3.2.6.6 Book Information

PAM includes book ISBN in the <head element. This should be mapped into the psv:bookInfo block. See Figure 3.12.

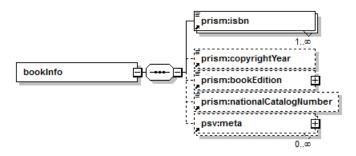


Figure 3.12 Book Info Structure

3.2.6.7 Chapter Information

PAM does not contain chapter information so no data should be mapped from PAM to PSV for this block. See Figure 3.13.

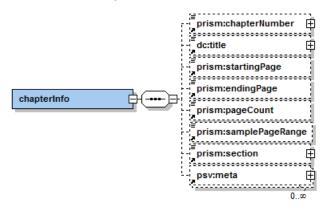


Figure 3.13 Chapter Info Structure

3.2.6.8 Blog Information

PAM does not contain blog information so no data should be mapped from PAM into PSV for this block. See Figure 3.14.

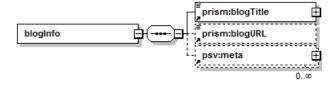


Figure 3.14 Blog Info Structure

3.2.6.9 Blog Entry Information

PAM does not contain blog entry information so no data should be mapped from PAM into PSV for this block. See Figure 3.15.

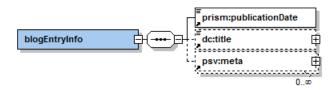


Figure 3.15 Blog Entry Info Structure

3.2.7 PAM to PSV Mapping Table

The following table contains all possible elements found in the PAM 2.2 <head and their mappings into PSV metadata blocks.

PAM Heading Element	PSV Metadata Mapping
dc:identifier	psv:uniqueID > psv:IDBlk
dc:contributor	psv:description
dc:creator	psv:description
dc:description	psv:description
dc:publisher	psv:whereUsed > psv:publicationInfo
dc:subject	psv:description
dc:title	psv:uniqueID > psv:IDBIk
dcterms:hasPart	psv:relations
dcterms:isPartOf	psv:relations
pam:status	NO Mapping
prism:academicField	psv:description
prism: aggregatels sue Number	psv:whereUsed > psv:issueInfo
prism:aggregationType	psv:whereUsed > psv:aggregationInfo
prism:alternateTitle	psv:uniqueID > psv:IDBIk
prism:channel	psv:whereUsed > psv:websiteInfo
prism:channel1	psv:whereUsed > psv:websiteInfo
prism:channel2	psv:whereUsed > psv:websiteInfo
prism:channel3	psv:whereUsed > psv:websiteInfo
prism:channel4	psv:whereUsed > psv:websiteInfo
prism:copyright	psv:usageRights = pur:copyright
prism:corporateEntity	psv:whereUsed > psv:publicationInfo
prism:coverDate	psv:whereUsed > psv:issueInfo
prism:coverDisplayDate	psv:whereUsed > psv:issueInfo
prism:dateReceived	NO Mapping
prism:doi	psv:whereUsed > psv:articleInfo
prism:edition	psv:whereUsed > psv:issueInfo
prism:elssn	psv:whereUsed > psv:issueInfo
prism:embargoDate	psv:usageRights = pur:embargoDate
prism:event	psv:description
prism:expirationDate	psv:usageRights = pur:expirationDate
prism:genre	psv:description
prism:hasCorrection	psv:relations
prism:industry	psv:description
prism:isbn	psv:whereUsed > psv:bookInfo
prism:issn	psv:whereUsed > psv:issueInfo

PAM Heading Element	PSV Metadata Mapping	
prism:issueldentifier	psv:whereUsed > psv:issueInfo	
prism:issueName	psv:whereUsed > psv:issueInfo	
prism:issueTeaser	psv:whereUsed > psv:issueInfo	
prism:issueType	psv:whereUsed > psv:issueInfo	
prism:keyword	psv:description	
prism:link	psv:description	
prism:location	psv:description	
prism:number	psv:whereUsed > psv:issueInfo	
prism:object	psv:description	
prism:organization	psv:description	
prism:originPlatform	No Mapping	
prism:pageRange	psv:whereUsed > psv:articleInfo	
prism:person	psv:description	
prism:profession	psv:description	
prism:publicationDate	psv:whereUsed >:issueInfo	
prism:publicationName	psv:whereUsed > psv:publicationInfo	
prism:section	psv:whereUsed > psv:articleInfo	
prism:seriesNumber	psv:whereUsed > psv:issueInfo	
prism:seriesTitle	psv:whereUsed > psv:issueInfo	
prism:sport	psv:description	
prism:startingPage	psv:whereUsed > psv:articleInfo	
prism:subsection1	psv:whereUsed > psv:articleInfo	
prism:subsection2	psv:whereUsed > psv:articleInfo	
prism:subsection3	psv:whereUsed > psv:articleInfo	
prism:subsection4	psv:whereUsed > psv:articleInfo	
prism:subtitle	psv:uniqueID > psv:IDBIk	
prism: supplement Display ID	psv:whereUsed > psv:articleInfo	
prism: supplement Starting Page	psv:whereUsed > psv:articleInfo	
prism:supplementTitle	psv:whereUsed > psv:articleInfo	
prism:teaser	psv:description	
prism:ticker	psv:description	
prism:timePeriod	psv:description	
prism:url	psv:whereUsed > psv:websiteInfo	
prism:uspsNumber	psv:whereUsed > psv:issueInfo	
prism:versionIdentifier	psv:whereUsed > psv:issueInfo	
prism:volume	psv:whereUsed > psv:issueInfo	
prism:wordCount	psv:articleInfo	
publication Display Date	psv:whereUsed > psv:issueInfo	
pur:adultContentWarning	psv:usageRights	

