San Francisco Bay Area Regional Prosperity Plan

Development Tracking "Dashboard" Project Final Report

A. Project Profile (1 page maximum)

Project Name:	Development Tracking Dashboard
Lead and Partner Organizations:	ABAG
Primary Contact Person:	Duane Bay, ABAG, duaneb@abag.ca.gov 101 8th Street, Oakland 94607
Sub-Grant Program:	Housing the Workforce
Project Type:	Mitigating displacement; supporting preservation and production of affordable housing
Total Grant Amount:	$29,372
Total Match (if any):	100% = $29,372
Geographic Coverage of Project:	Region
Brief Description (150 words maximum):	Key purpose: Develop a tool to identify and track communities likely to experience gentrification, putting low-income residents at risk of displacement. This tracking tool, which should be replicable at the local jurisdiction level, will allow monitoring of projects by affordability and neighborhood to reveal shifts in development patterns. Underlying System Necessary to Accomplish Key Purpose: Develop and document a stakeholder-generated requirement specification for, and prototypical example of, an open-data, frequently updated, region-wide database of key public information about real estate development projects planned and in progress, at both the parcel and geographical sub-area levels, along with user-friendly tools for passive viewing (reports, maps) and active uses (such as supporting development of community-based applications).
Images:	Landing page on City of SanFrancisco's Development Pipeline Report website Schematic diagram of datasets that comprise the proposed Open Data Regional Development Tracking System

B. Project Description

1. Goals and Objectives

   On behalf of the San Francisco Bay Area Regional Prosperity Consortium, funded through HUD's Sustainable Communities Initiative, ABAG conducted a feasibility assessment for, developed a rough architecture of, and proposed a general implementation approach for an Open-Data Real Estate Development Tracking System for the region, with the capacity to support "dashboard" summary views of infill development patterns and other tools for policy analysis and decision support. The system would be an open-data, up-to-date, region-wide database of key public information about real estate development projects in progress, at parcel and aggregate level, along with user-friendly tools for passive viewing and active use.

   Timely, comprehensive, region-wide development data is scarce or expensive, hampering analysis of the patterns of development—how much is built, and of which kind, and at which time frames. Only with such comprehensive and timely data can we analyze in depth how land use policy affects development in general and housing production in particular. Having this information resource can inform public discourse about the regional implications of development, its potential co-benefits such as greenhouse gases (GHG) reduction and access to opportunity, and its potential collateral consequences such as economic displacement. Although this data is in the public domain, it is not easily accessible to interested parties, much less in a standard form that allows comparison across jurisdictions or over time. Existing data sources are some combination of too expensive, too out-of-date, insufficiently detailed, insufficiently comprehensive, error prone or blind to the pre-permit entitlement process.

2. Work Plan

   The work plan for this project was to (1) develop a "template" example of desired end-product, (2) conduct technical feasibility studies, and (3) a write a workplan for a comprehensive collaborative effort. San Francisco's existing "development pipeline reports" served as the template, and is summarized and highlighted as "best practice" in the Project Report. However, to date, efforts to develop a similar *regional* information asset have been partial and sporadic. Development of the system as a whole requires an effort that is more coordinated, systematic, comprehensive, sustained and collaborative than has been attempted to date. The Project Report describes the system and recommends a collaborative development approach.

3. Role of Lead and Partners

   This report was researched and written by Association of Bay Area Governments (ABAG), and funded by a Sustainable Communities Partnership grant from the U.S. Department of Housing and Urban Development (HUD) to Metropolitan Transportation Commission (MTC) and ABAG to administer on behalf of the Regional Prosperity Plan Consortium (RPPC). ABAG thanks the project's numerous stakeholders and participants, especially the Housing Working Group, Joint Project Team and Steering Committee of the RPPC. Representatives of the following RPPC partners made substantial design and/or editorial contributions: MTC, City/County of San Francisco, City of San Jose, Contra Costa County, California Housing Partnership Corporation, Nonprofit Housing Association of Northern California, Council of Community Housing Organizations (San Francisco), East Bay Housing Organizations, Housing Leadership Council of San Mateo County, Greenbelt Alliance, Transform, Urban Habitat, San Francisco Foundation.

C. Challenges and Outcomes

4. Challenges

Probably the biggest challenge was managing expectations. Most of the end-users (many of whom were RPP Consortium members) vastly underestimated the complexity and scope of the project, although they could describe the end result they wanted rather precisely. In essence the never-explicitly-state problem statement was: the **desired datasets are commercially available but cost too much** to be of general use to the public sector, non-profit sector and general public users; and further, because these datasets are not readily available for free or at low cost many of the tools for reporting, mapping and analysis that these end users and application developers could and would build are of little value without collected datasets to analyze and display, and **so these public-interest tools don't get built.** Based on the feasibility analysis we believe a free, open, useful system can be built, and we have described an approach. But there was admittedly substantial disappointment for some stakeholders who expected that a useful system or at least a system prototype could be delivered as the project outcome. Only one jurisdiction of more than 100 in the region, San Francisco, actually has an online open-data housing development pipeline tracking system. It was developed over ten years, championed, designed, resourced and managed within one organization. The project was NOT to replicate that at regional scale, but rather to scope that task, evaluate feasibility and propose an approach.

5. Outcomes

The Project Report is a blueprint for cities, regional agencies, regional and community-based non-profits to collaboratively building the envisioned system that has come to be known as "the dashboard." A "dashboard" equipped with instrumentation that allows Bay Area residents to see indicators of the pace, direction, and extent of development is the most visible aspect of an underlying system of data collection, curation and access. A dashboard can be thought of as the top layer or outer layer of a supporting system. While the key components of the system will be the data collected and, importantly, the anchoring of the processes and standards for doing so within local jurisdictions, the system will also provide an interface to the data collected and thus serve as a platform for viewing the data, accessing the data with third-party applications and, possibly, as an interface for updating it. Once the foundation is in place, users and application developers can build custom dashboards independently to suit their own particular needs.

The layers of the proposed Open-Data Real Estate Development Tracking System are as follows:

- Dashboard: Summary indicators of development pipeline in tabular and graphical (e.g., charts, maps, icons) formats—with emphasis on accessibility and priority on visualization.
- Portal: Tools and standards that support multiple means of accessing and using the data—downloading files, performing searches, writing apps that access the data dynamically.
- Pipeline Monitoring: Tools and processes for collecting real-time data on progress of projects through the pre-permit and post-permit development processes of planning, entitlement and construction
- Development Project Database: database of parcels where development has happened, is happening or is planned.
- Land Use Database: Databases of key land use information at parcel and jurisdiction granularity, including baseline "existing conditions" data to support analysis of change, such as development in the recent past or current pre-development use on a particular parcel.
- Collection & Maintenance System: Collaboratively derived data standards, staff time to collect and "clean" data, and the underlying database system and its maintenance.

D. Replicability and Dissemination (1 ½ pages are used for D & E combined)

6. Replicability
   The proposed system would be region wide. Because one of the major challenges to implementation is the state of readiness and open-data infrastructure in local jurisdictions, success will require providing tools and technical assistance and facilitating spread of effective practices among jurisdictions.

7. Tools and Resources
   While there was already a high consensus on tool design in very general terms, much more detailed information was needed to scope the effort, describe the component parts and how they would interact, and evaluate feasibility of several key technical aspects. The proposed approach to development and proposed project scope and example template reports *are* the "tools and resources" produced. In addition, the appendix of the Project Report includes a great deal of requirement specification for particular datasets, as well as feasibility analysis memos.

8. Sharing and Dissemination
   Many of the active stakeholders (mentioned above in #3) have already volunteered to be part of implementation working groups going forward, as well as several private sector and non-profit application developers who would like to use this information resource, once developed, and therefore wish to participate to its development.

E. Recommendations and Next Steps

9. Recommendations
   Because there was a high level of consensus for *what* to develop and *why*, the project focused on just documenting that consensus then setting the stage for *how* to proceed with a collaborative implementation effort. The following guiding principles characterize a recommended general approach that takes into account key design considerations outlined in the previous section. The particular datasets to be assembled, as well as the charter of each of the working groups that would specify and assemble them, are detailed the Project Report.

   - *Open-data.* Create datasets that, to the greatest extent possible, are free and freely accessible to the public, can be viewed on-line in useful ways, and ultimately can be downloaded and/or used through a published application programming interface.
   - *Incremental Refinement Along Key Dimensions.* Starting with an adequate kernel in 2015, steadily extend and refine the system incrementally along these seven dimensions (order not indicative of priority): (1) automating collection, (2) covering more jurisdictions, (3) distinguishing more use and development types, (4) collecting more attributes and in more detail, (5) increasing accuracy through standardization and verification, (6) capturing more pre-permit data, and (7) increasing ease and utility of public access.
   - *Three Complementary Kernels*. Work from “three ends toward the middle” or “three kernels out,” respectively: (1) semi-automated capture of parcel-level development data to be used in combination with data from other sources to improve resolution for land use modeling, and for which geographical coverage and official verification are lesser concerns because data aggregated at the tract or subarea level is adequate for modeling; (2) verified parcel-level data from all jurisdictions for housing development and affordable-housing development, (3) verified local land use policies.

- *Collaborative Methods/Standards Process.* Develop methods and standards collaboratively. The many decisions necessary about which attributes to collect, attribute definitions, which jurisdictions to prioritize, and more will be the domain of several working groups and a coordinating committee. The groups will include regional agency staff, city/county staff, tracking software vendors, as well as data users and application developers from private and public-interest sectors. It is anticipated that several organizations that have been active in the Regional Prosperity Consortium will be active participants in this collaboration. Each working group will generate a work plan that is appropriate to the level of resource available.
- *Open Development of Data Visualization Tools and Other Applications*. Focus on data collection, quality and access rather than on data display. Encourage and support public, private and public-interest collaborators to develop applications.
- *Agency Commitment to Core Data Collection and Maintenance*. Rely on the regional public agencies, and ABAG in particular, to host the database(s), perform some basic data collection and cleaning of certain fundamental datasets (e.g., Regional Parcel Inventory, Local Geographical Subareas), and provide administrative support to assure continuity for the collaborative working groups.

10. Next Steps
    The Project Report includes an action plan and first year implementation workplan.

Date: March 31, 2015

For: Housing Working Group, Regional Prosperity Consortium

By: Association of Bay Area Governments

Subject: Scoping an Open-Data Real Estate Development Tracking System for the San Francisco Bay Area with the Capacity to Support "Dashboard" Summary Views of Infill Development Patterns

Deliverable: This memorandum fulfills the requirements of Deliverable 2.3a of the Development Tracking Dashboard task, "Housing Development Tracking Template." Appendix G maps memo content to contract tasks.

What do we want?

Subject to definition of italicized terms as discussed later in this memo, we want an *open-data*, *up-to-date*, *region-wide* database of *key public information* about real estate development projects in progress, *at parcel and aggregate levels*, along with *user-friendly tools* for passive viewing and active use.

When do we want it?

Now! For many years, development of an Open-Data Real Estate Development Tracking System has been on the wish lists or action plans of many parties, including many organizations in the Sustainable Communities Initiative's Regional Prosperity Consortium. To date, efforts to develop this regional information asset have been partial and sporadic. Development of the system as a whole requires an effort that is more coordinated, systematic, comprehensive, sustained and collaborative than has been attempted to date.

This memo describes the system and recommends a collaborative development approach.

Contents

Executive Summary

On behalf of the San Francisco Bay Area Regional Prosperity Consortium, funded through HUD's Sustainable Communities Initiative, ABAG conducted a feasibility assessment for, developed a rough architecture of, and proposed a general implementation approach for an Open-Data Real Estate Development Tracking System for the region, with the capacity to support "dashboard" summary views of infill development patterns and other tools for policy analysis and decision support. The system would be an open-data, up-to-date, region-wide database of key public information about real estate development projects in progress, at parcel and aggregate level, along with user-friendly tools for passive viewing and active use.

Timely, comprehensive, region-wide development data is scarce or expensive, hampering analysis of the patterns of development—how much is built, and of which kind, and at which time frames. Only with such comprehensive and timely data can we analyze in depth how land use policy affects development in general and housing production in particular. Having this information resource can inform public discourse about the regional implications of development, its potential co-benefits such as green house gases (GHG) reduction and access to opportunity, and its potential collateral consequences such as economic displacement. Although this data is in the public domain, it is not easily accessible to interested parties, much less in a standard form that allows comparison across jurisdictions or over time. Existing data sources are some combination of too expensive, too out-of-date, insufficiently detailed, insufficiently comprehensive, error prone or blind to the pre-permit entitlement process.

The multi-year collaborative project would:

- Enable capture of: (1) parcel-level data about residential and non-residential development, semi-automatically; (2) verified parcel-level data from all jurisdictions for housing development in general and affordable-housing development in particular; (3) verified inventory of local land use policies; and (4) time-stamped history of the adoption/approval pipeline process for policies and projects.

- Focus on data collection, quality, and access rather than on data display, creating datasets that are free and freely accessible, can be viewed on-line in useful ways, and ultimately can be downloaded and/or used by collaborators to develop applications.

- Start with a kernel in 2015, steadily extend and refine the system incrementally along these dimensions: automation, geographical coverage, types of development tracked, breadth of data, accuracy, pre-permit visibility and ease of public access.

- Develop common methods and standards collaboratively via linked working groups comprising regional agency staff, city/county staff, tracking software vendors, as well as data users and application developers from public, private, and non-profit sectors.
- Rely on ABAG to host the database(s), perform some basic data collection and cleaning, and provide administrative support to assure continuity for the working groups. Additional work would proceed on an *ad hoc* basis as working groups recruit resources.

This endeavor would build on a foundation of existing assets. All jurisdictions are required to send an Annual Progress Report to California Department of Housing and Community Development (HCD), and HCD has demonstrated interest in cooperating. Several existing datasets related to existing land use and building permit issuance are maintained by regional and county agencies already, and they work closely together. The project will also benefit from the confluence of certain technical trends, especially the open data movement and crowdsourcing. Several community based organizations and regional non-profits who have been active participants in the Regional Prosperity Consortium already have developed centers of technical expertise, datasets, and analytic tools related to his project and are eager to participate. Finally, the City and County of San Francisco has set the benchmark with its Development Pipeline Report.

This endeavor would have to deal with some major constraints as well. Although all of the target data is public information, only a subset of it is required to be published in a compiled form, and then only annually, and these reports use differing data definitions. The data is typically originated in separate local departments, many without benefit of specialized tracking software. Often data must be collected by reading meeting packets on city/county websites, attending public meetings or communicating directly with busy city/county staff. Although a few jurisdictions do now post development "pipeline" data on their websites, attributes reported and data definitions are not standardized. Commercial data subscription services face the same obstacles to data quality and availability, and understandably their most valuable compilations are not free. Also, they simply do not collect some important data such as the affordability level of newly developed housing, or development in pre-permit stages. Many interested parties who do track projects' progress regard their data compilations as proprietary.

To date, efforts to develop this regional information asset have been partial and sporadic. Development of the system as a whole requires an effort that is more coordinated, systematic, comprehensive, sustained and collaborative than has been attempted to date.

This memo describes the system and recommends a collaborative development approach.

Purposes — Why Do We Want this Information Resource

Most noteworthy among the motivations and objectives voiced by various different stakeholders that have called for development of this information resource are the following:

- In general, timely, comprehensive, region-wide development data is scarce or expensive, hampering analysis of the patterns of development—how much is built, and of which kind, and at which time frames.
- In particular, there is no such source for tracking production of affordable housing, tracking progress toward residential and non-residential development projections in Priority Development Areas, or tracking projects at pre-permit stages in the pipeline. Existing data sources are some combination of too expensive, too out-of-date, insufficiently detailed, insufficiently comprehensive, error prone or blind to the pre-permit entitlement process.
- Only with such comprehensive and timely data can we begin to analyze how land use policy affects development in general and housing production in particular.
- Regional agencies have, to a degree, begun to link the allocation of regional public funds to local adoption and implementation of plans and policies that will tend to result in communities and a region that is more sustainable; therefore we need to be able to monitor the performance of those plans and policies.
- Although this data is in the public domain, it is not easily accessible to interested parties, much less in a standard form that allows comparison across jurisdictions or over time. While local jurisdictions have made great strides in making much data available, there is little incentive, or support for, coming up with common, pan-jurisdictional dataschema that is open, well documented, and widely adopted. That is our ultimate hope for this project.
- Having this information resource can inform public discourse about development that can enable communities to realize local aspirations; as well as the potential co-benefits of such development, such as GHG reduction and access to opportunity, and its potential collateral consequences such as economic displacement.

Design Requirements — Defining What We Want

What is meant by "an open, up-to-date, region-wide database of key public information about real estate development projects in progress, at parcel and aggregate levels, along with user-friendly tools for passive viewing and active use?"

Open-data means data that is freely accessible, free to use and transparent with respect to source and methods of collection and verification.

Up-to-date means many things. The US Census updates real estate development information annually for larger cities, and for smaller locales interpolates between five-year snapshots. The State and Federal governments require cities and counties to submit certain information annually, but it is not readily accessible and can be seriously lagged (stale). Commercial data vendors poll jurisdictions monthly to compile reports to service their subscribers. For many years real estate industry reporters and spotters for each county's Building Trades Council track some projects' progress on municipal websites and at public hearings then issue reports that are somewhat public, somewhat timely but not comprehensive. But now the era of real-time development tracking data is at hand through the power of the internet. The premiere example in the Bay Area is San Francisco, where information on a building permit issued is available to the public virtually the instant it is entered into the city's permit tracking database internally. The proposed system will capture certain data elements annually or quarterly, but will also support development of real-time applications by application developers as well.

The meaning of *region-wide* would appear to be obvious, given that all of the data of interest is available locally in public documents that are available in all jurisdictions. However, because there is great variance among jurisdictions' data management practices, limited resources dictate that data collection and verification is spotty. Current commercial data sources give the appearance of region-wide coverage because their commercially reasonable practices of getting "close enough" for their customers' purposes are not subject to transparency requirements. The proposed system would include verified information from all jurisdictions.

What constitutes *key public information* is a very important question, the answer to which depends on who is asking and how much resource they have to pursue the answer. Appendix A presents a detailed list of attributes we propose to collect. In general terms, however, we want to track:

- Parcel-level residential and non-residential development with key attributes such as affordability, unit type, development stage and timing;
- All major local land use policy adoptions; and, eventually,
- Parcel-level existing use and allowable development.

Granularity of data, meaning *disaggregated parcel-level data versus aggregated data,* is important as well. For many purposes, it is sufficient to have summary data for each census tract, transportation analysis zone, Priority Development Area or jurisdiction. For other purposes parcel-level data is essential. Recently updated parcel-level data is also an important ingredient for improving the resolution, calibration and currency of summary data. Therefore we propose to collect and use: (a) parcel-level development data received from and directly

verified with city/county staff, (b) parcel-level data available from batch open-data sources such as city websites, and (c) aggregated data from several public and private sources.

By *user-friendly tools* we mean that we aim, over time, to make the data available in all of the many ways that information users and application developers have become accustomed to. Initially, however, we propose to give priority to developing the data source and making it readily available in relatively passive forms that enable third parties and affiliates to develop visualization tools and other applications, rather than emphasizing application building by the project *per se*.

One of the best ways to understand and visualize what this system would be like overall is to look at a quarterly Pipeline Report published by the City/County of San Francisco. A copy of a recent report is included as Appendix B. While, on one hand, it is sobering to realize that the underlying system of standards, procedures and software, which covers but one city not the whole region, took approximately ten years to develop, on the other hand, expertise has been developed, lessons have been learned and an exemplary system sets a clear target.

Dashboard — Why Has the Envisioned System Frequently Been Called a Dashboard?

A "dashboard" equipped with instrumentation that allows Bay Area residents to see indicators of the pace, direction, and extent of development is the most visible aspect of an underlying system of *data collection, curation and access*. A dashboard can be thought of as the top layer or outer layer of a supporting system. While the key components of the system will be the data collected and, importantly, the anchoring of the processes and standards for doing so within local jurisdictions, the system will also provide an interface to the data collected and thus serve as a platform for viewing the data, accessing the data with third-party applications and, possibly, as an interface for updating it. Once the foundation is in place, users and application developers can build custom dashboards independently to suit their own particular needs.

For simplicity, the layers of the proposed Open-Data Real Estate Development Tracking System, although not strictly a "stack," can be described as follows:

- Dashboard: Summary indicators of development pipeline in tabular and graphical (e.g., charts, maps, icons) formats—with emphasis on accessibility and priority on visualization.
- Portal: Tools and standards that support multiple means of accessing and using the data, including downloading files, performing searches, writing applications that access the data dynamically.
- Pipeline Monitoring: Tools and processes for collecting real-time data on progress of projects through the pre-permit and post-permit development processes of planning,

entitlement and construction—with emphasis on timeliness and priority on local and regional impact.

- Development Project Database: database of parcels where development has happened, is happening or is planned.
- Land Use Database: Databases of key land use information at parcel and jurisdiction granularity, including baseline "existing conditions" data to support analysis of change, such as development in the recent past or current pre-development use on a particular parcel.
- Collection & Maintenance System: Collaboratively derived data standards, staff time to collect and "clean" data to these standards, and the underlying database system and its maintenance.

To repeat, this is a data collection system being referred to by its most visible feature, the "dashboards" that system curators and third-party developers will create to view and use information in the database.

Design Considerations — What Assets Do We Build Upon? What Are Key Constraints?

The desired outcome has been on wish lists and action lists of many agencies and organizations for many years, yet has not been accomplished to date. There are many reasons why. This memo is not retrospective or investigatory, but a scan of assets and constraints that bear upon the proposed scope of work and the recommended approach is appropriate and useful.

Key Assets to Build Upon

All jurisdictions are required to send an Annual Progress Report to State HCD. Although the timeliness, completeness and quality of this data source is problematic, as discussed further below, coverage is substantial and HCD is making a concerted effort to upgrade and has demonstrated willingness to coordinate with local and regional upgrade efforts.

Several existing datasets related to existing land use and building permit issuance are maintained by ABAG, MTC and some county Congestion Management Agencies. These agencies formed an Interagency Modeling Group (IMG) in 2014 for coordination. The IMG has a strategic commitment to open-data and a direct interest in improving the timeliness and quality of local land use data; therefore it is anticipated that existing technical resources can be leveraged for the collaborative undertaking described in this memo.

In particular, usable system assets are already in place because ABAG and/or MTC must collect and maintain certain datasets to fulfill existing state mandates (e.g., land use and

transportation modeling necessary to comply with state Sustainable Communities Strategy legislation) or regional application of state law (e.g., linkage of housing production performance as a factor in allocating One Bay Area Grant funds.

The project will benefit from the confluence certain of technical trends, especially the "open data movement" and "crowdsourcing," and political trends, especially "transparency" of public information and "linkage" of allocation of public funds to policy performance measures.

Several community based organizations and regional non-profits in the San Francisco Bay Area, who have been active participants in the Regional Prosperity Consortium, already have developed centers of technical expertise, datasets, and analytic tools related to his project and are eager to contribute to and influence positive outcomes and momentum.

Key Constraints to Take Into Account

All of the target data is public information. However, only a *subset* of it must be published by the local jurisdiction in a compiled form, which is typically a submittal to a state or federal agency on an *annual* basis. The information that is required is typically generated by two or more different departments within a city/county (e.g., planning, housing, building, public works) that may or may not be using the same software system or same data definitions, in part because of differing requirements of each separate report recipient (e.g., HUD, State Dept of Finance, HCD). Historically many of these reports are incomplete, inaccurate or missing altogether. Appendix D includes a list of required reports.

To acquire relatively *current* information, in all but a few jurisdictions, data *must be collected on an ad hoc basis*, in person by attending public meetings or reading packets and minutes of meetings on-line, or communicating directly with busy city/county staff.

As a fortunate exception to this pattern, and a growing trend, some jurisdictions are now posting development "pipeline" data on their websites, which may be downloaded and compiled remotely and integrated into a regional database semi-automatically. However, at present, few jurisdictions do this and the attributes reported and definitions of attributes differ across jurisdictions. Appendix E discusses this issue in more depth.

Another fortunate factor is that many jurisdictions—and all of them that post reports to the web—use one of several popular commercial off-the-shelf software applications for permit tracking, and therefore could adhere to standards for data definitions and reporting formats once/if such standards are negotiated and programmed. However, less than half of jurisdictions use specialized software and only 25% to 20% use one of the two most common applications. Appendix E also discusses this issue in more depth.

There is currently no strong requirement for compliance or a funding source for incentives to increase standardization and automation. Therefore, the proposed project must proceed as a voluntary effort to provide the tools and technical support that can make incremental progress.

Some desired development data is commercially available, but not for free. Consequently non-profit organizations as well as individuals and smaller entities (public, private or non-profit) are at a disadvantage in acquiring and using it for analysis, planning or advocacy, which becomes an equity issue.

Commercial data aggregators collect only some of the information of interest, and for the data they do collect they must ultimately rely on the same local staff contacts that generate data for required public reports. In order to assemble complete and timely datasets they must interpret ambiguous definitions and must interpolate and compensate for data flaws and gaps. However, they are not subject to requirements of method transparency or data verification.

Much of the desired public data is simply not collected by commercial data aggregators. For example, there is no commercially available data source for the affordability level of newly developed housing, nor for all building permits issued at the parcel/address level (only available in aggregate), nor for the development "pipeline" pre-permit. This gap is partially closed by other interested parties and advocates (e.g., labor, pro-housing, anti-growth) who systematically track a project's progress through the entitlement process. However, the data they collect is only about some projects of interest, not comprehensive, and is generally not made available except to allies. Also, once the advocacy campaign is completed, the data is discarded (or at least not shared in any systematic way) even though it would be of value for regional planning and progress tracking.

Finally, the challenges of finding resources and incentives for a comprehensive effort toward standardization, automation, and making use of willing advocates' data are compounded by the misperception that the task at hand is relatively easy (because narrowly scoped collection can indeed be relatively easy) and that commercial data sources are more timely and probably higher quality even though they must ultimately rely on the same primary sources.

Unresolved Issues Add Complexity

Adding to the complexity of the contextual "landscape" of this project are many challenging unresolved issues that working groups will need to address. The following examples range from project scope to resource allocation to data definitions to politics:

- How much resource should be invested in populating and verifying baseline information on existing land uses (as distinct from information about development in progress) in order to refine the acuity of modeling and establish baselines for trend analysis?
- How much resource should be invested in retroactively capturing recent development data in order to have comprehensive development data from 2010 baseline forward?
- How important is it to track both gross development and net development, the later taking into account what is demolished in the process of development?
- To what extent are state and/or regional agencies willing to participate in efforts to standardize data requirements, and to enforce adherence to data standards and deadlines, and/or provide incentives for compliance?

Proposed Approach — General Principles to Guide the Collaborative Project

The following guiding principles characterize a recommended general approach that takes into account key design considerations outlined in the previous section. The particular datasets to be assembled, as well as the charter of each of the working groups that would specify and assemble them, are detailed in subsequent sections of this memo.

1. *Open-data.* Create datasets that, to the greatest extent possible, are free and freely accessible to the public, can be viewed on-line in useful ways, and ultimately can be downloaded and/or used through a published application programming interface.

2. *Incremental Refinement Along Key Dimensions.* Starting with an adequate kernel in 2015, steadily extend and refine the system incrementally along these seven dimensions (order not indicative of priority): (1) automating collection, (2) covering more jurisdictions, (3) distinguishing more use and development types, (4) collecting more attributes and in more detail, (5) increasing accuracy through standardization and verification, (6) capturing more pre-permit data, and (7) increasing ease and utility of public access. This concept is elaborated in Appendix C: System Development Priorities and Sequence.

3. *Three Complementary Kernels*. Work from "three ends toward the middle" or "three kernels out," respectively: (1) semi-automated capture of parcel-level development data to be used in combination with data from other sources to improve resolution for land use modeling, and for which geographical coverage and official verification are lesser concerns because data aggregated at the tract or subarea level is adequate for modeling; (2) verified parcel-level data from all jurisdictions for housing development and affordable-housing development, (3) verified local land use policies.

4. *Collaborative Methods/Standards Process.* Develop methods and standards collaboratively. The many decisions necessary about which attributes to collect, attribute definitions, which jurisdictions to prioritize, and more will be the domain of several working groups and a coordinating committee. The groups will include regional agency staff, city/county staff, tracking software vendors, as well as data users and application developers from private and public-interest sectors. It is anticipated that several organizations that have been active in the Regional Prosperity Consortium will be active participants in this collaboration. Each working group will generate a work plan that is appropriate to the level of resource available. Working groups and their scope of work are described in more detail below, in Proposed Gross User/Developer Working Group Structure.

5. *Open Development of Data Visualization Tools and Other Applications*. Focus on data collection, quality and access rather than on data display. Encourage and support public, private and public-interest collaborators to develop applications. In order to hedge against undervaluation of the underlying system, promote or require that third party application developers give full credit and attribution of the data source and the collaborative effort that supports and extends it.

6. *Agency Commitment to Core Data Collection and Maintenance*. Rely on the regional public agencies, and ABAG in particular, to host the database(s), perform some basic data collection and cleaning of certain fundamental datasets (e.g., Regional Parcel Inventory, Local Geographical Subareas), and provide administrative support to assure continuity for the collaborative working groups. The precise scope of work will depend on level of available resource and confirmation of priorities with the collaborative. However, provisionally, this foundation work might include the following: on a quarterly cycle, ABAG staff would collect incremental development data from all jurisdictions from which data can be sourced on a semi-automated basis; on an annual cycle, ABAG staff would use jurisdictions' Annual Progress Reports and other sources to compile parcel-level data in order to update a report of verified housing production by affordability level by jurisdiction with breakouts for within-PDA and on-RHNA-inventory-sites; on a quarterly basis, ABAG staff would host a day of working group meetings. Additional work would proceed on an *ad hoc* basis using staffing resources recruited by the working groups from participants' agencies and organizations and other willing partners. ABAG staff would also assure linkage between the working groups and the Regional Interagency Modeling Group.

Proposed Gross Data Structure — What Collections of Data Will Be Assembled

There is an important distinction between the structure of the *data* and the structure of the *project*. The distinction is a matter of perspective, whether one focuses on the data itself or the

use of the data. The proposed datasets are described in the table below and illustrated in the graphic that follows. The proposed centers of work necessary to systematize collection, management and open access of/to these datasets are presented in the next section, Proposed Gross User/Developer Working Group Structure. The terms Telescope, Microscope and Stethoscope in the table below are links to that next section and will be explained there.