PAM Heading Element	PSV Metadata Mapping
pur:agreement	psv:usageRights
pur:copyright	psv:usageRights
pur:creditLine	psv:usageRights
pur:embargoDate	psv:usageRights
pur:exclusivityEndDate	psv:usageRights
pur:expirationDate	psv:usageRights
pur:imageSizeRestriction	psv:usageRights
pur:optionEndDate	psv:usageRights
pur:permissions	psv:usageRights
pur:restrictions	psv:usageRights
pur:reuseProhibited	psv:usageRights
pur:rightsAgent	psv:usageRights
pur:rightsOwner	psv:usageRights

3.2.8 Example Metadata Tagging

Here is an example of the metadata head coded in PRISM/PAM markup:

```
<?xml version='1.0' encoding='UTF-8'?>
<pam:message>
  <pam:article xml:lang="en-US">
    <head>
      <dc:identifier>20110502044</dc:identifier>
      <prism:issueIdentifier>SI20110502</prism:issueIdentifier>
      <prism:originPlatform prism:platform="print"/>
      <dc:title>Oh, The Places They&apos;ll Go</dc:title>
      <dc:creator role="writer">Phil Taylor</dc:creator>
      <prism:publicationName>Sports Illustrated</prism:publicationName>
      <prism:issn>0038-822X</prism:issn>
      <prism:coverDate>2011-05-02</prism:coverDate>
      <prism:coverDisplayDate>May 2, 2011</prism:coverDisplayDate>
      <prism:volume>114</prism:volume>
      <prism:number>18</prism:number>
      <prism:startingPage>44</prism:startingPage>
      <prism:section>BASEBALL</prism:section>
      <prism:subsection1>THE ROCKIES</prism:subsection1>
      <prism:wordcount>2175</prism:wordcount>
      <prism:genre>feature</prism:genre>
      <prism:organization>Colorado Rockies</prism:organization>
```

Here is the new nextPub metadata block divided into blocks for easy access and to aid transformations into EPUB3:

```
<psv:description>
   <prism:genre>feature</prism:genre>
   <dc:creator prism:role="writer">Phil Taylor</dc:creator>
   <prism:wordcount>2175</prism:wordcount>
   sport>baseball
   <prism:organization>Colorado Rockies</prism:organization>
 </nextPub:description>
 <psv:whereUsed>
   <psv:aggregationInfo>
     <prism:aggregationType>magazine</prism:aggregationType>
     <prism:platform>tablet</prism:platform>
     <prism:device>iPad 1</prism:device>
   </psv:aggregationInfo>
   <psy:publicationInfo>
     <prism:publisher>IDEAlliance</prism:publisher>
     prism:publicationName>Spectrum
   </psv:publicationInfo>
   <psv:issueInfo>
     <prism:issueIdentifier>SI20110502</prism:issueIdentifier>
     <prism:coverDate>2011-05-02</prism:coverDate>
     <prism:coverDisplayDate>May 2, 2011</prism:coverDisplayDate>
     <prism:volume>114</prism:volume>
     orism:number>18
     </psv:issueInfo>
   <articleInfo>
     <prism:startingPage>44</prism:startingPage>
     <prism:section>BASEBALL</prism:section>
     subsection1>THE ROCKIES
   </articleInfo>
 </nextPub:whereUsed>
</nextPub:metadata>
```

4 TAGGING CONTENT

When the nextPub working group decided to base their tagging on HTML5, they had two choices. First they could extend HTML5 to provide special nextPub functionality. Or they could code content with valid HTML5 that could be handled by any HTML5 compliant browser or processor. The Working Group opted for the latter.

4.1 What Does this Mean?

The implication of this decision was that no direct use of special PAM or PIM elements for the body of the article can be included in nextPub HTML5. This means that pam:media and the <pim: elements could not be used. The PAM media structure was replaced with the new HTML5 figure element. And the <pim: elements can be mapped to a new PRISM Inline Class Markup Controlled Vocabulary to be used as classes for the HTML5 <p>span tag. See the PRISM Source Vocabulary Specification [PSVS] for details about nextPub tagging for the body of an article.

4.2 Sample Coding for Body Content

Here is PRISM/PAM content encoding:

```
<body>
     <h1>Oh, The Places They&apos;ll Go</h1>
     <span class="emphasis">Why are fans so aglow in Colorado? Because Tulo,
CarGo and Ubaldo-and a park that no longer plays like the game's highest
plateau-are making the Rockies' title hopes grow</span>
     <span class="prism:lead-in">It seemed</span> like a good idea to Troy
Tulowitzki when he allowed fans to choose his walk-up music-the song played
over the <pim:location>Coors Field</pim:location> public-address system as he
approaches the batter's box-in an online poll before the season. But after
the Rockies' prolific shortstop followed his standout 2010 (.315 average,
27 home runs and a fifth-place finish in the National League MVP voting) by
going hitless in his first eight at bats of this year, he quickly dumped the
people's choice, Katy Perry's <span class="emphasis">Firework,</span>
for something more in his comfort zone-<span class="emphasis">Baby,</span> by
<pim:person>Justin Bieber</pim:person>.
     <div prism:class="pullQuote">
       <quot; WE TALK ALL THE TIME ABOUT HOW TO SET THE RIGHT TONE, &quot;</p>
CARGO SAYS OF TULO. " WHAT CAN WE DO TODAY TO MAKE THIS TEAM
BETTER?"
     </div>
```

Here is the same content coded in HTML5:

</aside>

5 TAGGING MEDIA

This section provides documentation and guidelines for encoding media blocks. The <pam:media element is used to communicate with aggregators about the media that was present in the article when it was presented in the print magazine. In most cases the element <pam:media falls at the end of the article and users do not attempt to use the tag to actually place a media element in the flow of text. Very few publishers actually send the media elements to aggregators, they simply alert the aggregators about media that appeared in the original publication.

In nextPub, on the other hand, the HTML5 figure, which replaces <pam:media, is used to not only place media in the proper position within the flow of text, but also to call in media elements so they can be rendered with the text when delivered to a display technology. Hence the tagging within an HTML5 <figure is much different than tagging allowed in PAM <media.

5.1 About PAM Media

The structure of <pam:media is illustrated in Figure 5.1. This structure is made up of many optional metadata fields about the media object related to the article. Although it was *not* the intent, some publishers who are using PAM tagging as a content source format may be using the <pam:media tag to place the media (using the mediaReference) and to display it in the text. Figure 5.1 shows the structure of PAM media with indications about how each tag may be used to capture metadata about the media or to tag the content of the media block.

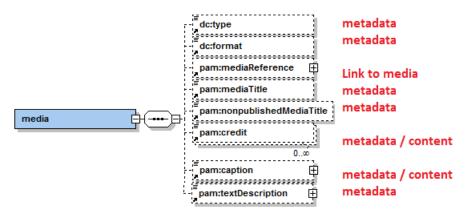


Figure 5.1 PAM Media Structure

5.2 About nextPub Media

The nextPub media block concept was initially based on the pam:media element currently used in the existing PRISM Aggregator Message. A media block presents *one or more media objects along with ancillary text content* so it is what is known as a "grouping" structure. After the decision was made to base nextPub on HTML5, the <figure element was selected for coding a media block because it contains richness required to express the wide variety of media blocks that are found in tablet editions. See Figure 5.2.

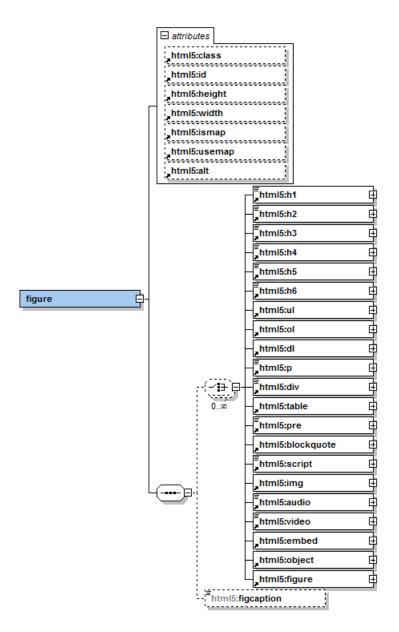


Figure 5.2 Figure Model

Two PRISM Figure Content Class metadata vocabulary terms have been developed for use with any text structure allowed in the figure. The terms included in this CV map directly to PAM content elements. In nextPub, this mechanism replaces the use of pam: content elements which are not HTML5 compatible:

#credit; An acknowledgement, appearing in the style of a caption.

#mediaTitle: The title of the media block or figure

5.3 Sample Coding

Here is a pam:media encoded media object:

Here is the equivalent HTML5 encoding:

5.4 About Tagging Media Metadata in nextPub

One feature of the pam:media model which is difficult to represent in nextPub HTML5 is the inclusion of metadata for each media object. In nextPub, each of these media objects (such as a photo or a video) is considered to be a *Component*. Metadata for each media component is encoded using the nextPub component metadata encoding model. See PRISM Source Vocabulary Specification [PSVS] for complete documentation about tagging component metadata.

In this model, some fields	have equivalents and	for most you can use	the new PSV <meta tag:<="" th=""/>

Field description	PAM element	nextPub Metadata field -
The title of the resource.	<pam:mediatitle< td=""><td>dc:title</td></pam:mediatitle<>	dc:title
An account of the content of the resource.	<pre><pam:nonpubishedmediatitle< pre=""></pam:nonpubishedmediatitle<></pre>	<meta <br="" prefix="pam"/> name="nonpublishedMediaTitle" content="xxx" />
Gives credit the provider of the image. This is not necessarily the same as the credit line that will appear in a publication.	<pam:credit< td=""><td><meta content="xxx" name="credit" prefix="pam"/></td></pam:credit<>	<meta content="xxx" name="credit" prefix="pam"/>
A brief publishable synopsis/summary of the contents of the photograph. This is not the same as the title.	<pam:caption< td=""><td><meta content="xxx" name="captione" prefix="pam"/></td></pam:caption<>	<meta content="xxx" name="captione" prefix="pam"/>
The main topic or topics of the content of the resource.	<pre><pam:textdescription< pre=""></pam:textdescription<></pre>	<meta content="xxx" name="textDescription" prefix="pam"/>

Note: For PSV, the component model requires a reference to a unique identifier for each component that appears in the body and a title.

While the majority of <pam:media is captured within the nextPub components metadata block and linked to the media object in the text with id/refines attribute pair, two key elements fall within the HTML5 body.

This table provides a mapping from <pam:media to HTML5:

PAM element	HTML5 Element		
<pre><pam:media< pre=""></pam:media<></pre>	<figure< td=""></figure<>		
<pam:mediaref< td=""><td>The media references within a figure (or media block) are coded with either <img, <audio="" <embed,="" <pam:mediaref<="" <video.="" article.="" body="" equivalent="" in="" into="" is="" link="" media="" object="" of="" or="" pull="" refid="*" src="is" td="" the="" this="" to=""></img,></td></pam:mediaref<>	The media references within a figure (or media block) are coded with either <img, <audio="" <embed,="" <pam:mediaref<="" <video.="" article.="" body="" equivalent="" in="" into="" is="" link="" media="" object="" of="" or="" pull="" refid="*" src="is" td="" the="" this="" to=""></img,>		

A figure is a media block and so the figure typically presents *one or more media objects along with ancillary text content*. This means that the figure is what HTML5 calls a "grouping structure." The figure bears an "id" that is unique within the article and links the media to a <psv:mediaComponent block with metadata for the media object.

Note: References to the media only appear in the HTML5 body text and not in the nextPub component media block.