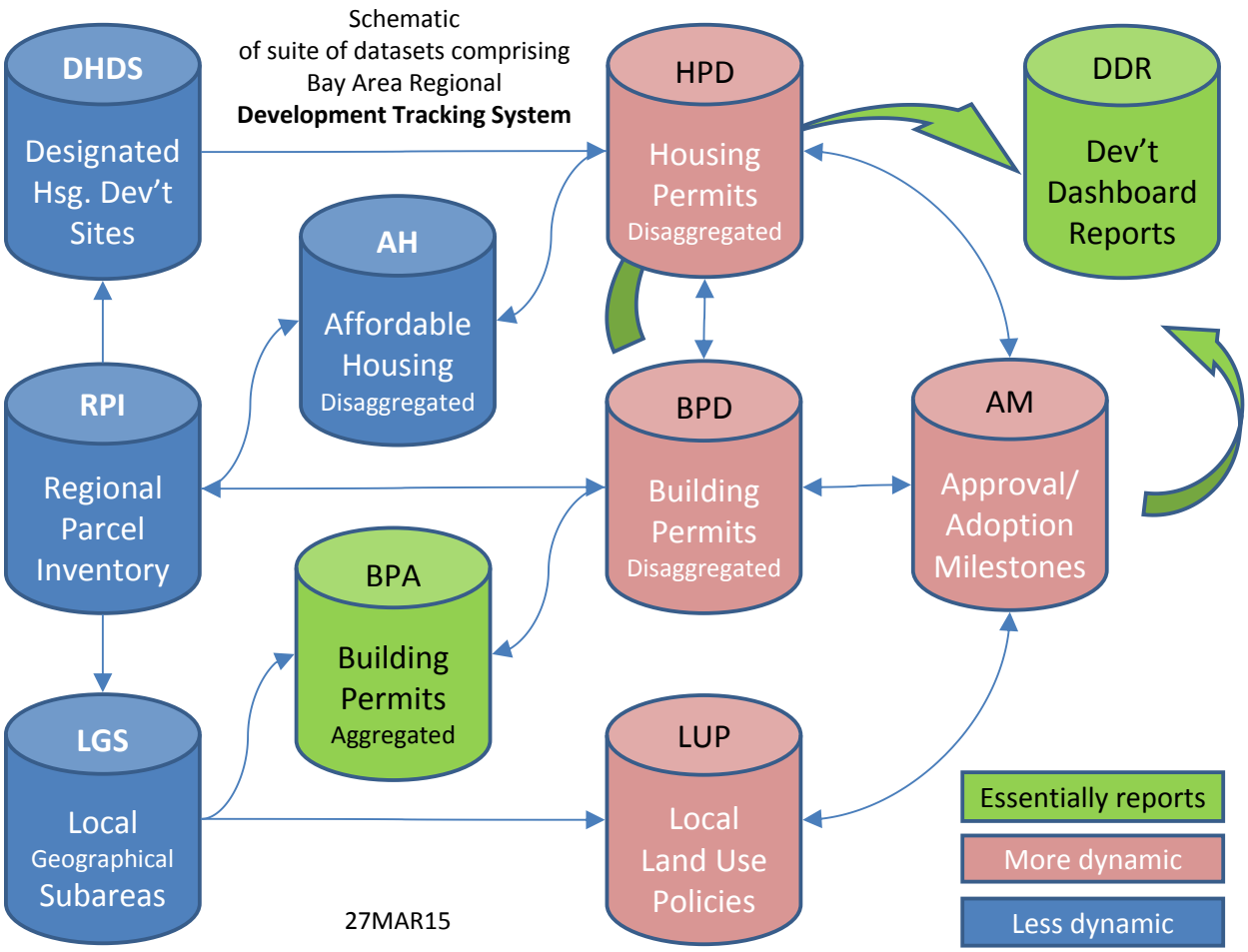
Caveats:

- Whether these datasets are maintained in separate databases or in separate tables within one database is an implementation detail that is outside the scope of this memo. Regardless, all of the datasets will have common elements, such as parcel number, that allow them to be linked and joined for searching, sorting, totaling, mapping, and so on.
- The question of how exactly to represent the distinction between parcels and multi-parcel sites, and the interplay of this distinction with project reporting and with geo-coding (which enables mapping and analysis in relation to points and areas of interest) is an important implementation detail that is also outside the scope of this memo.

Proposed Datasets		
Handle	**Unit of analysis**	**Description**
Regional Parcel Inventory	Parcel, disaggregated	All parcels in region (size, shape, location; also, as available, general plan designation, zoning, current land use). This existing dataset is fundamental, is maintained by ABAG, and is included here for context.
Local Geographical Subareas	Parcel, aggregated	GIS layers of boundaries for jurisdictions, census tracts, Priority Development Areas (PDAs), Transportation Analysis Zones (TAZs), Tansit Priority Areas (TPAs), etc. This existing dataset is fundamental, is maintained by ABAG, and is included here for context.
Building Permits Disaggregated (Telescope)	Parcel/project, disaggregated	Location (address or parcel number) and date of building permit issuance for all tracked development (i.e.,the subset of all development that we are able to track at a given time as we gradually improve the coverage of our system).
Building Permits Aggregated	Jurisdiction/ TAZ/ PDA, aggregated	Total development (residential and non-residential) each year in each jurisdiction and each PDA, compiled as the sum of all geo-coded development (collected in the

(Telescope)		Building Permits Disaggregated dataset described above) and non-geocoded aggregate residuals (i.e., development included in sub-area subtotals and/or jurisdiction totals for which we have not captured a precise location).
Housing Developments (Microscope)	Parcel/project, disaggregated	Location, date of completion , number of units and RHNA affordability level for all tracked housing development (i.e., for sites captured in dataset Building Permits Disaggregated above or reported in Annual Progress Reports to Cal HCD).
Designated Housing Sites (Microscope)	Parcel/project, disaggregated	Location (address or parcel number) of sites designated for housing development in local jurisdictions' Housing Elements.
Affordable Housing (Microscope)	Parcel/project, disaggregated	Location date of completion, number of units, affordability level and asset management entity for all designated (deed restricted) affordable housing.
Local Land Use Policies (Microscope)	Jurisdiction, aggregated	Significant local land use policies and specific plans by jurisdiction, and sub-area where applicable. Initially ordinance capture will be yes/no, and ultimately will link to primary source documents and indicators based on key provisions. Initially plan capture of zoning, specific plan and EIR readiness will be yes/no; ultimately more detailed.
Adoption/ Approval Milestones (Stethoscope)	Parcel/project, disaggregated	Dates upon which monitored development projects or local land use policies reached particular milestones, including pre-permit milestones. This dataset can hold milestone date-stamps collected for any project or policy at any of a dozen pipeline stages, not just building permit issuance date or policy adoption date.

Schematic of suite of datasets comprising a Bay Area Regional Development Tracking System


Schematic
of suite of datasets comprising
Bay Area Regional
Development Tracking System
DHDS
Designated Hsg. Dev't Sites
HPD
Housing Permits
Disaggregated
DDR
Dev't Dashboard Reports
AH
Affordable Housing
Disaggregated
RPI
Regional Parcel Inventory
BPD
Building Permits
Disaggregated
AM
Approval/ Adoption Milestones
BPA
Building Permits
Aggregated
LGS
Local
Geographical
Subareas
LUP
Local Land Use Policies
Essentially reports
More dynamic
Less dynamic
27MAR15

Proposed Gross User/Developer Working Group Structure

We propose that work proceed in working groups organized around four major aspects of the system's user/benefit functionality, coordinated by a steering committee. The proposed work structure is intended to favor broad participation and coordinated work through multiple projects that can proceed semi-independently as resources become available. Once built, a fully-featured open-data system could serve the combined interests of all of these stakeholders, and work to maintain and improve the system might be coalesced under a single working group. A fanciful lense-type name has been given to each of the working groups for ease of reference.

- *Telescope Working Group* would be the focal point for stakeholders who are primarily interested in aggregate information about development location, size, type, and pace at the jurisdiction level, or district level such as Priority Development Area or Transportation Analysis Zone (TAZ).
- *Microscope Working Group* would be the focal point for stakeholders who are primarily interested in detailed information about housing developments at the project level.
- *Stethoscope Working Group* would be the focal point for stakeholders who are primarily interested in detailed information about the progress of projects through the development stage "pipeline" process, or local land use policies through the adoption process.
- *Oscilloscope/Monitor Working Group* would be the focal point for stakeholders who are primarily interested in tools for visualization and analysis of the data collected.

Telescope. The ABAG/MTC Interagency Modeling Group must regularly update development data to refine the resolution of the UrbanSim land use model. Although the model can and does generate synthetic parcel-level "virtual development" data that reconciles to census tract and jurisdiction totals in aggregate, the model's accuracy and resolution can be increased by supplanting this synthetic data with real development data where and when available. Regional agency staff already collects project-specific data about the progress of major developments on a spot basis, often using local development reports available on some city/county websites. *The objective of the Telescope will be to capture building permit issuance for all development (residential and non-residential) using periodically updated parcel-level data available on city/county websites that can be collected through semi-automatic processes.* This primary data will be augmented by additional parcel-level data from ancillary sources (e.g., business journals and bulletins from interested parties). Aggregate development information for each geographical area of interest (e.g., census tract, TAZ, PDA, jurisdiction) will be represented as a composite of this "collected" parcel-level data and "calculated" information to

reconcile to area totals. *Emphasis will be on geographical coverage of jurisdictions with most readily accessible data and most substantial planned development; and on improving the quality of aggregate data (i.e., totals for geographical areas).*

Microscope. Periodically, ABAG must provide a definitive statement of each jurisdiction's progress toward RHNA targets, by affordability level, for use in allocation formulas for regional public funds. Although grant allocation cycles that use this information are currently multi-year, ABAG intends to update this statement annually. *The objective will be to capture accurate unit counts by affordability level for all new housing development, all incremental affordable housing (newly built or newly restricted), on all actual and planned housing development sites*, using annually updated parcel-level data from Annual Housing Reports submitted by jurisdictions to California HCD. This primary data will be augmented by data, as available, from the Telescope and Stethoscope work. *Emphasis will be on verified accuracy and project-scale detail of housing development*, with geographical focus on Priority Development Areas, especially in the inner bay corridors, and jurisdictions where there are large discrepancies among benchmark data sources (i.e., US Census, California HCD, Construction Industry Research Board or other popular commercial sources). Appendix A presents more detail about proposed datasets and the development project attributes collected in each.

Stethoscope. A major impetus for this project has been the desire for transparency into the development "pipeline." There has also been great interest in knowing how long it takes for projects to move through the pipeline, and to be able to compare among jurisdictions. To these ends, *the objective will be to capture the dates on which developments reach development milestones such as application submittal, zoning entitlements, and building permits. Emphasis will be on timeliness, verification, transparency, and especially for access by interested parties to develop applications that use this data.*

Monitor / Oscilloscope. Policy analysts in many agencies and organizations as well as freelance developers of open-data software applications are eager to use the datasets under development by Microscope, Telescope and Stethoscope working groups for many purposes. They will build applications that use tables, maps, graphics, and other functionality to represent information of interest; as will the regional agencies (ABAG, MTC, BCDC, BAAQMD) themselves. *The objective will be to assure that the system developed supports third-party use and application development. Emphasis is on usefulness, accessibility and method/source transparency of data.*

Proposed Development Sequencing and 2015 Priorities

The table below outlines proposed tasks for 2015 that are potentially feasible given current level of available resource. Developing an actual workplan for 2015 and beyond would be within the purview of the Steering Committee. The pace of progress would depend on level of available resource identified by theSteering Committee from among stakeholders.

Appendix C presents draft sequencing of important milestones.

Dataset	**Proposed Tasks for 2015**
Overall	1.1.1 Assemble working groups and Steering Committee. 1.1.2 Scope staffing for institutional support for working groups, and identify sufficient specific staff resources and funding to accomplish 2015 scope. 1.1.3 Steering Committee adopt an overall workplan for FY2015-2016 and oversee the development of workplans for each working group. 1.1.4 Establish a framework for progress measurement related to geographical coverage, data accuracy, level of detail, degree of public accessibility, and extent of collaboration; and oversee adoption of performance metrics and targets by each working group. 1.1.5 Write an annual progress report and circulate to all working group members and other stakeholders identified by the Steering Committee.
Building Permits Disaggregated (Telescope)	2.1.1 Assemble a subcommittee of the Telescope Working Group. 2.1.2 Adopt a workplan to, at minimum: (1) convene the standards process described below in next bullet, (2) coordinate closely with Microscope Working Group on standards for automating housing development tracking, and (3) adopt metrics by which implementation progress can be measured. 2.1.3 Convene a process that involves vendors and city/county users of one or two of the most-used tracking-software applications to develop data standards and procedures that will enable participating pilot jurisdictions to post quarterly development pipeline reports.
Building Permits Aggregated (Telescope)	2.2.1 Assemble a subcommittee of the Telescope Working Group to focus on this dataset. 2.2.2 Adopt a dataschema based on the constructs that (a) total residential and non-residential development in any particular geographical area is the sum of location-specific development data collected into the Building Permits Disaggregated dataset (see above) plus residual additional development at

	locations for which we have not captured parcel-level locations; and (b) each jurisdiction composed of a set of PDAs plus residual non-PDA space. 2.2.3 Adopt a workplan to, at minimum: (1) enable the 2014 housing production progress report (by jurisdiction and PDA vs non-PDA) to be generated from a database that uses the "location-known + residuals" dataschema described above, and (2) adopt metrics by which implementation progress can be measured. 2.2.4 Convene a process that involves vendors and city/county users of one or two of the top two tracking-software applications to develop data standards and procedures that will enable participating pilot jurisdictions to automatically generate one of the five most pervasively required state and federal reports.
Housing Developments (Microscope)	3.1.1 Assemble Microscope Working Group. 3.1.2 Adopt a workplan to, at minimum: (1) systematize annual summer compilation of annual regional progress report, (2) identify all required reports and all attributes in these reports, (3) design training program to improve quality of data input to jurisdictions' APRs, and (4) adopt metrics by which implementation progress can be measured. 3.1.3 Gain commitment from Cal HCD to cooperate on elements of a joint workplan to improve data quality (e.g., data definitions, input tools, jurisdiction training, automated upload of compliant local datasets). 3.1.4 Coordinate closely with data standards / tracking software subcommittee (see Telescope above) 3.1.5 Publish dataset comprising all housing development permits issued in 2014 in all jurisdictions, with all attributes required by Cal HCD for Annual Progress Reports; accessible by downloading spreadsheet and/or downloading shapefiles linked to site attributes.
Designated Housing Sites (Microscope)	3.2.1 Publish the dataset comprising all sites, all jurisdictions, all attributes required by Cal HCD for both 2007-1014 and 2014-2022 RHNA cycles; accessible by downloading spreadsheet, downloading pdf file, downloading shapefiles linked to site attributes, and accessing dynamically through Google Maps or equivalent.
Affordable Housing (Microscope)	3.3.1 Assemble a subcommittee of the Microscope Working Group to focus on this dataset. 3.3.2 Complete a workplan for (1) defining phased extension of types of affordable housing covered and available data sources for each, (2) phased addition of attributes, and (3) adopt metrics by which implementation

	progress can be measured.
Local Land Use Policies (Microscope)	3.4.1 Assemble a subcommittee of the Microscope Working Group to focus on this dataset. 3.4.2 Publish an inventory, by jurisdiction, of local adoption of any of the 30 most common local land use policies; accessible by downloading a spreadsheet, or shapefiles linked to site attributes. 3.4.3 Complete a workplan for: (1) adding indicators or PDA entitlement readiness to the dataset, (2) adding map-able "strength of ordinance" or other key-provision indicators to the dataset, (3) linking to adoption milestone database (see Stethoscope), and (4) adopt metrics by which implementation progress can be measured.
Adoption/ Approval Milestones (Stethoscope)	3.5.1 Assemble Stethoscope Working Group. 3.5.2 Adopt a workplan to, at minimum: (1) confirm development stage definitions, (2) adapt dataschemas to be able to hold milestone attributes if/as they are collected, (3) enter current milestone data for projects in pilot categories (see next bullet) , and (4) adopt metrics by which implementation progress can be measured. 3.5.3 Identify all affordable housing developments in the development pipeline in PDAs in the East Bay Corridors and Grand Boulevard Corridor, and collect current and historic milestone data for all of them.