Appendix A PAM Sample Tagging

```
<?xml version='1.0' encoding='UTF-8'?>
<pam:message xmlns="http://www.w3.org/1999/xhtml"</pre>
xmlns:TimeInc="http://www.timeinc.com/PRISM/2.1/"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:pam="http://prismstandard.org/namespaces/pam/2.0/"
xmlns:pim="http://prismstandard.org/namespaces/pim/2.0/"
xmlns:prism="http://prismstandard.org/namespaces/basic/2.0/"
xmlns:prl="http://prismstandard.org/namespaces/prl/2.0"
xmlns:pur="http://prismstandard.org/namespaces/prismusagerights/2.1/"
xmlns:xhtml="http://www.w3.org/1999/xhtml"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" prism:schemaVersion="String"
xsi:schemaLocation="http://www.w3.org/1999/xhtml
http://emagweb.timeinc.com/PRISM_2.1/PAMTimeInc.xsd">
  <pam:article xml:lang="en-US">
    <head>
      <dc:identifier>20110502044</dc:identifier>
      <prism:issueIdentifier>SI20110502</prism:issueIdentifier>
      <pam:status/>
      <prism:originPlatform prism:platform="print"/>
      <dc:title>Oh, The Places They&apos;ll Go</dc:title>
      <dc:creator role="writer">Phil Taylor</dc:creator>
      <prism:publicationName>Sports Illustrated</prism:publicationName>
      <prism:issn>0038-822X</prism:issn>
      <prism:coverDate>2011-05-02</prism:coverDate>
      <prism:coverDisplayDate>May 2, 2011</prism:coverDisplayDate>
      <prism:volume>114</prism:volume>
      <prism:number>18</prism:number>
      <prism:startingPage>44</prism:startingPage>
      <prism:section>BASEBALL</prism:section>
      <prism:subsection1>THE ROCKIES</prism:subsection1>
      sm:wordCount/>
    </head>
    <body>
      <span class="emphasis">Why are fans so aglow in Colorado? Because Tulo, CarGo
and Ubaldoâ€"and a park that no longer plays like the game's highest
plateauâ€"are making the Rockies' title hopes grow</span>
      >
        <pam:media>
          <dc:type>PHOTO</dc:type>
          <pam:mediaReference pam:refid="110540_opy6-95787-rawCOMBO.jpg"/>
          <pam:mediaTitle>PEAK PERFORMERS</pam:mediaTitle>
          -
<pam:credit>ROBERT BECK</pam:credit>
          <pam:caption>Gonzalez, Tulowitzki and Jimenez are all entering their primes
and locked up through at least 2014, giving the Rockies the game's best young
core.</pam:caption>
        </pam:media>
      <span class="prism:lead-in">It seemed</span> like a good idea to Troy
Tulowitzki when he allowed fans to choose his walk-up musicâ€"the song played over the
Coors Field public-address system as he approaches the batter's boxâ€"in an
online poll before the season. But after the Rockies' prolific shortstop followed
his standout 2010 (.315 average, 27 home runs and a fifth-place finish in the National
League MVP voting) by going hitless in his first eight at bats of this year, he
quickly dumped the people's choice, Katy Perry's <span
class="emphasis">Firework,</span> for something more in his comfort zoneâ€"<span
class="emphasis">Baby,</span> by Justin Bieber.
```

```
It might be a stretch to say that Tulo, as he's widely known, has Bieber
fever, but he is partial to the teen pop star's music and doesn't care who
knows it. He and teammate Jason Giambi took in a Bieber concert during spring
training, and Tulo is willing to put up with the inevitable ribbing from the rest of
the Rockies, who keep his locker at Coors stocked with Bieber-obilia. Last week his
space was adorned with a glittery backpack and a T-shirt bearing the singer's
likeness that most middle school girls would surely be proud to own. "Lots of
comedians in this clubhouse, " Tulowitzki says. " I just go with the
flow."
     >
       <pam:media>
         <dc:type>PHOTO</dc:type>
         <pam:mediaReference pam:refid="110540_opy6-98460-raw.jpg"/>
         <pam:mediaTitle>DOUBLE THREAT</pam:mediaTitle>
         <pam:credit>ROBERT BECK</pam:credit>
         <pam:caption>Tulowitzki and Gonzalez have a thriving partnership in the
lineup, in the clubhouse and in the NL MVP balloting, where both finished in the top
five last year.</pam:caption>
       </pam:media>
     Even though they might not want to trade iPods with Tulowitzki, 26, none of
the Rockies would dream of suggesting he change his tune. He went on one of his
familiar Tulo tears shortly after the sound track switch, almost single-handedly
demolishing the Mets with home runs in four straight games, and through Sunday he was
hitting .333 with seven homers, tying him with the Cardinals' Albert Pujols and
Ryan Braun of the Brewers for the National League home run lead. Besides, although his
new signature song isn't exactly a lyrical masterpiece, it does have one
lineâ€"in fact, it's repeated so much it seems like the only lineâ€"that's
especially appropriate for the Rockies: <span class="emphasis">Baby, baby, baby,
oh.</span>
     >
       <pam:media>
         <dc:type>VIDEO</dc:type>
         <pam:mediaReference</pre>
pam:refid="http://mlb.com/video/play.jsp?content_id=2586034"/>
         <pam:caption>A HUMOROUS LOOK AT THE COORS HUMIDOR</pam:caption>
       </pam:media>
      Put the emphasis on the <span class="emphasis">oh.</span> That&apos;s a
joyful sound in Colorado these days, thanks not just to Tulo, but to CarGo and Ubaldo.
CarGo is otherwise known as leftfielder Carlos Gonzalez, 25, last year' s NL
batting champ. The linchpin of the Rockies' staff is Ubaldo Jimenez, 27, who
finished third in the NL Cy Young Award voting in 2010. Due mostly to a thumb injury
that forced him to miss two starts, Jimenez has started slowly, but the Rockies
haven't. Tulowitzki's performance, both with bat and glove, and solid
pitching from starters Jhoulys Chacin and Jorge De La Rosa, as well as the entire
bullpen, helped put Colorado atop the NL West with a 14â€"7 record through Sunday. The
Rockies appear set to contend not just this year but for the foreseeable future, and
it's largely, as the Biebs might say, because of their O's, baby.
      It's also because of a long-term philosophy that ensures the O trio
won't be leaving anytime soon. Back in the '90s, when the Rockies were
playing something closer to pinball than baseball because of the way the ball carried
in the mile-high altitude of their home park, finding one of baseball' s most
sensible organizational blueprints in Colorado would have seemed as likely as
discovering a gourmet meal in an Easy-Bake Oven. But just as the introduction of a
humidor at Coors in 2002â€"baseballs stored in it don't carry like golf
ballsâ€"has normalized the game in Denver, the Rockies have devised a smart, down-to-
earth approach.
     >
       <pam:media>
         <dc:type>SLIDESHOW</dc:type>
         <pam:mediaReference pam:refid="110540_ROCKIES_iPad_ss.jpg"/>
```

```
<pam:mediaTitle>SLIDESHOW: SEE THE VISITING PITCHERS WHO PERFORM BEST AT
COORS FIELD</pam:mediaTitle>
          <pam:credit>ROBERT BECK</pam:credit>
          <pam:caption>TIM LINCECUM In nine career starts at Coors Field, the Freak
has a 3.49 ERA, the best among all hurlers who have thrown at least 40 innings there
and never played for the Rockies.</pam:caption>
       </pam:media>
      >
        <pam:media>
          <dc:type>SLIDESHOW</dc:type>
          <pam:mediaReference pam:refid="110540_ROCKIES_iPad_ss2.jpg"/>
          <pam:mediaTitle>TOM GLAVINE</pam:mediaTitle>
          <pam:credit>BRIAN BAHR/GETTY IMAGES</pam:credit>
          <pam:caption>The Braves lefty made 13 starts at Coors (including eight in
the prehumidor days) and had a 3.68 ERA there, second best among
visitors.</pam:caption>
        </pam:media>
      >
        <pam:media>
          <dc:type>SLIDESHOW</dc:type>
          <pam:mediaReference pam:refid="110540_ROCKIES_iPad_ss3.jpg"/>
          <pam:mediaTitle>DAN HAREN</pam:mediaTitle>
          <pam:credit>GARRETT W. ELLWOOD/GETTY IMAGES</pam:credit>
          <pam:caption>The righthander (now with the Angels) has made seven starts at
Coors and gone 4\hat{a} \in 2 with a 3.80 ERA there, third best among visiting
pitchers.</pam:caption>
       </pam:media>
      >
        <pam:media>
          <dc:type>SLIDESHOW</dc:type>
          <pam:mediaReference pam:refid="110540_ROCKIES_iPad_ss4.jpg"/>
          <pam:mediaTitle>BRETT MYERS</pam:mediaTitle>
          <pam:credit>DUSTIN BRADFORD/ICON SMI</pam:credit>
          <pam:caption>The righty (now with the Astros) has a career ERA of 4.18"but
that figure drops to 3.92 (with a 5"0 record) in his six starts at
Coors.</pam:caption>
       </pam:media>
      <q>
        <pam:media>
          <dc:type>SLIDESHOW</dc:type>
          <pam:mediaReference pam:refid="110540_ROCKIES_iPad_ss5.jpg"/>
          <pam:mediaTitle>BRAD PENNY</pam:mediaTitle>
          <pam:credit>ROBERT BECK</pam:credit>
          <pam:caption>In 11 starts at Coors, the righty (now with the Tigers) is 6'"1
with a 3.92 ERA, tying him with Myers for the fourth-best mark there among
visitors.</pam:caption>
        </pam:media>
      Having been burned in the past by throwing big bucks at older free agents who
failed miserably on the field and sometimes embarrassed the franchise away from it,
Colorado's front office has become more selective about making major financial
commitments. They focus on retaining players who appear to be entering their prime
rather than signing older, more established stars, and before the Rockies offer a
long-term deal, they have to be satisfied not just with a player's skills but
also with the content of his character. " We found that talent that isn' t
also accompanied by other qualities, such as humility, accountability and integrity,
really didn't work for us," says general manager Dan O'Dowd.
" We' ve tried to build this team not just with a certain kind of player but a
certain quality of person."
```

Gonzalez, Jimenez and Tulowitzki are perfect templates for the Rockies' vision"young, gifted, industrious and squeaky-clean. That's why Colorado locked up CarGo with a seven-year, \$80 million contract extension in January, two months after it extended Tulowitzki's contract to make it a 10-year deal worth \$157.75 million. They got in on the ground floor with Jimenez as well, signing him to a fouryear deal in 2009 (with team options for '13 and '14) that could earn him \$22.75 million and keep him in Denver two years after he would have been eligible for free agency.

The pair complement each other so well that when Gonzalez bought a new
Ferrari after his contract extension, he chose red instead of his preferred black
because Tulowitzki already owns a black one. "Got to keep things balanced
out," CarGo says. They work together as team leaders just as well.
"Obviously the Latin guys on this team and in the minors look up to Carlos,"
Tulowitzki says. "Some of the other young guys come up to me and ask me
questions, so that part of it works out real well."

<pam:media>

<dc:type>PHOTO</dc:type>

<pam:mediaReference pam:refid="110540_opy6-96398-raw.jpg"/>

<pam:mediaTitle>PRECIOUS CARGO</pam:mediaTitle>

<pam:credit>ROBERT BECK</pam:credit>

</pam:media>

The greater emphasis on personal character post-Neagle was at first seen as a religious-based movementâ€"a 2006 USA Today article described the organization as following a "Christian-based code of conduct"â€"but O'Dowd says that decisions about players have never been based on the nature of their faith. "Do we like players with character? Yes. With strong moral values? Yes," he says. "It is not required that a player be deeply religious to have those qualities."

The approach is less about religion than about an intelligent design. The effort to pay a little more to lock up promising players before they can break the bank in free agency makes good financial sense, and the ability to retain budding stars is especially appreciated by Colorado fans, who have seen some of the market's most prominent athletes, like Carmelo Anthony of the Nuggets, the Broncos' Jay Cutler and the Rockies' Matt Holliday, leave town in the last three years.

peers marvel. "One of the best things about my job is I get to watch him for free," says Tulowitzki.