Appendices

- Appendix A: Datasets and Attributes in each Dataset (i.e., database fields, table columns)
- Appendix B: Description of San Francisco's Housing Development Tracking System
- Appendix C: System Development Priorities and Sequence by Dataset
- Appendix D: Data Required by Commonly Applicable State and Federal Reports
- Appendix E: Feasibility of Generating a Regionwide Development Activity Database
- Appendix F: Acknowledgements
- Appendix G: Linking Memo to Contract Requirements

Appendix A: Datasets and Attributes (aka data fields, table columns) in Each Dataset

Negotiating a consensus in the collaborative working groups as to precisely what data to collect initially, and ultimately, will require time and diligent effort. Therefore, the material presented in this appendix should be viewed as a starting place, not as a proposed resolution. Many factors impinge on setting priorities for which attributes to collect, and at what level of detail. One influential factor will be which attributes are required for filing mandatory reports to state and federal agencies. A list of the state and federal reports required of most jurisdictions is included as Appendix D.

To illustrate the complexity of the data definition task facing working groups, Table A-1 below presents an analysis of the "development stage" or "project status" attribute. The project would seek to create a very fine-grained common-reference list of possible stages so that each stage captured by any jurisdiction could be mapped onto the common-reference. These would be clustered into fewer generic categories for ease of reporting and comparison, without loss of local detail. Two designations, one at each end of the completion spectrum, require special explanation. The three sub-types of "built" are actually required in some federal reporting to address the eventuality that a development is built but no one has moved in yet, or that it is not yet fully occupied. The "rumored" designation is designed to capture journalistic or crowdsourced reports of local development activity that is not yet official.

Table A-2 presents a summary description of the seven (draft) proposed datasets described in the memo.

Table A-3 lists data attributes that are a preliminary proposal of the data to be collected for each datasets.

Table A-1: Project Status Field Options

A generic set of project status values should be agreed upon can be captured and reported at more coarse AND more fine granularity.
The Adoption/Approval Milestone dataset should reflect the date on which a project completed each development stage.

Coarse-grain Generic (6)	Medium-grain Generic (10)	Fine-grain Generic (15)	San Francisco Status Fields	Sunnyvale Planning Permit Status Fields	Concord Current Projects Report	Saratoga Permits Issued	Fremont Status Fields
Pre-Project		Rumored					
	Designated (HE)	Designated (HE)					
	Zoned	Zoned					
Project Proposed & Staff Review		Withdrawn					
	Proposed	Proposed	PL Filed	Pending Review	Proposed		
	Staff Review	Staff Review		Comments Provided			Prelim.Review Procedure
Public Review & Entitlement	Public Review	Approved by ARB					Open for Public Comment
		Approved by PC					
	Approved	Approved by CC	PL Approved	Approved	Approved		
	Entitled	Entitled					Entitlement Approved
Permitting	Permitted		BP Filed			Applied	Building Permit Review
			BP ISSUED			Issued	
			BP REINSTATED				
		Permitted	BP APPROVED			Finaled	
Construction	Construction	Construction	CONSTRUCTION		Under Construction		Under Construction
Built	Built	Built					
		Occupied					
		Fully Occupied					

Table A-2: Summary of (Draft) Proposed Datasets

Regional Parcel Inventory (disaggregated parcels)	**RPI**	**Fundamental**
All parcels in region (size, shape, location; also, as available, general plan designation, zoning, current land use)		
Local Geographical Subareas (aggregated parcels)	**LGS**	**Fundamental**
GIS layers of boundaries for jurisdictions, census tracts, PDAs, TAZs, TPAs, and more		
Building Permits D (disaggregated, parcel/project)	**BPD**	**Telescope Working Group**
Location (address or parcel number) and date of building permit issuance for all tracked development.		
Building Permits A (aggregated, jurisdictions/areas)	**BPA**	**Telescope Working Group**
Total development (residential and non-residential) each quarter in each jurisdiction and each PDA, compiled as the sum of all geo-coded, disaggregated development data (collected in the Building Permits Disaggregated dataset described above) and non-geocoded aggregate residuals (i.e., development included in sub-area subtotals and/or jurisdiction totals for which we have not captured a precise location).		
Housing Developments (disaggregated, parcel/project)	**HPD**	**Microscope Working Group**
Location (address or parcel number), date of completion , number of units and RHNA affordability level for all tracked housing development (i.e., for sites captured in dataset Building Permits Disaggregated above or reported in Annual Progress Reports to Cal HCD).		
Affordable Housing (disaggregated, parcel/project)	**AH**	**Microscope Working Group**
Location (address or parcel number), date of completion, number of units, affordability level and asset management entity for all designated (deed restricted) affordable housing, both new developments and existing (pre-2010).		
Designated Housing Sites (disaggregated, parcel/project)	**DHDS**	**Microscope Working Group**
Location of sites designated for housing development in local jurisdictions' Housing Elements, plus additional information required in the Annual Progress Report submitted to State HCD, most importantly, the planned housing capacity of each site.		
Local Land Use Policies (aggregated, jurisdiction)	**LUP**	**Microscope Working Group**
Significant local land use policies and specific plans by jurisdiction, and sub-area where applicable. Initially ordinance capture will be yes/no, and ultimately will link to primary source documents and indicators based on key provisions. Initially plan capture of zoning, specific plan and EIR readiness will be yes/no; ultimately more detailed.		
Adoption/Approval Milestones (disaggregated)	**AM**	**Stethoscope Working Group**
Dates upon which monitored development projects or local land use policies reached particular milestones. This dataset can hold milestone date-stamps collected for any project or policy at any of a dozen pipeline stages, not just building permit issuance date or policy adoption date. This dataset is a heap of "reports" or "sightings" as a policy or project achieves successive milestones in a formal approval process.		

Table A-3: Attributes by Dataset

Building Permits D (disaggregated, parcel/project)			BPD	Telescope Working Group
Location and date of building permit issuance for all tracked development. This dataset is conceptualized as an extension (e.g., collecting more attributes), upgrade (e.g., automating collection, standardizing data definitions) and expansion (i.e., cover more cities more frequently) of existing databases maintained by ABAG. In current (2015) form, information about each project's location, description and status are all in one dataset. Going forward, project description information would remain and would be linked to a Milestone database and a Global Parcel Inventory.				
*Fields marked in the "GPI" column are presented as part of this dataset, but probably would be implemented as part of a Global Parcel Inventory (GPI) that is referenced by multiple datasets.. In other words, as implemented, this dataset, which focuses on planned housing capacity of certain designated parcels, would not repeat generic information about each parcel such as location, size, zoning, but rather would be linked to it through a parcel ID number.				
Attribute Category	**Attribute**	**GPI***	**Notes**	
Geo Reference	JoinNumA	Link		
	Jurisdiction	✔	Lookup?	
	County	✔	Lookup?	
	PDA	✔	Geo-calculate?	
	Address	✔		
	APN	✔		
	ParcelProjectRelation		Proposed addition. Single-parcel vs multi-parcel site; index parcel vs adjunct parcel	
	Lat	✔	Lookup? Proposed addition	
	Long	✔	Lookup? Proposed addition	
	Neighborhood	✔	Geo-calculate?	
	CensusTract2010	✔	Lookup? Geo-calculate?	
	CensusBlock2010	✔	Lookup? Geo-calculate?	
Project Status	Devt Stage Status	Link	Or link to Milestone database	
	Devt Stage Status Date		Or link to Milestone database	
	Built		Yes/No. Could also be link to Milestone database	
	Year Built		Year complete, actual or estimated	

Project Description	Project Name	
	Res_Units	
	Sqft Total	
	Sqft_Res	
	Sqft_Non_Res	
	Sqft Comm	Proposed addition
	Sqft Insti	Proposed addition
	Sqft Indus	Proposed addition
	Proposed Zoning	
	Project Description	
	Action	New Construction / Demolition / Rehab
	Stories	Number of stories
	Agency Sponsor	Public agency sponsor, if any
	Res_Tenure	Rent or Own for residential units
	Res_Below_Market	Number of Below Market residential units (see HPD for affordability details)
	Res_Unit_Rent	Average rent for residential units initially
	Res_Unit_Price	Average price for residential units initially
	Res_Total Bedrooms	
	Res_BedroomsPerUnit	Calculated
	Res_AreaPerUnit	Calculated

Building Permits A (aggregated, jurisdictions/areas)		BPA	Telescope Working Group

Total development (residential and non-residential) each quarter in each jurisdiction and each PDA, compiled as the sum of all geo-coded, disaggregated development data (collected in the Building Permits Disaggregated dataset described above) and non-geocoded aggregate residuals (i.e., development included in sub-area subtotals and/or jurisdiction totals for which we have not captured a precise location).

This dataset is a collection of geographic sub-area / sub-total "pieces" from which summary reports can be generated easily without need of GIS engine calculations. It draws detailed parcel level data from the Building Permits Disaggregated database, it draws jurisdiction totals from other sources, then it calculates subtotals for residual pieces. For example, for a given jurisdiction that happened to have three PDAs, it would hold four sub-area sub-totals for housing units permitted in a quarter: one for each PDA plus one "residual" for housing units not known to be in a PDA.

Attribute Category	Attribute	Notes
	Area ID	
	AreaType	Is this for a jurisdiction, or for a sub-area. Which type of sub-area, PDA, TAZ, TPA, Residual
	Jurisdiction	
	Calendar Quarter	Quarterly time granularity supports both fiscal years and calendar year reportain
	Res_Units	
	Sqft Total	
	Sqft_Res	
	Sqft_Non_Res	
	Sqft Comm	Proposed addition
	Sqft Insti	Proposed addition
	Sqft Indus	Proposed addition

Housing Developments (disaggregated, parcel/project)			HPD	Microscope Working Group

Location (address or parcel number), date of completion, number of units and RHNA affordability level for all tracked housing development (i.e., for sites captured in dataset Building Permits Disaggregated above or reported in Annual Progress Reports to Cal HCD).

Much of this information is imported from or linked to the Global Parcel Inventory (GPI) or Building Permits Disaggregated (BPD) databases. This dataset comprises additional detail about housing developments. In other words, as implemented, this dataset is the *residential* subset of the building permits data for which additional detail has been collected, and which has been directly verified with jurisdictions.

Attribute Category	Attribute	Links	Notes
Geo Reference	JoinNumA	GPI	Link to other databases
	Jurisdiction	GPI	From Global Parcel Inventory
	County	GPI	From Global Parcel Inventory
	PDA	GPI	From Global Parcel Inventory
	Address	GPI	From Global Parcel Inventory
	APN	GPI	From Global Parcel Inventory
	ParcelProjectRelation	BPD	From Building Permits Disaggregated database (applies to multi-parcel sites)
Project Status	Date Permitted	AAM	From Milestone database
	Date Built	AAM	From Milestone database
	Verified by	AAM	From Milestone database
	Date Verified	AAM	From Milestone database
Project Description	Project Name	BPD	From Building Permits Disaggregated database
	Res_Units	BPD	From Building Permits Disaggregated database
	Sqft Total	BPD	From Building Permits Disaggregated database
	Sqft_Res	BPD	From Building Permits Disaggregated database
	Project Description	BPD	From Building Permits Disaggregated database
	Action	BPD	New Construction / Demolition / Rehab
	Stories	BPD	Number of stories
	Agency Sponsor	BPD	Public agency sponsor, if any
	Res_Tenure	BPD	Rent or Own for residential units
	Res_Below_Market	BPD	Number of Below Market residential units

Developer	Developer
Residents	Served population (e.g., general, general seniors, assisted living, mental disabilities)
BMR_MI	Number of BMR units at moderate-income level
BMR_LI	Number of BMR units at low-income level
BMR_VLI	Number of BMR units at very-low-income level
BMR_ELI	Number of BMR units at extremely-low-income level
Bdrm_Cong	Number of beds, if congregate facility
Bdrm_SRO	Number of SRO units
Bdrm_Studio	Number of studio units
Bdrm_1	Number of one bedroom units
Bdrm_2	Number of two bedroom units
Bdrm_3	Number of three bedroom units
Bdrm_4	Number of four or more bedroom units

Affordable Housing (disaggregated, parcel/project)			AH	Microscope Working Group
Location (address or parcel number), date of completion, number of units, affordability level and asset management entity for all designated (deed restricted) affordable housing, both new developments and existing (pre-2010).				
For new developments almost all of this information is imported from or linked to the Global Parcel Inventory (GPI) or Building Permits Disaggregated (BPD) or Housing Development Disaggregated (HDD) databases. This database is designed to also hold information collected about existing (pre-2010) affordable housing complexes. For both existing and new developments, this dataset adds some attribute feilds as well.				
Attribute Category	**Attribute**	**Links**	**Notes**	
Geo Reference	JoinNumA	GPI	Link to other databases	
	Jurisdiction	GPI	From Global Parcel Inventory	
	County	GPI	From Global Parcel Inventory	
	PDA	GPI	From Global Parcel Inventory	
	Address	GPI	From Global Parcel Inventory	
	APN	GPI	From Global Parcel Inventory	
Project Status	Year Built	AAM	From Milestone database	
Project Description	Project Name	BPD	From Building Permits Disaggregated database for new developments	
	Res_Units	BPD	From Building Permits Disaggregated database for new developments	
	Sqft Total	BPD	From Building Permits Disaggregated database for new developments	
	Sqft_Res	BPD	From Building Permits Disaggregated database for new developments	
	Project Description	BPD	From Building Permits Disaggregated database for new developments	
	Action	BPD	New Construction / Demolition / Rehab	
	Stories	BPD	Number of stories	
	Agency Sponsor	BPD	Public agency sponsor, if any	
	Res_Tenure	BPD	(Rent or Own)	
	Res_Below_Market	BPD	Number of Below Market residential units	
	Developer	BPD	Developer	
	Residents	HDD	Served population (e.g., general, general seniors, assisted living, mental disabilities)	
	BMR_MI	HDD	Number of BMR units at moderate-income level	
	BMR_LI	HDD	Number of BMR units at low-income level	
	BMR_VLI	HDD	Number of BMR units at very-low-income level	

BMR_ELI	HDD	Number of BMR units at extremely-low-income level
Bdrm_Cong	HDD	Number of beds, if congregate facility
Bdrm_SRO	HDD	Number of SRO units
Bdrm_Studio	HDD	Number of studio units
Bdrm_1	HDD	Number of one bedroom units
Bdrm_2	HDD	Number of two bedroom units
Bdrm_3	HDD	Number of three bedroom units
Bdrm_3	HDD	Number of four or more bedroom units
Website	HDD	Link to description of project maintained by current property owner or manager
Property Mgt		Entity responsible for property management
Property Mgt Contact		Contact information for purposes of refreshing data
Asset Mgt		Entity responsible for maintaining affordability covenants
Funding Sources		Could be link to table of sources; could have amounts from each source
Est_Expiration_Year		Estimated year in which affordability covenants expire
Expiring Use Link	Link	Link or Index to CHPC's database of expiring affordability covenents

Designated Housing Sites (disaggregated, parcel/project)			DHDS	Microscope Working Group

Location of sites designated for housing development in local jurisdictions' Housing Elements, plus additional information required in the Annual Progress Report submitted to State HCD, most importantly, the planned housing capacity of each site.