Signed at 16 by the Diamondbacks, he spent four years in the minors as a
highly valued prospect until he was dealt in 2007 to Oakland with a package of players
for pitcher Dan Haren. Despite showing flashes of his prodigious talent in Oakland,
Gonzalez lasted only 11 months there before the A's traded him to Colorado with
pitchers Huston Street and Greq Smith for Holliday.

"It feels like it took a long time," Gonzalez says, "and then
it came all at once." The Rockies sent him to Triple A Colorado Springs when they
first acquired him, and after they brought him up, he scuffled along, batting around
.200 for a long stretch. But once he got settled in Colorado, stardom did come in a
rush. He hit .284 in 89 games in 2009 before tearing through the league last year with
a .336 average, 34 homers, 117 RBIs and 26 stolen bases. The key to his breakout?
"Being comfortable," he says. "Getting traded twice was hard; it kind
of shakes you up. Once I knew I was going to be here for good, that they really wanted
me here, everything fell into place."

The one quality that Gonzalez hasn't yet developed is patience at the
plate, where he shows off that sweet swing a little too often. "When he got here,
his strike zone was from the bill of his cap to the top of his spikes," says
manager Jim Tracy. "He still needs to be more selective, but he's getting
better." According to Fangraphs.com, 37% of Gonzalez's swings last season
were at pitches outside the strike zone. Only 14 hitters chased bad balls more often.
So far this season he has reduced that number to 30.9%.

"I am trying to be patient and make the pitchers throw strikes, but at
the same time I don't want to lose my aggressiveness," Gonzalez says.
"It takes time. You have to learn from your mistakes, and in time you figure it
out." That's as true for a franchise as it is for a ballplayer. In fact, the
Rockies' trio of O's are the perfect symbols of their franchiseâ€"young,
learning, on the cusp of something special. You get the feeling it won't be long
now. They're just about to figure it out.

```
5.47 2002
5.07 2003
6.27 2004
5.18 2005
4p>4.72 2006
```

```
4.34 2007
        4.83 2008
       4.41 2009
       4.25 2010
       3.91 2011
       >
         <pam:media>
           <dc:type>PHOTO</dc:type>
           <pam:mediaReference pam:refid="110540_opy6-96359-raw.jpg"/>
            <pam:credit>ROBERT BECK</pam:credit>
         </pam:media>
       </div>
      <pam:media>
       <dc:type>WEBPAGE</dc:type>
        <pam:mediaReference</pre>
pam:refid="http://sportsillustrated.cnn.com/2011/magazine/04/25/colorado.rockies.relat
ed.may2/index.html"/>
      </pam:media>
      <pam:media>
        <dc:type>WEBPAGE</dc:type>
        <pam:mediaReference</pre>
pam:refid="http://sportsillustrated.cnn.com/2011/magazine/04/25/related.galleries.may2
/index.html"/>
      </pam:media>
      <pam:media>
       <dc:type>WEBPAGE</dc:type>
        <pam:mediaReference</pre>
pam:refid="http://sportsillustrated.cnn.com/vault/article/magazine/MAG1184832"/>
      </pam:media>
    </body>
  </pam:article>
</pam:message>
```