This dataset is conceptualized as a heap of "reports" or "sightings" as a policy or project achieves successive milestones in a formal approval process. It is linked to the other datasets that contain descriptive information about the actual policies and projects. Initially "reports" would be entered by project staff exclusively. Later, reports could be crowdsourced, while verification would still be done by project staff exclusively.

*Fields marked in the "Link" column are presented as part of this dataset, but probably would be implemented as part of a Global Parcel Inventory (GPI) that is referenced by multiple dataset. In other words, as implemented, this dataset, which focuses on planned housing capacity of certain designated parcels, would not repeat generic information about each parcel such as location, size, zoning, but rather would be linked to it through a parcel ID number.

Attribute Category	Attribute	Links	Notes
Location / ID	JoinNumA	Link	ABAG Global Parcel Inventory ID
	APN	GPI	
	County	GPI	
	Jursidiction	GPI	
	Address	GPI	
	Local Housing Element Site ID		
	Address Notes		
	Project ID		For multi-parcel sites
	ParcelProjectRelation		Proposed addition. Single-parcel vs multi-parcel site; index parcel vs adjunct parcel
	PDA	GPI	
Zoning	General Plan Designation	GPI	
	Zone Designation	GPI	
	Notes(Sub Zone, etc)	GPI	
Existing Conditions	Lot Size (Acres)	GPI	
	Square Feet	GPI	
	Existing Use	GPI	
	Existing Residential Units	GPI	
	etc.)	GPI	

Housing Capacity	Minimum Capacity (units)	
	Maximum Capacity (units)	
	Realistic Capacity (units)	
	Net Yield	Calculated
	Minimum Density (UPA)	Calculated
	Maximum Density (UPA	Calculated
	Realistic Density (UPA)	Calculated

Local Land Use Policies (aggregated, jurisdiction)			LUP	Microscope Working Group
Significant local land use policies and specific plans by jurisdiction, and sub-area where applicable. Initially ordinance capture will be yes/no, and ultimately will link to primary source documents and indicators based on key provisions. Initially plan capture of zoning, specific plan and EIR readiness will be yes/no; ultimately more detailed.				
This dataset is relatively self-contained, but does link to two other datasets. In particular, it links to a list of geographical areas and subareas so applicability can be mapped, and it links to the Adoption/Approval Milestone database for tracking its approval status.				
Attribute Category	**Attribute**	**Links**	**Notes**	
	Project-Policy_ID			
	AAM_ID	AAM	Link to Adoption/Approval Milestone database	
	Project-Policy_Name	AAM	Name as represented in Adoption/Approval Milestone database	
	Adoption Status	AAM	Is policy adopted or somewhere in the adoption process	
	Adoption Date	AAM	Date current version was adopted	
	Active		Is it active and in force	
	Effective Date		Date current revision became effective	
	Expiration Date		Date expired or superceded by amendment	
	Policy_Official Name		Official Name	
	Policy Type		Is this a land use plan (general, specific, etc) or policy (ordinance) or program	
	Policy Subtype		Categories that make sense in context such as "anti-displacement" or "specific plans"	
	Jurisdiction	BPA		
	Geography	BPA	Link to area/subarea list in Building Permits Aggregated. Area where ordinance applies.	
	Policy Description		Summary description	
	Key Provisions			
	Document Link	Link	Link to local ordinance, could be internal (in the database) or external	

Adoption/Approval Milestones (disaggregated)	AM	Stethoscope Working Group

Dates upon which monitored development projects or local land use policies reached particular milestones, including pre-permit milestones. This dataset can hold milestone date-stamps collected for any project or policy at any of a dozen pipeline stages, not just building permit issuance date or policy adoption date.

This dataset is conceptualized as a heap of "reports" or "sightings" as a policy or project achieves successive milestones in a formal approval process. It is linked to the other datasets that contain descriptive information about the actual policies and projects. Initially "reports" would be entered by project staff exclusively. Later, reports could be crowdsourced, while verification would still be done by project staff exclusively.

*Fields marked in the "Link" column are presented as part of this dataset, but probably would be implemented as residing in a separate table or dataset that is linked by reference. In other words, as implemented, this dataset, which focuses on milestone accomplishment for policies and projects does not really contain information like the name or description of each policy or project, but rather would be linked to it through a Project-Policy ID number.

Attribute Category	Attribute	Link	Notes
Identification	Project-Policy ID	Link	
	ProjectOrPolicy	✔	Is this a policy or a development project?
	Project-Policy Name	✔	
	Jurisdiction	✔	
Progress	Dev Stage Milestone		See appendix C-2 for more detail
	Date Accomplished		
Audit Trail	Reported by		Initially by project staff only, eventually crowdsourced.
	Date Reported		
	Documentation Type		What is the basis (evidence) this milestone is achieved.
	Verified by		
	Date Verified		

Appendix B: Description of San Francisco's Housing Development Tracking System

What follows is a copy of a recent quarterly Pipeline Report from City/County of San Francisco, available directly on-line as:

http://www.sf-planning.org/modules/showdocument.aspx?documentid=9338

The report may also be found on the Pipeline Report landing page at:

http://www.sf-planning.org/index.aspx?page=1691

Table of Contents

Contents

List of Tables

List of Figures

List of Maps

WHAT IS THE PIPELINE?

The San Francisco consolidated pipeline consists of development projects that would add residential units or commercial space, applications for which have been formally submitted to the Planning Department or the Department of Building Inspection (DBI). Pipeline projects encompass various stages of development: from applications filed to entitlements secured, building permits issued to projects under construction. The pipeline includes only those projects with a land use or building permit application. It does not include projects undergoing preliminary Planning Department project review or projections based on area plan analysis. When a project is issued a Certificate of Final Completion by DBI, it is taken out of the pipeline.

To filter inactive projects, the current pipeline only includes projects filed during the last five years, projects approved in the last four years (with the exception of large projects, which are kept for seven years), and projects for which construction has begun during the past three years.

Data sources for the pipeline are chiefly the project databases maintained by the Planning Department and the Department of Building Inspection, respectively, but data is also periodically obtained from the (now Successor Agency to the) San Francisco Redevelopment Agency. Affordable housing projects sponsored by the Mayor's Office of Housing figure in the pipeline database only after an application has been filed with either the Planning Department or the Department of Building Inspection.

The *Pipeline Report* measures housing production in terms of housing units. Non-residential development, on the other hand, is measured in terms of building square footage. Depending on the proposed development project, square footage can be added with new construction or expansion, reduced with demolition or alteration, or re-allocated with conversion to other uses. This report counts *net change*, or new space or units minus existing space lost through conversion or demolition.

Time Frame and Certainty of Development

As the pipeline spans the entire project development life cycle for small and large projects ranging from addition of an extra unit in the rear yard to multi-structure complexes of residential and commercial development needing environmental impact reports and transportation studies, it follows that the entitlement and ultimate actualization of some projects is several years and occasionally decades into the future, while some projects are abandoned altogether prior to receiving a permit or completion. The pipeline, then, represents a particular scenario that assumes that all proposed development projects are eventually entitled and all entitled development projects eventually built. In reality, this is not the case.

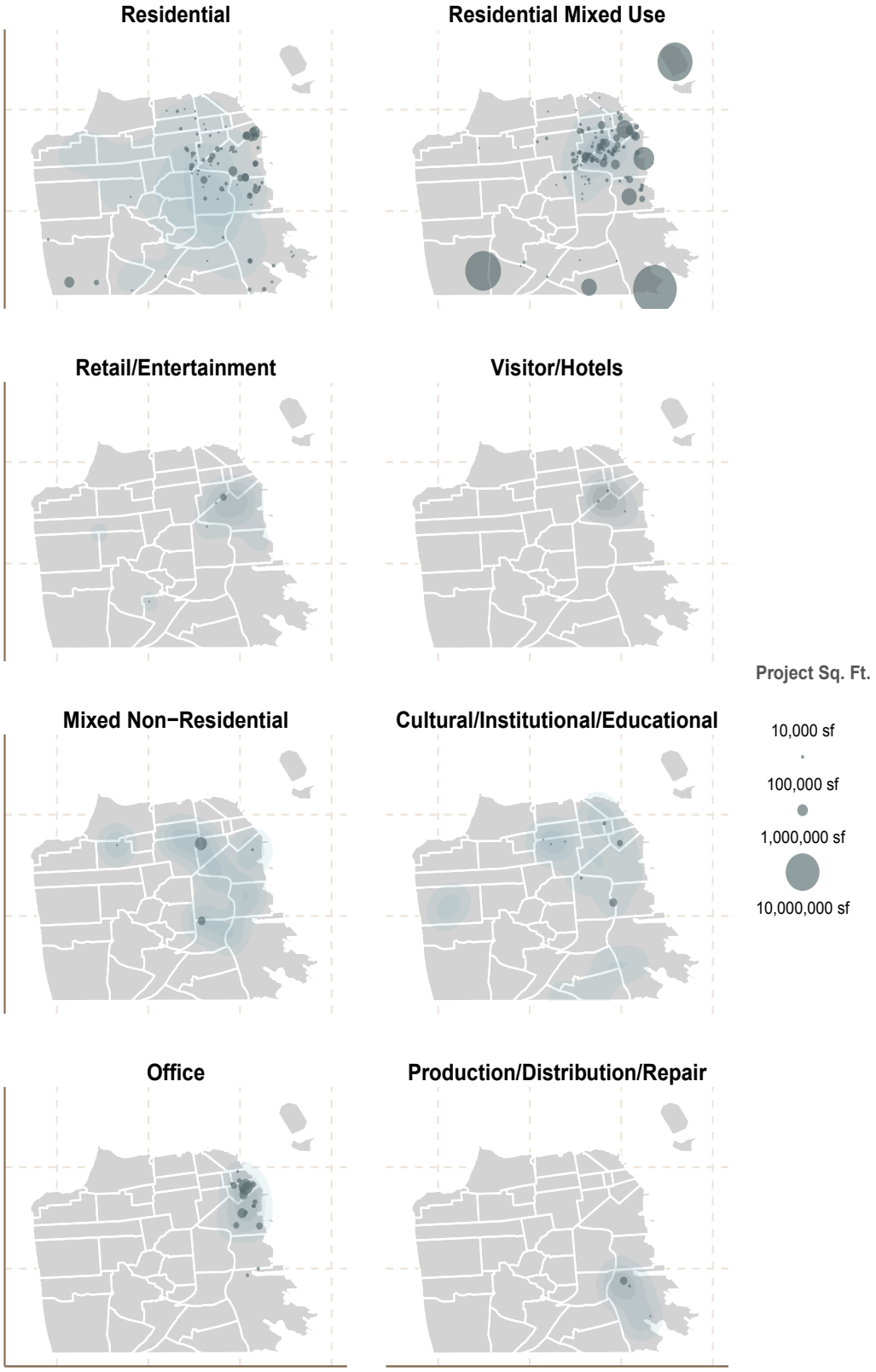
The Relevance of the Pipeline

The pipeline serves as a barometer of development trends in the medium to long term time horizon. It illustrates the location and scale of current and proposed future construction and reveals where new land uses are being established; it also records demolition and a partial listing of conversion of existing land uses. In sum, the pipeline provides a short- to medium-term picture of changing land uses, specifically tracking the changes to the city's housing stock and commercial uses. This report is meant to be a short overview.

Accuracy and Timeliness

The pipeline is compiled and consolidated from different data sources and is subject to errors due to varying accuracy and currency of original sources. The data in this report is pulled from original sources current through September 30, 2014. While we make an effort to consolidate multiple permits for different components of the same project from different agencies, it is not possible to validate the accuracy of all projects. Should you find inaccuracies and omissions, please e-mail your comments to *aksel.olsen@sfgov.org.*

Map 1. General Overview of Magnitude and Location of Development, by Major Land Use Type



THE DEVELOPMENT PIPELINE

There are currently 958 projects in the pipeline. Of these, 75 percent are exclusively residential and 17 percent are mixed-use projects with both residential and commercial components. Only 8 percent of projects are non-residential developments. Map 1 (left) gives the general location and magnitude of this development across is many stages.

A net total of 50,600 new housing units would be added to the city's housing stock according to current data. This is high relative to historical numbers and is largely due to the filing and entitlement of applications during the past five years for new large scale, long term development programs for Parkmerced, Treasure Island and the Bayview Waterfront. These projects, as well as their expected development over the course of decades must be kept in mind when considering the overall totals. The vast majority of pipeline projects, however, are small scale consisting of one to three units. The number of new projects slowed down during the Great Recession of 2007-2009 and beyond, but has since recovered in earnest as evidenced by both new project applications as well as the construction of projects with "older" entitlements. The "hot spot" for much of this development is Market Street at various sections of it. While this may seem a response to the recent acceleration of technology companies locating in the area, many development projects here predate the last recession, during which they were idle. As financing improved, many projects came back.

Projects by Overall Status

Table 1 breaks down projects, housing units and non-residential space by planning stage. First are non-entitled projects. A number of projects file building permit applications even as their projects have not cleared planning entitlements. The second major group include entitled projects; those which have completed the planning process and obtained necessary approvals. These are then divided into different stages of the building permiting process. Table 1 shows the following:

- Around 21 percent of all projects, representing 6,700 net added housing units and 5,400,000 sq ft of commercial space, are under construction.
- Around 20 percent of projects (with another 4,100 net units and 2,1 million sq ft of commercial space) have received building permit approvals. As of the time of writing, some may have moved to the construction phase.
- Around one in three projects (including 900 net new units and an net loss of 60,000 sq ft of commercial space) have filed building permit applications with the Department of Building Inspections. A small number of projects have filed applications but have yet to receive planning approvals.
- One in eight projects and 55 percent of the units and 37 percent of the non-residential space have received Planning Department approvals. These projects now must secure building permits.