Appendix B Sample PSV Tagging

```
<?xml version='1.0' encoding='UTF-8'?>
  <psv:psv>
   <psv:metadata>
     sm:contentType>article
     <psv:uniqueID>
      <psv:IDBlk>
        <dc:identifier>20110502044</dc:identifier>
        <dc:title> Oh, The Places They&apos;ll Go</dc:title>
      </nextPub:IDBlk>
     </nextPub:uniqueID>
     <psv:description>
       <prism:genre>feature</prism:genre>
       <prism:genre>opinion</prism:genre>
       <dc:creator prism:role="writer">Phil Taylor</dc:creator>
       sm:wordcount>2175
       <prism:sport>baseball</prism:sport>
       <prism:organization>Colorado Rockies</prism:organization>
     </nextPub:description>
     <psv:whereUsed>
       <psv:aggregationInfo>
         <prism:aggregationType>magazine</prism:aggregationType>
         <prism:platform>tablet</prism:platform>
         <prism:device>iPad 1</prism:device>
       </psv:aggregationInfo>
       <psv:publicationInfo>
         <prism:publisher>IDEAlliance</prism:publisher>
         <prism:publicationName>Spectrum</prism:publicationName>
       </psv:publicationInfo>
       <psv:issueInfo>
         <prism:issueIdentifier>SI20110502</prism:issueIdentifier>
         <prism:coverDate>2011-05-02</prism:coverDate>
         <prism:coverDisplayDate>May 2, 2011</prism:coverDisplayDate>
         <prism:volume>114</prism:volume>
         <prism:number>18</prism:number>
         <prism:issn>0038-822X</prism:issn>
       </psv:issueInfo>
       <articleInfo>
         <prism:startingPage>44</prism:startingPage>
         <prism:section>BASEBALL</prism:section>
         subsection1>THE ROCKIES
       </articleInfo>
     </nextPub:whereUsed>
    </nextPub:metadata>
    <head>
      <base href="http://www.timeinc.com/SI/SI20110502" target="_blank" />
      <link rel="stylesheet" type="text/css" href="SI_style.css" />
      <link rel="alternate" type="application/pdf" href="Article44.pdf" />
      <script src="js/scripts.js" type="text/javascript"/>
    </head>
    <body>
     <article>
     <span class="emphasis">Why are fans so aglow in Colorado? Because Tulo, CarGo
and Ubaldoâ€"and a park that no longer plays like the game's highest plateau"are
Rockies' title hopes grow</span>
     >
       <figure class="prism:photo">
```

```
<img src="110540_opy6-95787-rawCOMBO.jpg" alt="Peak Performers"/>
         PEAK PERFORMERS
         ROBERT BECK
         <figcaption>Gonzalez, Tulowitzki and Jimenez are all entering their primes
and locked up through at least 2014, giving the Rockies the game's best young
core.</figcaption>
       </figure>
      <span class="prism:lead-in">It seemed</span> like a good idea to Troy
Tulowitzki when he allowed fans to choose his walk-up musicâ€"the song played over the
Coors Field public-address system as he approaches the batter's box"in an online
poll before the season. But after the Rockies' prolific shortstop followed his
standout 2010 (.315 average, 27 home runs and a fifth-place finish in the National
League MVP voting) by going hitless in his first eight at bats of this year, he
quickly dumped the people's choice, Katy Perry's <span
class="emphasis">Firework,</span> for something more in his comfort zone"<span
class="emphasis">Baby,</span> by Justin Bieber.
      It might be a stretch to say that Tulo, as he's widely known, has Bieber
fever, but he is partial to the teen pop star's music and doesn't care who
knows it. He and teammate Jason Giambi took in a Bieber concert during spring
training, and Tulo is willing to put up with the inevitable ribbing from the rest of
the Rockies, who keep his locker at Coors stocked with Bieber-obilia. Last week his
space was adorned with a glittery backpack and a T-shirt bearing the singer's
likeness that most middle school girls would surely be proud to own. "Lots of
comedians in this clubhouse, " Tulowitzki says. " I just go with the
flow."
      >
       <figure class="prism:photo">
         <img src="110540_opy6-98460-raw.jpg" alt="Double Threat"/>
         DOUBLE THREAD
         ROBERT BECK
         <figcaption>Tulowitzki and Gonzalez have a thriving partnership in the
lineup, in the clubhouse and in the NL MVP balloting, where both finished in the top
five last year.</figcaption>
       </figure>
      Even though they might not want to trade iPods with Tulowitzki, 26, none of
the Rockies would dream of suggesting he change his tune. He went on one of his
familiar Tulo tears shortly after the sound track switch, almost single-handedly
demolishing the Mets with home runs in four straight games, and through Sunday he was
hitting .333 with seven homers, tying him with the Cardinals' Albert Pujols and
Ryan Braun of the Brewers for the National League home run lead. Besides, although his
new signature song isn't exactly a lyrical masterpiece, it does have one
lineâ€"in fact, it's repeated so much it seems like the only lineâ€"that's
especially appropriate for the Rockies: <span class="emphasis">Baby, baby, baby,
oh.</span>
      >
        <figure class="prism:video">
         <embed width="320" height="240"</pre>
src="http://mlb.com/video/play.jsp?content_id=2586034"/>
         A HUMOROUS LOOK AT THE COORS HUMIDOR
        </figure>
      Put the emphasis on the <span class="emphasis">oh.</span> That&apos;s a
joyful sound in Colorado these days, thanks not just to Tulo, but to CarGo and Ubaldo.
CarGo is otherwise known as leftfielder Carlos Gonzalez, 25, last year' s NL
batting champ. The linchpin of the Rockies' staff is Ubaldo Jimenez, 27, who
finished third in the NL Cy Young Award voting in 2010. Due mostly to a thumb injury
that forced him to miss two starts, Jimenez has started slowly, but the Rockies
haven't. Tulowitzki's performance, both with bat and glove, and solid
pitching from starters Jhoulys Chacin and Jorge De La Rosa, as well as the entire
bullpen, helped put Colorado atop the NL West with a 14â€"7 record through Sunday. The
```

```
Rockies appear set to contend not just this year but for the foreseeable future, and
it's largely, as the Biebs might say, because of their O's, baby.
     It's also because of a long-term philosophy that ensures the O trio
won't be leaving anytime soon. Back in the '90s, when the Rockies were
playing something closer to pinball than baseball because of the way the ball carried
in the mile-high altitude of their home park, finding one of baseball's most
sensible organizational blueprints in Colorado would have seemed as likely as
discovering a gourmet meal in an Easy-Bake Oven. But just as the introduction of a
humidor at Coors in 2002â€"baseballs stored in it don't carry like golf
ballsâ€"has normalized the game in Denver, the Rockies have devised a smart, down-to-
earth approach.
     <q>
     <figure class="prism:slideshow>
       SLIDESHOW: SEE THE VISITING PITCHERS WHO PERFORM
BEST AT COORS FIELD
       <figure class="prism:photo">
         <img src="110540_ROCKIES_iPad_ss.jpg"/>
         TIM LINCECUM
         ROBERT BECK
         <figcaption>TIM LINCECUM In nine career starts at Coors Field, the Freak has
a 3.49 ERA, the best among all hurlers who have thrown at least 40 innings there and
never played for the Rockies.</figcaption>
       </figure>
      <figure class="prism:photo">
         <img src="110540_ROCKIES_iPad_ss2.jpg" />
         TOM GLAVINE
         ROBERT BECK
         <figcaption>The Braves lefty made 13 starts at Coors (including eight in the
prehumidor days) and had a 3.68 ERA there, second best among visitors.</figcaption>
       </figure>
      <figure class="photo">
         <img src="110540_ROCKIES_iPad_ss3.jpg" />
         DAN HAREN
         GARRETT W. ELLWOOD/GETTY IMAGES
         <figcaption>The righthander (now with the Angels) has made seven starts at
Coors and gone 4"2 with a 3.80 ERA there, third best among visiting
pitchers.</figcaption>
       </figure>
      <figure class="prism:photo">
         <img src="110540_ROCKIES_iPad_ss4.jpg" />
         BRETT MYERS
         DUSTIN BRADFORD/ICON SMI
         <figcaption>The righty (now with the Astros) has a career ERA of 4.18"but
that figure drops to 3.92 (with a 5â€"0 record) in his six starts at
Coors.</figcaption>
       </figure>
      <figure class="prism:photo">
         <img src="110540_ROCKIES_iPad_ss5.jpg" />
         BRAD PENNY
         ROBERT BECK
         <figcaption>In 11 starts at Coors, the righty (now with the Tigers) is 6'"1
with a 3.92 ERA, tying him with Myers for the fourth-best mark there among
visitors.</figcaption>
       </figure>
     </figure>
     Having been burned in the past by throwing big bucks at older free agents who
failed miserably on the field and sometimes embarrassed the franchise away from it,
Colorado's front office has become more selective about making major financial
commitments. They focus on retaining players who appear to be entering their prime
rather than signing older, more established stars, and before the Rockies offer a
long-term deal, they have to be satisfied not just with a player's skills but
also with the content of his character. " We found that talent that isn' t
```

also accompanied by other qualities, such as humility, accountability and integrity, really didn't work for us," says general manager Dan O'Dowd. "We've tried to build this team not just with a certain kind of player but a certain quality of person."

Gonzalez, Jimenez and Tulowitzki are perfect templates for the Rockies' visionary"young, gifted, industrious and squeaky-clean. That's why Colorado locked up CarGo with a seven-year, \$80 million contract extension in January, two months after it extended Tulowitzki's contract to make it a 10-year deal worth \$157.75 million. They got in on the ground floor with Jimenez as well, signing him to a four-year deal in 2009 (with team options for '13 and '14) that could earn him \$22.75 million and keep him in Denver two years after he would have been eligible for free agency.

The pair complement each other so well that when Gonzalez bought a new
Ferrari after his contract extension, he chose red instead of his preferred black
because Tulowitzki already owns a black one. "Got to keep things balanced
out," CarGo says. They work together as team leaders just as well.
"Obviously the Latin guys on this team and in the minors look up to Carlos,"
Tulowitzki says. "Some of the other young guys come up to me and ask me
questions, so that part of it works out real well."

```
<figure class="prism:photo">
    <img src="110540_opy6-96398-raw.jpg" />
    PRECIOUS CARGO
    ROBERT BECK
```

```
</figure>
```

The greater emphasis on personal character post-Neagle was at first seen as a religious-based movementâ€"a 2006 USA Today article described the organization as following a "Christian-based code of conduct"â€"but O'Dowd says that decisions about players have never been based on the nature of their faith. "Do we like players with character? Yes. With strong moral values? Yes," he says. "It is not required that a player be deeply religious to have those qualities."

The approach is less about religion than about an intelligent design. The effort to pay a little more to lock up promising players before they can break the bank in free agency makes good financial sense, and the ability to retain budding stars is especially appreciated by Colorado fans, who have seen some of the market's most prominent athletes, like Carmelo Anthony of the Nuggets, the Broncos' Jay Cutler and the Rockies' Matt Holliday, leave town in the last three years.

5.47 20025.07 2003

Signed at 16 by the Diamondbacks, he spent four years in the minors as a
highly valued prospect until he was dealt in 2007 to Oakland with a package of players
for pitcher Dan Haren. Despite showing flashes of his prodigious talent in Oakland,
Gonzalez lasted only 11 months there before the A's traded him to Colorado with
pitchers Huston Street and Greg Smith for Holliday.

"It feels like it took a long time," Gonzalez says, "and then
it came all at once." The Rockies sent him to Triple A Colorado Springs when they
first acquired him, and after they brought him up, he scuffled along, batting around
.200 for a long stretch. But once he got settled in Colorado, stardom did come in a
rush. He hit .284 in 89 games in 2009 before tearing through the league last year with
a .336 average, 34 homers, 117 RBIs and 26 stolen bases. The key to his breakout?
"Being comfortable," he says. "Getting traded twice was hard; it kind
of shakes you up. Once I knew I was going to be here for good, that they really wanted
me here, everything fell into place."

The one quality that Gonzalez hasn't yet developed is patience at the
plate, where he shows off that sweet swing a little too often. "When he got here,
his strike zone was from the bill of his cap to the top of his spikes," says
manager Jim Tracy. "He still needs to be more selective, but he's getting
better." According to Fangraphs.com, 37% of Gonzalez's swings last season
were at pitches outside the strike zone. Only 14 hitters chased bad balls more often.
So far this season he has reduced that number to 30.9%.

"I am trying to be patient and make the pitchers throw strikes, but at
the same time I don't want to lose my aggressiveness," Gonzalez says.
"It takes time. You have to learn from your mistakes, and in time you figure it
out." That's as true for a franchise as it is for a ballplayer. In fact, the
Rockies' trio of O's are the perfect symbols of their franchiseâ€"young,
learning, on the cusp of something special. You get the feeling it won't be long
now. They're just about to figure it out.

```
<div class="now-on-si">
        >
         <span class="prism:lead-in">Now on Twitter</span>
       Follow <span class="pim:link">@SI_JonHeyman</span> and <span</p>
class="pim:link">@Jposnanski</span> for their up-to-the-minute baseball thoughts.
      <aside class="prism:pullQuote">
        <quot; WE TALK ALL THE TIME ABOUT HOW TO SET THE RIGHT TONE, &quot; CARGO</p>
SAYS OF TULO. " WHAT CAN WE DO TODAY TO MAKE THIS TEAM BETTER? " 
      </aside>
      <aside class="prism:box">
       <h1>ROCKY MOUNTAIN LOW</h1>
        <span class="emphasis">In their first nine seasons, the Rockies&apos; home
ERA was a lofty 6.07. But since '02, when they began using a humidor at Coors
Field to make balls behave more like they do at sea level, <span
class="pim:person">Ubaldo Jimenez</span> and the team&apos;s other pitchers have had
that mark in free fall.</span>
```

```
 6.27 2004 
       5.18 2005
       4.72 2006
       4.34 2007
       4.83 2008
       4.41 2009
       4.25 2010
       3.91 2011
       >
       <figure class="prism:photo">
        <img src="110540_opy6-96359-raw.jpg" />
        ROBERT BECK
      </figure>
      </aside>
   <div class="relatedWebContent">
href="http://sportsillustrated.cnn.com/2011/magazine/04/25/colorado.rockies.related.ma
y2/index.html">Colorado Rockies (Related Stories)</a>
href="http://sportsillustrated.cnn.com/2011/magazine/04/25/related.galleries.may2/inde
x.html">Related Galleries</a>
href="http://sportsillustrated.cnn.com/vault/article/magazine/MAG1184832">Related
Content</a>
   </div>
   </article>
   </body>
</psv:psv>
```