Table 1. Residential and Commercial Pipeline, by Pipeline Status and Land Use Category

Entitlement Status	Status	Total No. of Projects	Net Housing Units	Net Comm'l Sq. Ft.	Net Commercial Gross Square Footage					
					CIE	Medical	Office	PDR	Retail	Visitor
Under Planning Review	Planning Filed	121	8,900	4,568,300	(20,200)	-	3,806,300	(444,400)	711,900	514,600
	BP Filed	329	3,100	634,900	913,000	-	141,800	(98,800)	(56,300)	(264,800)
	Total, Not Entitled	450	12,000	5,203,200	892,800	-	3,948,200	(543,200)	655,700	249,700
Approved by Planning	PL Approved	80	26,900	6,131,200	33,600	-	3,762,000	296,700	1,612,700	426,200
	BP Filed	30	900	(61,500)	(5,700)	-	(19,600)	(58,100)	38,200	(16,300)
	BP Approved/ Issued/ Re-Instated	183	4,100	2,055,500	111,300	20,000	1,325,700	7,900	463,400	127,200
	Construction	215	6,700	5,405,900	567,800	1,767,400	2,986,700	(8,100)	86,600	5,400
	Total, Entitled	508	38,600	13,531,100	707,000	1,787,400	8,054,900	238,400	2,201,000	542,500
	Total	**958**	**50,600**	**18,734,300**	**1,599,800**	**1,787,400**	**12,002,900**	**(304,800)**	**2,856,500**	**792,300**

Quarter 3, 2014

Subset of pipeline where project adds either more than 10 units or 10,000 GSF

CONSTRUCTION

Block Lot	Address	District	Net Comm'l sq ft	Net Units	Land Use	Largest Comm'l	Best date
3720001	TRANSBAY TOWER	TB Combo	1743k	0	MIPS	MIPS	6/25/2014
3746001	390 Main St	Rincon Hill	0k	669	Resident	--N/A--	9/30/2014
3735063	222 02ND ST	TB Combo	623k	0	MIPS	MIPS	9/29/2014
4154001	1001 POTRERO AV	Showpl/Potrero	419k	0	CIE	CIE	2/10/2014
3710017	350 MISSION ST	TB Combo	416k	0	MIPS	MIPS	9/23/2014
3747320	The Californian	Rincon Hill	-2k	452	Mixres	MIPS	8/7/2014
3738004	280 BEALE ST	TB Combo	0k	479	Resident	--N/A--	9/3/2014
0814020	100 VAN NESS AV	Market Octavia	-424k	399	Mixres	Retail/Ent	9/24/2014
3833002	1006 16TH ST	Showpl/Potrero	0k	393	Resident	--N/A--	9/30/2014
3749059	45 LANSING ST	Rincon Hill	-14k	320	Resident	--N/A--	9/18/2014
3765015	One Rincon Hill Phase II	Rincon Hill	0k	312	Resident	--N/A--	6/9/2014
3722367	151 THIRD ST	Downtown	67k		CIE	CIE	9/24/2014
3721122	535 MISSION ST	TB Combo	296k	0	MIPS	MIPS	9/12/2014
3732009	900 FOLSOM ST	East SoMa	6k	282	Mixres	Retail/Ent	8/8/2014
3701064	55 9TH ST	Downtown	0k	273	Resident	--N/A--	12/9/2013
8710007	718 LONG BRIDGE ST	Mission Bay	0k	263	Resident	--N/A--	9/30/2014
3717019	120 HOWARD ST	Downtown	67k	0	MIPS	MIPS	6/13/2012
0857001A	218 BUCHANAN ST	Market Octavia	0k	191	Resident	--N/A--	7/31/2014
3507039	1420 MISSION ST	Downtown	12k	190	Resident	--N/A--	9/10/2014
0831023	MARKET OCTAVIA - PARC	Market Octavia	4k	182	Mixres	Retail/Ent	9/29/2014
7331003	800 Brotherhood Way	Park Merced	0k	182	Resident	--N/A--	9/11/2012
3509043	104 9th Street	Downtown	-8k	160	Mixres	Retail/Ent	10/1/2014
0794028	555 FULTON ST	Market Octavia	0k	139	Resident	--N/A--	9/30/2014
0857001	55 Laguna Street	Market Octavia	28k	133	Mixres	Retail/Ent	7/31/2014
4991277	833-881 Jamestown	Candlestick	0k	14	Resident	--N/A--	9/17/2007
3510001	1415 MISSION ST	Downtown	0k	121	Mixres	MIPS	9/17/2014
5431A001	5800 03RD ST	BVHP Area A,B	13k	121	Resident	--N/A--	9/29/2014
0870003	100 BUCHANAN ST	Market Octavia	0k	116	Resident	--N/A--	9/30/2014

Appendix C: System Development Priorities and Sequence

This appendix presents an orderly progression for the incremental extension and refinement of the system. As outlined in the body of the memo (See Proposed Approach — General Principles to Guide the Collaborative Project), coordinated working groups will make incremental progress to upgrade the system along seven key dimensions: (1) automating collection, (2) covering more jurisdictions, (3) distinguishing more use and development types, (4) collecting more attributes and in more detail, (5) increasing accuracy through standardization and verification, (6) capturing more pre-permit data, and (7) increasing ease and utility of public access.

Also presented in the memo is a draft workplan for the first year of a collaborative effort (See Proposed Development Sequencing and 2015 Priorities). During the first year the premiere accomplishment would be to establish the collaborative structure and adopt a collective workplan. In a sense, any workplan with a horizon beyond the first year is speculative. However, with that in mind, and in order to convey a tangible sense of how this work could be accomplished incrementally in orderly stages, the matrix below presents a sketch of two feasible functionality upgrade plateaus for each of the major datasets. The column entitled Proposed Tasks for 2015 merely restates the table presented in the body of the memo.

Appendix C: System Development Priorities and Sequence for Collbotative Development of a Development Tracking System

Dataset	Baseline	Proposed Tasks for 2015	First Major Functionality Plateau	Eventually
Institutional Infrastructure *Steering Committee* This row is not about a dataset, but about the collaborative structure and process. Development sequence for each major dataset is outlined on the following pages.	There is no organized process or structure. Many jurisdictions have open-data initiatives. There is an emerging regional open-data practice, with several jurisdictions leading the way, notably San Francisco. State HCD is actively working on improving data quality and access. An informal working group of regional NGOs and CBOs has started, most of them participants in the Regional Prosperity Consortium.	1.1.1 Assemble working groups and Steering Committee. 1.1.2 Scope staffing for institutional support for working groups, and identify sufficient specific staff resources and funding to accomplish 2015 scope. 1.1.3 Steering Committee adopt an overall workplan for FY2015-2016 and oversee the development of workplans for each working group. 1.1.4 Establish a framework for progress measurement related to geographical coverage, data accuracy, level of detail, degree of public accessibility, and extent of collaboration; and oversee adoption of performance metrics and targets by each working group. 1.1.5 Write an annual progress report and circulate to all working group members and other stakeholders identified by the Steering Committee.	The start-up objectives are listed in Proposed Tasks for 2015. One additional First Plateau objective is to establish a sustainable institutional infrastructure for oversight, and evolution of the systems, tools, standards and user-group relations.	Institutional infrastructure for sustainable collaborative established.

Dataset	Baseline	Proposed Tasks for 2015	First Major Functionality Plateau	Eventually
Building Permits Disaggregated **BPD** *Telescope Working Group* Location (address or parcel number) and date of building permit issuance for all tracked development.	Currently many of the pieces exist, but in multiple semi-connected databases and spreadsheets. However, almost none of the standards and protocols across jurisdictions have been established. Due to resource constraints, data collection has been episodic and project-driven	2.1.1 Assemble a subcommittee of the Telescope Working Group. 2.1.2 Adopt a workplan to, at minimum: (1) convene the standards process described below in next bullet, (2) coordinate closely with Microscope Working Group on standards for automating housing development tracking, and (3) adopt metrics by which implementation progress can be measured. 2.1.3 Convene a process that involves vendors and city/county users of one or two of the most-used tracking-software applications to develop data standards and procedures that will enable participating pilot jurisdictions to post quarterly development pipeline reports.	Implement the data structures, standards and protocols to launch with partial data. Expand data collection later. Launch a basic system when data from three sources is available systematically and sustainably: - quarterly semi-automated import of new permits from local open-data sources in a pilot group that includes at least three cities. - link to parcel-level housing development data in HPD dataset - quarterly manual update of "major developments" in Top 15 Growth cities, and all jurisdictions in the East Bay Corridors Initiative	Over time, increase the sophistication and ease of use. Set and meet goals for % of total development that is captured in BPD. Put priority on capture of development in PDAs. Suggest edtargets: 80% of all development; 100% of PDA development; 100% pre-permit tracking for major developments in PDAs.

Dataset	Baseline	Proposed Tasks for 2015	First Major Functionality Plateau	Eventually
Building Permits Aggregated **BPA** *Telescope Working Group* Total development (residential and non-residential) each quarter in each jurisdiction and each PDA, compiled as the sum of all geo-coded, disaggregated development data (collected in the Building Permits Disaggregated dataset described above) and non-geocoded aggregate residuals (i.e., development included in sub-area subtotals and/or jurisdiction totals for which we have not captured a precise location).	Currently the pieces exist, but in multiple unconnected spreadsheets. We have subareas, we have totals, we have some information at the parcel level and some at the district level, but not integrated. The concept of total = itemized + residual is not consistently implemented.	2.2.1 Assemble a subcommittee of the Telescope Working Group to focus on this dataset. 2.2.2 Adopt a dataschema based on the constructs that (a) total residential and non-residential development in any particular geographical area is the sum of location-specific development data collected into the Building Permits Disaggregated dataset (see above) plus residual additional development at locations for which we have not captured parcel-level locations; and (b) that each jurisdiction composed of a set of PDAs plus residual non-PDA space. 2.2.3 Adopt a workplan to, at minimum: (1) enable the 2014 housing production progress report (by jurisdiction and PDA vs non-PDA) to be generated from a database that uses the "location-known + residuals" dataschema described above, and (2) adopt progress metrics. 2.2.4: Pilot to help local jxns generate required local state/fed reports automatically. See next column.	Implement the data structure, even with partial data. Make dataset downloadable; generate one downloadable summary report as a template and example. 2.2.3: Generate the 2014 Housing Production Report (released in 2015) from this database / report generator. 2.2.4: Convene a process that involves vendors and city/county users of one or two of the top two tracking-software applications to develop data standards and procedures that will enable participating pilot jurisdictions to automatically generate one of the five most pervasively required state and federal reports.	Over time increase the sophistication and ease of use. Set and meet goals for % of total development that is generated from BPD; diminish aggregate residuals.

Dataset	Baseline	Proposed Tasks for 2015	First Major Functionality Plateau	Eventually
Housing Developments **HPD** *Microscope Working Group* Location (address or parcel number), date of completion, number of units and RHNA affordability level for all tracked housing development (i.e., for sites captured in dataset Building Permits Disaggregated above or reported in Annual Progress Reports to Cal HCD).	Currently data for most jxns is hand-converted from HCD's pdf APRs annually with 12-24 month lag.	3.1.1 Assemble Microscope Working Group. 3.1.2 Adopt a workplan to, at minimum: (1) systematize annual summer compilation of annual regional progress report, (2) identify all required reports and all attributes in these reports, (3) design training program to improve quality of data input to jurisdictions' APRs, and (4) adopt progress metrics. 3.1.3 Gain commitment from Cal HCD to cooperate on elements of a joint workplan to improve data quality (e.g., data definitions, input tools, jurisdiction training, automated upload of compliant local datasets). 3.1.4 Coordinate closely with data standards / tracking software subcommittee (see Telescope above) 3.1.5 Publish dataset comprising all housing development permits issued in 2014 in all jurisdictions, with all attributes required by Cal HCD for Annual Progress Reports; accessible by downloading spreadsheet and/or downloading shapefiles linked to site attributes.	Data from semi-automated jxns (see coverage goals above) would be imported directly from BPD and verified quarterly with 3-6 month lag. Data from APRs would be received for 90% of jxns in some importable format, and lag would be reduced to 6-18 months Data downloadable, basic canned summary reports downloadable. At least one dynamically generated info graphic published by a third-party app developer.	Eventually get semi-automated data. Data from 50% of jxns that account for 80% of housing development is generated semi-automtically through BPD. Access: Numerous info-graphics, including "dashboards"

Dataset	Baseline	Proposed Tasks for 2015	First Major Functionality Plateau	Eventually
Affordable Housing **AH** *Microscope Working Group* Location (address or parcel number), date of completion, number of units, affordability level, asset management entity, and other attributes for all designated (deed restricted) affordable housing, both new developments and existing (pre-2010).	Several organizations in the Regional Prosperity Consortium maintain databases that are relatively current and complete, but for only a subset of all "species" of deed-restricted affordable housing. No existing set is open, comprehensive and well-integrated with other key datasets.	3.3.1 Assemble a subcommittee of the Microscope Working Group to focus on this dataset. 3.3.2 Complete a workplan for (1) defining phased extension of types of affordable housing covered and available data sources for each, (2) phased addition of attributes, and (3) adopt metrics progress metrics.	Get basic dataset up and accessible. Support access via downloading spreadsheet, downloading summary reports in spreadsheet and pdf formats, downloading shapefiles linked to site attributes, and accessing dynamically through Google Maps or equivalent. Establish protocols and resources for ongoing collection and maintenance.	Expand to include all units that have received federal or state funding; or exemption from local property tax; or (by jurisdiction) tenant-based rental assistance. Info-graphic of AH as % of housing by jurisdiction.
Designated Housing Sites **DHDS** *Microscope Working Group* Location of sites designated for housing development in local jurisdictions' Housing Elements, plus additional information required in the Annual Progress Report submitted to State HCD, most importantly, the planned housing capacity of each site.	Dataset is relatively complete for 2007-2014 RHNA cycle, but with limitations: - It is stand-alone, not integrated with RPI, BPD or AM datasets. - Ambiguous data definitions compromise data quality. - Limited support for info-graphics and basing maps on parcel shapes is problematic.	3.2.1 Publish the dataset comprising all sites, all jurisdictions, all attributes required by Cal HCD for both 2007-1014 and 2014-2022 RHNA cycles; accessible by downloading spreadsheet, downloading pdf file, downloading shapefiles linked to site attributes, and accessing dynamically through Google Maps or equivalent.	Complete collection and cleaning for 2007-2014 RHNA cycle. Link DHDS to RPI, BPD, BPA, AM. Because this dataset is relatively static with relatively few attributes it can serve as a vehicle for enabling access and third-party applications, so make it available in the various ways listed in 2015 task 3.2.1.	Add 2014-2022 RHNA cycle. Publish paper comparing the two cycles, and inviting other analyses. Publish paper analyzing extent to which housing development occurs on designated housing sites versus other sites.

Dataset	Baseline	Proposed Tasks for 2015	First Major Functionality Plateau	Eventually
Local Land Use Policies **LUP** *Microscope Working Group* Significant local land use policies and specific plans by jurisdiction, and sub-area where applicable. Initially ordinance capture will be yes/no, and ultimately will link to primary source documents and indicators based on key provisions. Initially plan capture of zoning, specific plan and EIR readiness will be yes/no; ultimately more detailed.	Existing dataset has yes/no for top 30 policies by jurisdiction. Limitations: No sub-area breakout. Policies but not plans (sometimes policies are codified in the plans), for instance, can't map "entitlement readiness." No breakout of ordinance details to enable comparative analysis across jurisdictions. No platform to capture commentary.	3.4.1 Assemble a subcommittee of the Microscope Working Group to focus on this dataset. 3.4.2 Publish an inventory, by jurisdiction, of local adoption of any of the 30 most common local land use policies; accessible by downloading a spreadsheet, or shapefiles linked to site attributes. 3.4.3 Complete a workplan for: (1) adding indicators or PDA entitlement readiness to the dataset, (2) adding map-able "strength of ordinance" or other key-provision indicators to the dataset, (3) linking to adoption milestone database (see Stethoscope), and (4) adopt progress metrics.	Link to subarea geographies schema underlying the BPA dataset. Add area-plan status data & link to PDA Showcase website. Enable collection of detail-level analysis (i.e., salient features and points of comparison). Link to Milestone (AM) database.	Release a platform to enable third-party apps for : - Crowd-sourcing apparent accomplishment of milestones. - Private or public commentary and networking about policies and plans. Info-graphics related to policy coverage and comparison.

Dataset	Baseline	Proposed Tasks for 2015	First Major Functionality Plateau	Eventually
Adoption/ Approval Milestones **AM** Stethoscope Working Group Dates upon which monitored development projects or local land use policies reached particular milestones, including pre-permit milestones. This dataset can hold milestone date-stamps collected for any project or policy at any of a dozen pipeline stages, not just building permit issuance date or policy adoption date. This dataset is a heap of "reports" or "sightings" as a policy or project achieves successive milestones in a formal approval process.	This functionality, which was a core motivation for this "development dashboard" study is not implemented in any form currently at the regional level. No systematic tracking of pre-permit approval process. Some existing reports report completion while others report issuance of permits. No systematic capture of pace of entitlement & permitting process.	3.5.1 Assemble Stethoscope Working Group. 3.5.2 Adopt a workplan to, at minimum: (1) confirm development stage definitions, (2) adapt dataschemas to be able to hold milestone attributes if/as they are collected, (3) enter current milestone data for projects in pilot categories (see next bullet) , and (4) adopt metrics by which implementation progress can be measured. 3.5.3 Identify all affordable housing developments in the development pipeline in PDAs in the East Bay Corridors and Grand Boulevard Corridor, and collect current and historic milestone data for all of them.	Implement the data structure, even with partial data. Launch with whatever geographical scope is currently supported in the BPD dataset.* Make dataset downloadable; generate one downloadable summary report as a template and example. Create one compelling info-graphic to attract users and application developers. *Initially data would include the data in BPD dataset, which is harvested semi-automatically from jurisdictions that publish relatively comprehensive pipeline data on the web at least quarterly (three to five large jurisdictions), plus all the manually collected data for housing developments in the HPD and AH datasets.	Release a platform to enable third-party apps for : - Crowd-sourcing apparent accomplishment of milestones. - Private or public commentary and networking about development policies and projects in the pipeline. Info-graphics on development location, type and pace.

Appendix D: Commonly Applicable State and Federal Annual Reports

The chart on the next page lists mandatory annual state and federal reports related to residential and non-residential development, applicable to many or all cities and counties.

The chart also lists some additional development reports that are typically required by internal city/county processes, or are routinely provided by municipal staff to requesting special interests or commercial data aggregators as a proactive alternative to having to respond to formal Public Information Act requests, which would be more time consuming and expensive.

What is apparent from the chart below is that there are many reporting requirements for development data. What is also true, but not readily apparent is that there is little standardization of data definitions across these reports, and that these reports are typically generated by different persons in different departments within a given jurisdiction.

Goal: A goal of the overall project described in this memo should be to develop more unified and streamlined housing production reporting definitions and collection practices so that cities, counties, and housing authorities (a) can capture data in a manner that is less demanding of staff time, (b) could generate all currently mandated reports related to housing production from one integrated process, and (c) can yield data that is more accurate, timely and accessible.

Project: Toward that end, a valuable project that could be launched by the Microscope Working Group would be to work with city/county staff, State HCD staff and vendors of tracking software, to standardize data definitions for the state-mandated Annual Progress Report, and generate input for the most common mandatory reports.

Project: Another valuable project, which could be launched by the Telescope Working Group, would be to work with representatives of the most-widely-used development tracking software applications, city/county staff, and regional and subregional agency staff to agree on standards for quarterly web posting of development reports. Although only possible in jurisdictions with relatively robust data integration and open-data practices in place, the acceleration of reporting to a quarterly basis from lagged annual timing (August regional compilation of April reports for previous calendar year) would be a major step forward.

Chart is positioned on next page to avoid page break.

Report	Required by	Purpose	Notes
Mandatory Federal			
C-404	US Dept. of Interior	Census	Sent to 9000 larger jurisdictions nationwide annually, and to 12000 smaller every five years.
Comprehensive Annual Performance Evaluation Report (CAPER)	US Housing & Urban Dev.	Performance on 5-year Consolidated Plan goals (Fair Housing, Housing programs/ projects	Jurisdictions entitled to receive CDBG funds directly; the 33 largest jurisdictions.
Mandatory State			
DOF Housing Unit Survey	Cal. Dept. of Finance	Growth and populations trends	Annual page length form
Annual Progress Report (APR)	Cal. Housing & Comm. Dev.	Performance on Housing Element goals	Annual report of number of units completed. Includes location, affordability level, tenure.
Strong Motion Instrumentation and Seismic Hazard Mapping Fee Report	California Department of Conservation	Collect fees for seismic mapping.	Quarterly a simple one-page form is mailed in to State with a fee. Reports number and valuation of permits for certain categories of buildings
Jurisdiction Internal			
Building Permit Data	County Assessor	Property tax assessment	Immediately as each building permit is issued a one page of a triplicate form is sent to Assessor. Many jurisdictions are automating.
Annual Performance Report	City/County Manager/ Council/Board	Internal progress reports	Each jurisdiction does this differently, but most jurisdictions generate internal progress reports as part of their annual budgeting and goal setting process.
External Stakeholder Requests			
Locally-generated Permitting Reports (CIRB form if needed)	Construction Industry Review Board	Data subscription service Policy analysis. Advocacy campaign support.	CIRB tracks "housing starts." Vendor polls for and expedites completion of a relatively simple form.
Housing Permits	Affordable-housing and fair-housing advocacy organizations	Policy analysis. Advocacy campaign support.	Because the official Annual Progress Reports to the State are time-lagged, are not in a readily usable format, and may be incomplete, advocates often poll cities and counties directly to receive parallel reports or refinements of reports when standard State data definitions are insufficient for their purposes.

Appendix E: Feasibility of Generating a Regionwide Development Activity Database

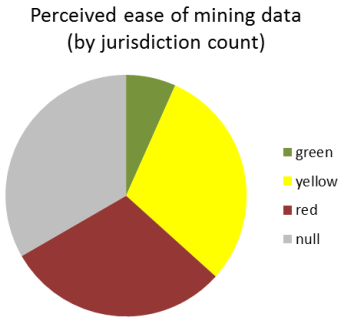
Following is an internal ABAG memo that summarizes findings about the relative accessibility of data through jurisdictions' websites, the types and prevalence of development tracking software applications in use by cities and counties, the degree of standardization across jurisdictions and other technical issues relevant to implementation of a regional development tracking database.

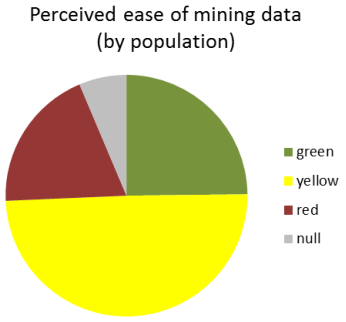
Subject: The Feasibility of Generating a Region-wide Development Activity Database

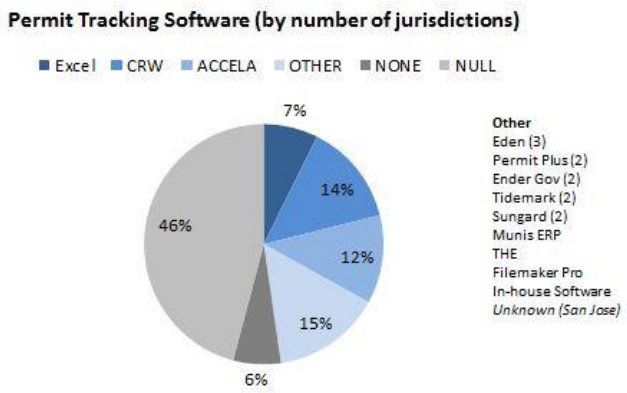
Section I: Assessing Feasibility of Data Mining Development Activity from Municipal Websites

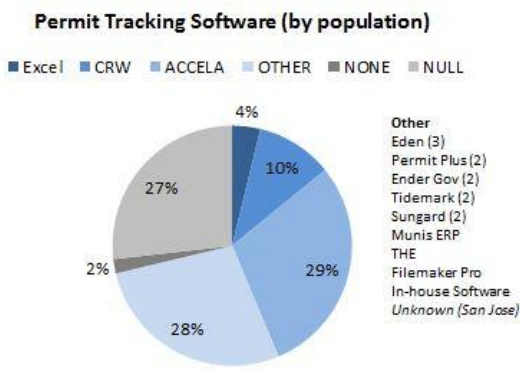
Summary

Thirty jurisdiction websites were explored to gain a high level understanding of the difficulty in directly mining their website for permitted development data. The 15 most populous cities and a cross sample of 15 jurisdictions ranging in population from 75,000 to 2,000 were used to develop an understanding of the scale of the problem we face to produce a regional development database. Of the 30 jurisdictions, only two (San Francisco, and Fremont) have a perceived data mining ease of *green* with the remainder of jurisdictions evenly split between *yellow, red, and null (no online data report)*. Given the difficult path ahead this memo explores the difficulties with *yellow* and *red* datasets, and potential approaches. Additionally, this memo explores potential opportunities to alter the profile to be more *green* than *red*.









This memo should be dynamic over the course of the next few weeks. The process moving forward will likely inform this discussion, with new challenges uncovered, or new opportunities presented. The memo is broken in four brief sections: (1) data challenges, (2) ABAG challenges, (3) Emerging Opportunities, (4) Next Steps.

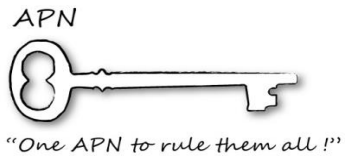
Data Challenges

When the number of jurisdictions are discussed, it is in reference to the 30 included in the study sample.

Only 7/30 jurisdictions have an archive of data records online. The remainder of the jurisdictions only have a single (most recent) data set available. Of the seven with an archive of records, half have a systematic record (i.e. every quarter) while the others have disparate records (i.e. two one year, zero the next). Of those without any archive section of the website, most had a report produced within the past three months. Most had a date the report was produced.

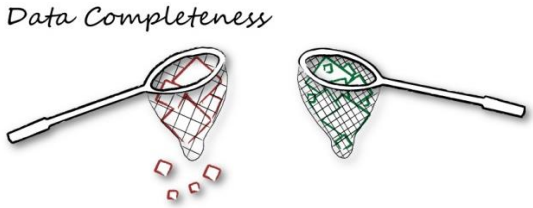
For jurisdictions that already report permitting activity they likely could produce past reports as part of a formal data request. Internally ABAG must determine how frequently archive records should be gathered (i.e. Can we capture everything we need with annual reports, or are their projects that are on the list one month, and off the list by the next?)

For our effort having information in a data form is critical to the efficiency and accuracy of data collection. Only four jurisdictions in the sample reported information in an excel file type. Another four jurisdictions display the information in an .html table. Methods exist to extract information from non-data formats; however, it increases the likelihood of errors, and is inefficient and unsustainable. 12 jurisdictions report their information in .pdf files, nine of which are in a tabular format, likely generated in a unique permit tracking software, or directly from excel. For these jurisdictions I see two options, (1) ABAG requests raw data be sent to ABAG, (2) ABAG requests the jurisdiction upload a data file as part of their public reporting procedure (in addition to or in replacement of current .pdf reports).



Of the 20 jurisdictions that report data, 12 don't list the property APN. Furthermore, four don't list a specific address. For sustainable, interoperable data, jurisdictions must include APN in their data tables. Using address for comparison is feasible, but would drastically slow the geotagging of the data, and more importantly would hinder the ability to compare archive data sets against one another.

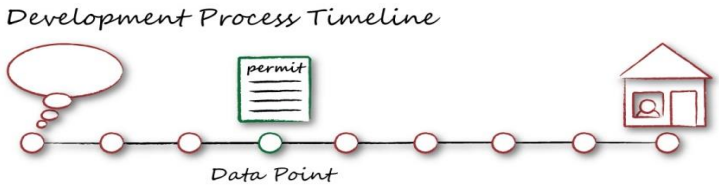
In developing simple scripts for aggregating data, APNs offer the most accurate referencing method. Additionally, if we develop small tools to aid in automating data updates and data cleaning, it will be key to only use a single attribute across all city data.



None of the jurisdictions clearly call out what their data is and what their data is not. There is a wide spectrum of data completeness across sampled jurisdictions. Some cities have a data set that includes all permits (i.e. reroofing, new garage, new development), while others only list hallmark projects (i.e. developments with >50 units).

It is not just scale of development that is reported unevenly across jurisdictions, but also the land use. Some jurisdictions only report residential use permits. It is unclear, especially in smaller jurisdictions, if the city has chosen to omit certain scales or uses from their permit records.

It will be important in labeling data, what the data represents. We need to know if zero single family homes were built in a community, or if the city does not report this permit type in their data set. Without an understanding of what is and is not included, the database cannot be used for drawing direct conclusions.

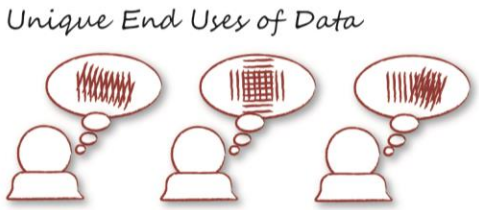


This section is important for the future of the development activity dashboard. For the dashboard to function it is important to include time-stamped milestones as development evolves from *inception* to an *occupied unit*.

Some jurisdictions have a status attribute for each permitted development. If these jurisdictions also have a successful data archive, it is possible to generate timestamps at the granularity of the report update frequency (for the timestamps recognized by the city). There are only a handful of cities that have both archived reports and a status attribute.

There are two methods that could be used to collect this data (1) have a single status attribute in the data and have archived records, or (2) have multiple status attributes that have dates associated with their progress. This method wouldn't require frequent archived records, but would increase the attributes in the data set.

ABAG Challenges



There are a lot of different users for this data. The good news is that there is overlap in what people want. Unfortunately, even small specifics of their needs can double the effort of collecting and managing data. It will be important to understand how this work can be leveraged by others without resulting in huge amounts of overlapping effort.

The best example of this is the relationship between UrbanSim and the recent generation of housing data. Housing users want high accuracy, regionally complete data, with information on affordability type. UrbanSim users want moderate accuracy, large developments only, all land uses (not just housing). The issue is that the data for both these users is likely to come from the same local jurisdictional source, but one process of data mining requires coarse data collection, while the other requires fine. Finding a way to do both without overlapping effort is the challenge.

Emerging Opportunities

In 2014 City of San Francisco hired their first Chief Data Officer, Joy Bonaguro. The position works out of the Mayor's office and works across all departments, attempting to make all their data *open*.

> *Open Data is data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and share alike. The dataset must allow for interoperability to ensure individuals can manipulate the data and work across datasets.*

In less than a year Joy has been able to change perspectives of Open Data within San Francisco Departments as well as build datasf.org. True to her charge she is also making her process open and is an advocate toward supporting similar efforts across the region. This not only includes their business strategy, but also legal implications and how to manage sensitive information. Her office will be a strong resource for cities interested in following their lead.

The Sunlight Foundation, created to standardize and advocate for open data, has created guidelines for organizations to follow when approaching the open data goal. The following webpage outlines 31 guidelines for an ideal open data platform http://sunlightfoundation.com/opendataguidelines/ This resource along with San Francisco's willingness to share their work, provide ABAG and local jurisdictions with the framework to approach open data.

Next Steps

These steps will run in parallel.

Dig into San Francisco & Fremont data sets

Using the two best known sets of data in the region I will play with the data and develop a sample database(s). This process will explore the importing of different city data into the same database. I will also reach out to users of the data (Mike Reilly, Pedro Galvao, Duane Bay) to determine if they can collect the information they need from the database without having to duplicate work.

Reach out to a handful of "yellow" cities

There are a number of cities who were labeled as yellow for only one or two reasons. With guidance from ABAG's research team I will determine which cities to approach with a formal data request. In most cases it will be a matter of asking for archived records, and asking for a raw data sets rather than a .pdf file. My short list of cities will be ones I believe already have the existing files we want, but simply have not made the data available online. Obtained city data sets will be included in the above database exploration.

Section II: Key Questions: Attributes and Approaches

Summary:

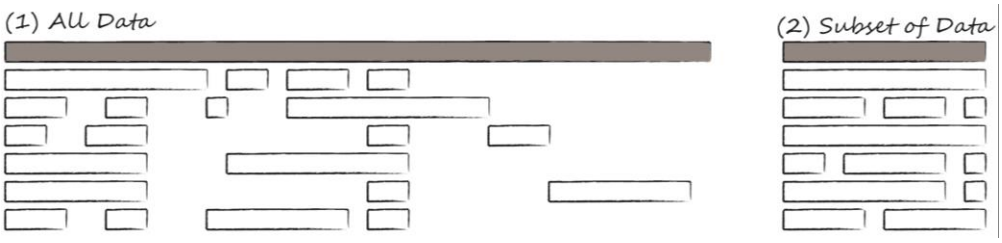
I've outlined three key questions that should be answered before moving forward. These questions are introduced in points 1, 2, and 3 in this memo. Question 1 explores which attributes to collect. It is difficult because it is a question of tradeoffs. Question 2 explores how to take city vernacular and make it consistent in the database. I don't envision it to be difficult, but it will require a high quality of documentation. Question 3 explores what software we'll use to hold, organize, and manipulate the database. This question will likely require technical staff at ABAG and MTC to meet and come to agreement on what solution is best.

I've also outlined two points, 4 and 5, that I believe are useful in developing the process for reque sting data from cities. Point 4 illustrates that even if we start by only collecting data from a few large jurisdictions the database will cover a large percentage of the population. The data in Point 5 suggests that a number of cities (both by jurisdiction count, and by population) use software to track permitting. Accela and CRW provide the majority of contracts in the region. This may make data requests easy for cities, and will provide ABAG staff with data we can manipulate.

1.0 Which data attributes make up this database?

[*Note: For brevity, this discussion has been removed from the memo as it is largely redundant with material presented in Appendix A-1.]*

1.1 Do we collect it all, or do we thin down to consistent values?



There are two options

(1) We collect all development database attributes that cities collect data for.
(2) We only collect a defined set of attributes.

I'm typically not a fan of deleting raw data (storage is cheap, and there are some applications that could someday be useful), however, I believe a streamlined database can lead to fewer errors in database management, and more importantly fewer errors when those unfamiliar with the data set use the data for analysis purposes. A middle ground may be the best route forward.

1.2 Agree on how timestamp will be generated.

One product for the data set will be to track the stage various planned developments are in, and how long it takes for them to go from inception to occupied. I see two options:

(1) Ask local jurisdictions to track this information.

Pro: The dates a development moves from one stage to the next will be exact.

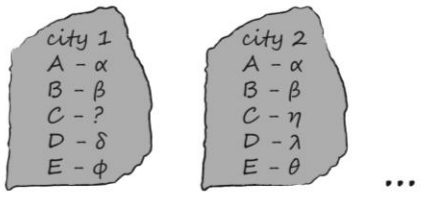
Con: I expect the majority of jurisdictions don't currently collect this data value. The data set that is generated will have a number of null values for jurisdictions.

(2) ABAG automates a process to generate the data with each database update (quarterly – recommended)

Pro: If jurisdictions list the developments stage, a date can be applied resulting in a regionally complete set of data.

Con: The granularity of the data set will be quarterly (i.e. if development moves through multiple steps in a month, steps will be missed). This will require a simple visual basic code for a limited program like excel. Another database software may provide a simple solution (see 3.0).

2.0 Develop a Rosetta Stone for data



Determine a single operating language for data and agree on how to translate each jurisdictions own data. Automated translators will need to be created for each jurisdiction. Every time data is brought in from a jurisdiction the data will need to go through a cleaning process, that should become nearly 100% automated. The process should be designed to take raw city data, and places it into a consistent ABAG format in less than ten minutes. This is possible. Each city will have a procedure written up, to insure consistency over time, and as a reference when using the data.

The graphic below shows the many different ways a project status is tracked in different cities. The same is true when jurisdictions name addresses, zoning, or even how they break apart building types (industrial, institutional, medical, retail, commercial). I believe the status designation will be the hardest field to untangle. The others for the most part seem straight forward.

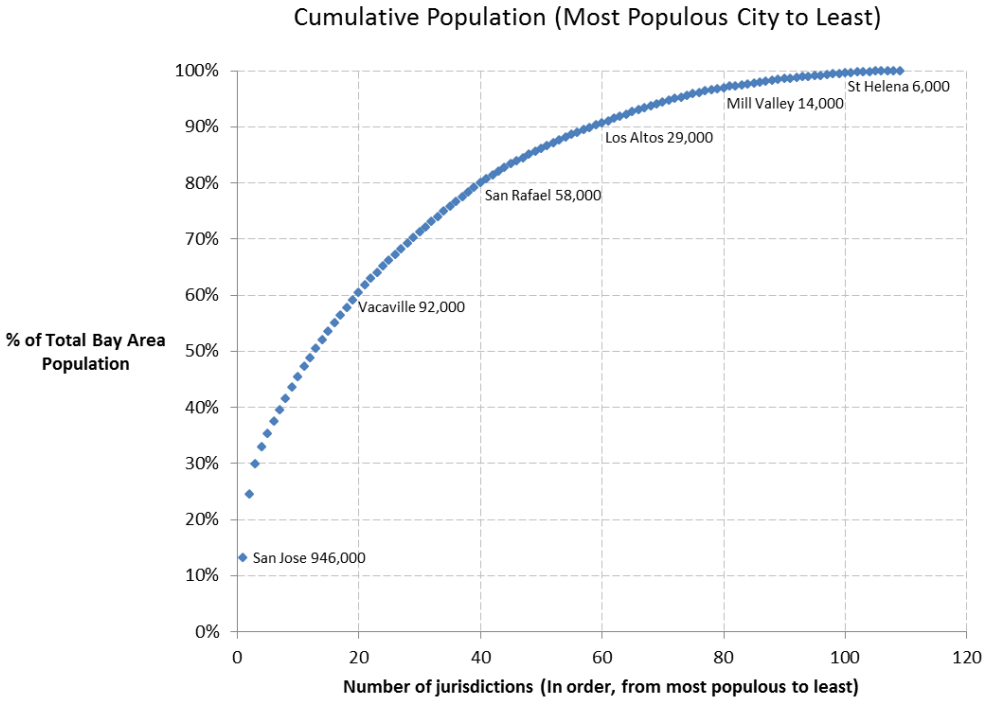
Reduced Categories		Duanes Categories	San Francisco Status Fields	Fremont Status Fields	Sunnyvale Planning Permit Status Fields	Concord Current Projects Report	Saratoga Permits Issued
		Rumored					
Early Discussions	0 / 5	Designated					
		Zoned					
Project Proposed	3 / 5	Proposed	PL Filed (13%)		Pending Review (22%)	Proposed	
Review	2 / 5	Staff Review		PRP - Preliminary Review Proce	Comments Provided (18%)		
		Public Review		OPC - Open for Public Comment			
Approved &/or Entitled	4 / 5	Approved	PL Approved (8%)		Approved (58%)	Approved	
		Entitled		APV - Entitlement Approved (18			
Permitted	3 / 5	Permitted	BP Filed (37%)	BPR - Building Permit Review (1			Applied
			BP ISSUED 129 (13%)				Issued
			BP REINSTATED (2%)				Finaled
			BP APPROVED (3%)				
Construction	3 / 5	Construction	CONSTRUCTION (22%)	UC - Under Construction (25%)		Under Construction	
		Built					
Built	0 / 5	Occupied					
		Fully Occupied					

3.0 Determine the Database Backend

This is an area I defer all recommendations to ABAG staff in the Research and IT teams. Based on several conversations, I believe my recommendations are well within ABAG's technical capacity.

4.0 Determine Breadth of Data Request

Number of Jurisdictions: I recommend an abbreviated request for at least the first quarterly database report. Based on staff resources that could be committed to this effort, goals could be set to cover 5%, 33%, 50%, or 60% of all development; or of all jurisdictions. These values can change with each update, depending on resources. By population, a large percentage of the region can be collected with a small percentage of the region's jurisdictions



4.1 How far back in time: At the start there was a desire to grab data from 2010 onwards. This will be possible for some jurisdictions, but it should be clear when we reach out to jurisdictions

that there is a current request to get backlogged data, and then there will be subsequent quarterly requests asking for only current data.

4.2 An iterative update process: The first time a jurisdiction is added to the database it will require additional resources. We need to find the right person within the city to make the request to, that person then has to (for the first time) assemble the data, and then send to ABAG. ABAG then has to figure out how to uniquely clean that city's data. For example, if ABAG can only commit 60 hrs to the first update, we may only reach ten jurisdictions. At the next update however if 60 hrs is committed again, it may only take 20 hrs to get updates from the previous ten jurisdictions, allowing for the expansion of the database to additional jurisdictions, increasing the reach of the database with each update.

5.0 Optimism! Working with Tracking Software Vendors

A large percent of the Bay Area uses one of two software packages to meet their permit tracking needs. Accela and CRW Trackit account for at least 40%. The data summarized below was pulled from an ABAG data request to jurisdictions in 2014. Anecdotal reports suggest that the jurisdictions that are reported as "null" (i.e., ABAG wasn't able to determine tracking software) have a similar pattern of use as those reporting. In other words, if 40% out of 73% reporting used Accela or CRW Trackit, we can extrapolate that perhaps 53% of all jurisdictions use one of these two systems. [Note: See graphics summarizing this data at the beginning of this appendix.]

I've reached out to both Accela and CRW, both of which have been reasonably helpful, and have mentioned we can come back to them with further questions / requests. Accela will send us a list of all their clients any day, and CRW is mulling over if they can/should share that information with us. A big piece of the pie is San Jose. By population, San Jose represents 13% of the region. The city uses a unique software, Amanda, a CSDC product. Amanda is more popular in the policing and justice systems for case management, and should be a software that can export appropriate information to ABAG if set up.

All of this is to say that a large percent of the Bay Area (by population) is collecting permitting data in database software. Both CRW and Accela believe jurisdictions should be able to very easily export their databases and send us raw data files. Uncovering this made me optimistic that this project may be sustainable, and may cover a significant cross section of the Bay Area.

Appendix F: Acknowledgements

This report was researched and written by Association of Bay Area Governments (ABAG), and funded by a Sustainable Communities Partnership grant from the U.S. Department of Housing and Urban Development (HUD) to Metropolitan Transportation Commission (MTC) and ABAG to administer on behalf of the Regional Prosperity Plan Consortium. ABAG thanks the project's numerous stakeholders and participants, especially the Housing Working Group, Joint Project Team and Steering Committee of the Consortium and the following individual contributors.

Gillian Adams, ABAG
Duane Bay, ABAG, *Project Lead*
Wayne Chen, City of San Jose
Miriam Chion, ABAG
Peter Cohen, Council of Community Housing Organizations
Kara Douglass, Contra Costa County
Pedro Galvao, ABAG
Adam Garcia, Greenbelt Alliance
Michael Germeraad, ABAG
Joshua Hugg, Housing Leadership Council
Johnny Jaramillo, ABAG
Doug Johnson, MTC
Brian Kirking, ABAG
Cynthia Kroll, ABAG
Sailija Kurrella, ABAG
Pilar Lorenzana-Campo, Nonprofit Housing Association
Jason Munkres, ABAG
Aksel Olsen, ABAG
James Pappas, California Housing Partnership Corporation
Carlos Romero, Urban Habitat
Mark Shorett, ABAG
Michael Smith, ABAG
Evelyn Stivers, Urban Habitat
Vikrant Sood, MTC
Bryan Tyler, ABAG
Matt Vander Sluis, Greenbelt Alliance
Elizabeth Wampler, San Francisco Foundation
Hing Wong, ABAG

Appendix G: Linking Memo to Contract Requirements

The form and content of the deliverable products for Task 2 of the Funding Agreement, Development Tracking Dashboard have evolved during the contract period. Nonetheless, this comprehensive memo, which comprises the final deliverable, includes all of the deliverables that were originally specified. The purpose of this appendix is to demonstrate explicitly the linkage and coverage between the requirements as stated in the scope of work and the contents of this memo

As stated above, the form and content of the deliverable products evolved during execution of the project. This evolution was due to early findings related to feasibility—what aspects of developing the desired system were determined to be more difficult and less difficult—and input from Regional Prosperity Consortium stakeholders in the Housing Working Group. In the grant application and the scope of work, great emphasis was placed on creating a *template* of a real estate development pipeline report. It was imagined that one of the most challenging and time-consuming tasks in designing such a template would be negotiating among multiple stakeholders exactly what data to report (i.e., data attributes and data definitions). As it turned out, general agreement on a desirable template was relatively easy: literally, "like San Francisco's *Development Pipeline Report*." Appendix B is a sample quarterly pipeline report.

As to form, San Francisco's development pipeline report is exemplary. As to data content, the San Francisco pipeline report is also exemplary, with two important caveats: (1) a multi-purpose, region-wide report will probably have to be restricted to the subset of desired data that is widely available from almost all jurisdictions; and (2) we aspire to collect and display pre-permit data (status of projects during the approval process) as well as the post-permit data that is more readily available.

Under a literal interpretation of the scope of work, Appendix B fulfills the requirements of Deliverable 2.3a of the Development Tracking Dashboard task, "Housing Development Tracking Template."

If the degree of challenge to describe a pipeline report template was overestimated, the degree of challenge to develop the underlying technical and procedural capacities necessary to generate a region-wide report comparable to the San Francisco report were grossly underestimated. Therefore, to honor the spirit and intent of the scope of work, a feasibility analysis for a comprehensive system and a proposal for an approach to developing a comprehensive system—in other words, this memo as a whole—became, functionally, the primary deliverable of the contract.

While Appendix B fulfills the literal requirements of the contract, this memo as a whole fulfills the substantive requirements.

Two other deliverables were mentioned in the project narrative, although not called out explicitly in the itemized scope of work. The first was a preliminary list of development tracking indicators [attributes], including at minimum, the number, location and estimated affordability level of all housing in the development pipeline. These particular attributes are so basic that they are mentioned in several places in the memo, most prominently in the high-level descriptions of the Housing Developments and Adoption/Approval Milestones datasets listed in the Proposed Datasets table in the Proposed Gross Data Structure section of the memo. The second was an assessment of the feasibility of collecting the necessary data from available sources, including cost, frequency of data updates, and data limitations. Appendix E, *Feasibility of Generating a Regionwide Development Activity Database*, fulfills this requirement.